

CHAPITRE

V

VARIABLES ALÉATOIRES DISCRÈTES ET ESPÉRANCE

Sommaire

A	Rappels sur les espaces probabilisés	2
A.1	Le cadre théorique	2
A.2	Conditionnement et indépendance	4
A.3	Résultats fondamentaux	5
A.4	Problématiques et savoir-faire fondamentaux	7
B	Notion de variable aléatoire réelle	17
B.1	Définitions et premières propriétés	17
B.2	Le cas des variables aléatoires discrètes	21
B.3	Lois discrètes usuelles	24
C	Espérance et variance d'une variable aléatoire discrète	29
C.1	Espérance et formule de transfert	29
C.2	Moments et variance	31
C.3	Espérance conditionnelle	34

A - Rappels sur les espaces probabilisés

A.1 - Le cadre théorique

Tout d'abord, on rappelle qu'un ensemble E est *dénombrable* lorsqu'il existe une bijection entre $\mathbb{N}$ et E ce qui revient à dire que l'ensemble E peut s'écrire $E = \{x_n ; n \in \mathbb{N}\}$.

Par exemple, les ensembles $\mathbb{Z}$, $\mathbb{N} \times \mathbb{N}$ et $\mathbb{Q}$, par exemple, sont dénombrables. En revanche, $\mathbb{R}$ n'est pas dénombrable. Une partie de $\mathbb{N}$ est soit finie, soit dénombrable.

⇒ Ensemble des événements

On considère un ensemble non vide Ω .

On appelle *ensemble des événements* de Ω tout sous-ensemble $\mathcal{A}$ de $\mathcal{P}(\Omega)$ tel que :

- i. $\Omega \in \mathcal{A}$;
- ii. si $A \in \mathcal{A}$ alors $\overline{A} \in \mathcal{A}$;
- iii. pour toute famille finie ou dénombrable $(A_k)_{k \in I}$ dans $\mathcal{A}$, on a $\bigcup_{k \in I} A_k \in \mathcal{A}$.

Les éléments de $\mathcal{A}$ sont appelés les *événements* de Ω et le couple $(\Omega, \mathcal{A})$ est appelé un *espace probabilisable*.

Remarques

- 1> Un tel ensemble de parties porte le nom de *tribu* ou de σ -*algèbre* mais cette terminologie n'est pas au programme d'ECG.
- 2> Notons que si $(A_k)_{k \in I}$ est une famille (finie ou) dénombrable de parties de Ω alors :

$$\overline{\bigcup_{k \in I} A_k} = \bigcap_{k \in I} \overline{A_k} \text{ et } \overline{\bigcap_{k \in I} A_k} = \bigcup_{k \in I} \overline{A_k},$$

et pour toute partie B de Ω :

$$B \cap \bigcup_{k \in I} A_k = \bigcup_{k \in I} (B \cap A_k) \text{ et } B \cup \bigcap_{k \in I} A_k = \bigcap_{k \in I} (B \cup A_k).$$

On en déduit que si $(A_k)_{k \in I}$ est une famille finie ou dénombrable d'éléments de $\mathcal{A}$ alors $\bigcap_{k \in I} A_k \in \mathcal{A}$.

- 3> Un *système complet d'événements* est une famille finie ou dénombrable $(A_k)_{k \in I}$ d'événements non vides et deux à deux incompatibles dont la réunion fait Ω :

- $\forall k \in I, A_k \neq \emptyset$ et $A_k \in \mathcal{A}$;
- $\forall (k, \ell) \in I^2, (k \neq \ell \Rightarrow A_k \cap A_\ell = \emptyset)$;
- $\bigcup_{k \in I} A_k = \Omega$.

Notons que si $(A_k)_{k \in I}$ est un système complet d'événements et si $B \in \mathcal{A}$, alors : $B = \bigcup_{k \in I} (B \cap A_k)$.

Exemples

- 1 ▶ Pour tout événement A non vide de Ω avec $A \neq \Omega$, $\{A, \bar{A}\}$ est un système complet d'événements (fini) de Ω .
- 2 ▶ Si Ω est fini ou dénombrable alors un système complet d'événements est donné par les $\{\omega\}$ pour tous les ω de Ω .

⇒ Probabilité

On considère un espace probabilisable $(\Omega, \mathcal{A})$.

On appelle **probabilité** sur $(\Omega, \mathcal{A})$, toute application $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$ vérifiant :

- i. $\mathbb{P}(\Omega) = 1$;
- ii. $\mathbb{P}$ est σ -additive i.e. pour toute famille finie ou dénombrable $(A_i)_{i \in I}$ d'événements deux à deux incompatibles, on a :

$$\mathbb{P}\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mathbb{P}(A_i) \quad (\sigma\text{-additivité}).$$

Pour tout événement A , le réel positif $\mathbb{P}(A)$ s'appelle la **probabilité de l'événement** A .

On dit que le triplet $(\Omega, \mathcal{A}, \mathbb{P})$ est un *espace probabilisé*.

Un événement de probabilité 1 est dit *presque sûr*.

Un événement de probabilité 0 est dit *négligeable*.

Remarques

- 1 ▶ Notons que (dans le cas où I n'est pas fini) la somme $\sum_{i \in I} \mathbb{P}(A_i)$ est sans ambiguïté. En effet, la convergence est assurée par le fait qu'il s'agisse de la probabilité d'un événement (la réunion des A_i) et tous les termes étant positifs, il s'agit d'une convergence absolue donc la somme ne dépend pas de l'ordre de sommation.
- 2 ▶ Pour tous événements A et B , on a :
- $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$;
 - $\mathbb{P}(\emptyset) = 0$;
 - Si $B \subset A$ alors $\mathbb{P}(B) \leq \mathbb{P}(A)$ (*croissance*) ;
 - $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
- 3 ▶ Un **système quasi-complet d'événements** est une famille finie ou dénombrable $(A_k)_{k \in I}$ d'événements non vides et deux à deux incompatibles dont la probabilité de la réunion est 1 :
- $\forall k \in I, A_k \neq \emptyset$ et $A_k \in \mathcal{A}$;
 - $\forall (k, \ell) \in I^2, (k \neq \ell \Rightarrow A_k \cap A_\ell = \emptyset)$;
 - $\mathbb{P}\left(\bigcup_{k \in I} A_k\right) = 1$.
- Si $(A_k)_{k \in I}$ est un système quasi-complet d'événements mais n'est pas un système complet d'événements alors $\bigcup_{k \in I} A_k \neq \Omega$ mais en ajoutant à cette famille l'événement $\bigcup_{k \in I} \bar{A}_k$, on obtient un système complet d'événements.

A.2 - Conditionnement et indépendance

⇒ Probabilité conditionnelle

Soit A un événement de probabilité non nulle et B un événement quelconque.

On définit **la probabilité de B sachant A**, notée $\mathbb{P}(B|A)$ ou $\mathbb{P}_A(B)$, par :

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

Soit A un événement de probabilité non nulle. L'application :

$$\mathbb{P}_A : \mathcal{A} \rightarrow \mathbb{R}, B \mapsto \mathbb{P}(B|A)$$

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

Remarques

- 1 ▶ Si A est un événement non négligeable, alors $(\Omega, \mathcal{A}, \mathbb{P}_A)$ est un espace probabilisé.
- 2 ▶ La définition donne une première version de la formule de Bayes.

Si $\mathbb{P}(A) \neq 0$ et $\mathbb{P}(B) \neq 0$, alors :

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

⇒ Indépendance

- 1 ▶ Deux événements A et B sont dits **indépendants** lorsque :

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

ce qui revient à dire que $\mathbb{P}(B) = \mathbb{P}_A(B)$ ou $\mathbb{P}(A) = \mathbb{P}_B(A)$.

- 2 ▶ On dit que des événements $A_1, \dots, A_{n-1}$ et A_n sont **deux à deux indépendants** lorsque :

$$\forall (i, j) \in \llbracket 1, n \rrbracket, i \neq j \implies A_i \text{ et } A_j \text{ sont indépendants};$$

- 3 ▶ On dit que des événements $A_1, \dots, A_{n-1}$ et A_n sont **mutuellement indépendants** lorsque, pour tout sous-ensemble I de $\llbracket 1, n \rrbracket$, on a :

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

Exemple - cas particulier

Trois événements A, B et C sont deux à deux indépendants si :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B), \mathbb{P}(B \cap C) = \mathbb{P}(B) \cdot \mathbb{P}(C) \text{ et } \mathbb{P}(C \cap A) = \mathbb{P}(C) \cdot \mathbb{P}(A);$$

pour l'indépendance mutuelle, on ajoute cette quatrième égalité :

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

A.3 - Résultats fondamentaux

Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements d'un espace probabilisé, on dit que :

- $(A_n)_{n \in \mathbb{N}}$ est **croissante** si : $\forall n \in \mathbb{N}, A_n \subset A_{n+1}$ (i.e. si A_n est réalisé alors A_{n+1} l'est également) ;
- $(A_n)_{n \in \mathbb{N}}$ est **décroissante** si : $\forall n \in \mathbb{N}, A_{n+1} \subset A_n$ (i.e. si A_{n+1} est réalisé alors A_n l'est également).

⇒ Théorème de la limite monotone

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

1. Si la suite $(A_n)_{n \in \mathbb{N}}$ est croissante alors :

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

2. Si la suite $(A_n)_{n \in \mathbb{N}}$ est décroissante alors :

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

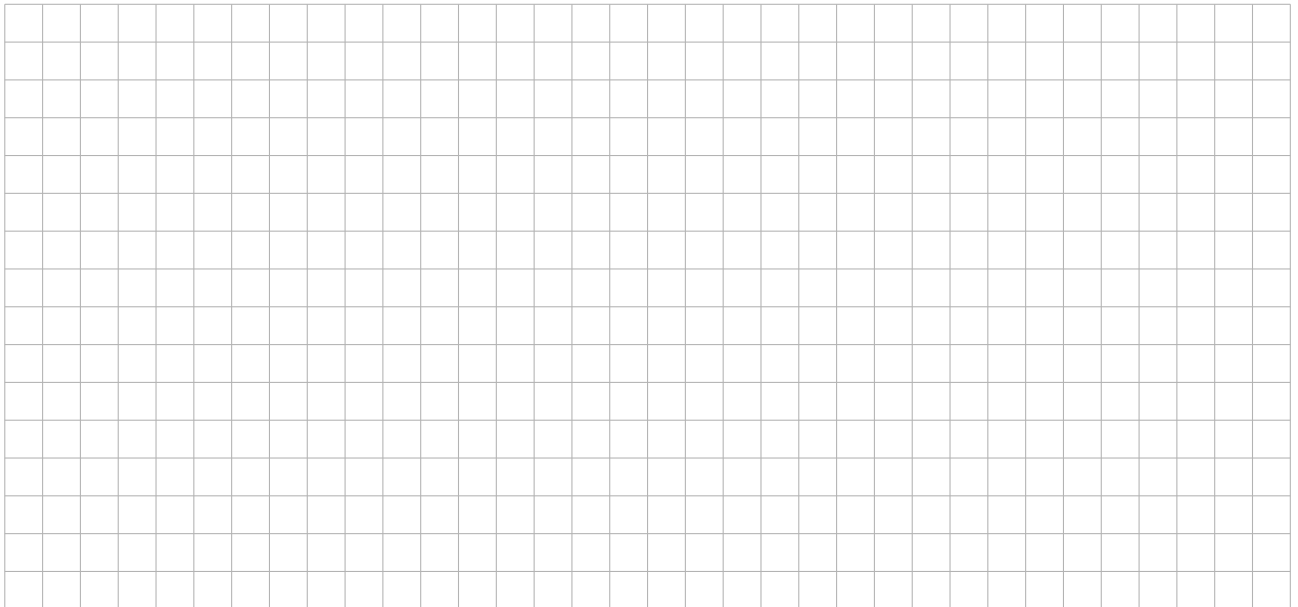
Notons qu'il n'y a aucune perte de généralité à considérer, dans un cas comme dans l'autre, que la réunion ou l'intersection «commence» à partir d'un rang non nul (i.e. pour $n \geq n_0$).

Exemple

On lance un dé équilibré à 6 faces (en supposant les lancers indépendants) tant que l'on n'obtient pas le chiffre 6. On suppose qu'il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ permettant de modéliser cette expérience.

Montrons que l'expérience s'arrête presque sûrement (c'est-à-dire que l'on finit par obtenir 6 presque sûrement).

Pour tout $n \in \mathbb{N}$, on note A_n l'événement «on lance le dé au moins $n+1$ fois» (ce qui signifie que l'on n'a pas obtenu 6 aux n premiers lancers).



Remarque

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(A_n)_{n \in \mathbb{N}}$ une suite d'événements de Ω , alors :

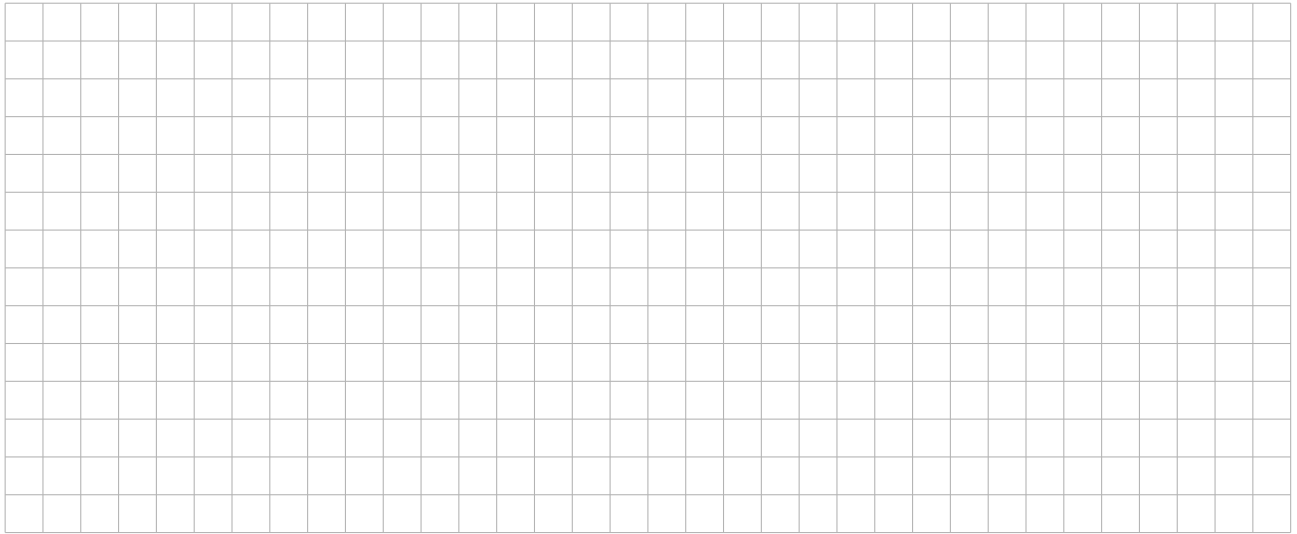
$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{N \rightarrow +\infty} \mathbb{P}\left(\bigcup_{n=0}^N A_n\right) \text{ et } \mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{N \rightarrow +\infty} \mathbb{P}\left(\bigcap_{n=0}^N A_n\right).$$

Exemples

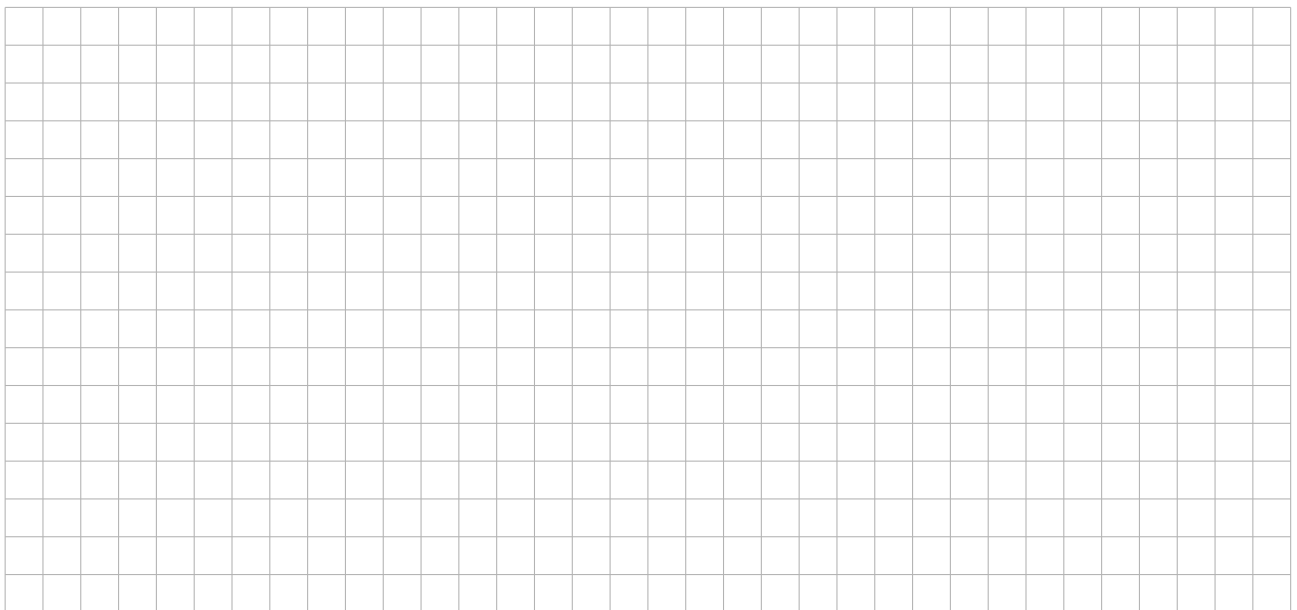
- 1 ▶ On dispose d'une pièce donnant *face* avec une probabilité $p \in]0, 1[$, on la lance une infinité de fois en supposant les lancers indépendants et on suppose qu'il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ permettant de modéliser cette expérience.

Montrons que l'événement «on n'obtient que des *faces*» est négligeable (autrement dit, on obtient au moins une fois *pile* presque sûrement).

Pour tout $n \in \mathbb{N}^*$, on note F_n l'événement «on obtient *face* au n -ième lancer».



- 2 ▶ Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Montrer qu'une union dénombrable d'événements négligeables est également un événement négligeable puis en déduire qu'une intersection dénombrable d'événements presque-sûrs est également un événement presque-sûr.



⇒ Formule des probabilités composées

Soit $A_1, A_2, \dots, A_{n-1}$ et A_n des événements d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ tels que :

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_{n-1}) \neq 0,$$

alors :

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(A_1) \cdot \mathbb{P}_{A_1}(A_2) \cdot \mathbb{P}_{A_1 \cap A_2}(A_3) \dots \mathbb{P}_{A_1 \cap \dots \cap A_{n-1}}(A_n).$$

⇒ Formule des probabilités totales

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé et $(A_n)_{n \in I}$ un système complet d'événements non négligeables (I est un ensemble fini ou dénombrable). Pour tout événement B :

$$\mathbb{P}(B) = \sum_{n \in I} \mathbb{P}(A_n \cap B) = \sum_{n \in I} \mathbb{P}(A_n) \mathbb{P}_{A_n}(B).$$

A.4 - Problématiques et savoir-faire fondamentaux

■ Comment mettre en place le bon cadre de travail ?

SF132 Reconnaître une éventuelle équiprobabilité

Raisonnement avec une probabilité uniforme offre le cadre de travail le plus confortable pour les calculs puisque ces derniers se ramènent à des situations de dénombrement. Il est donc impératif de reconnaître de telles situations quand elles se présentent.

Dans un énoncé, l'équiprobabilité se cache derrière des expressions telles que « pièce équilibrée », « dé non pipé », « boules indiscernables au toucher »...

Exemple

Quel est l'univers associé à l'expérience consistant à lancer deux dés usuels non pipés afin d'en considérer la somme ?

Il convient de bien choisir l'univers. Ce sont tous les lancers différents qui sont équiprobables, et non pas les sommes des deux dés (pour s'en convaincre, on se rappellera les parties de jeux de société où le 12 sortait rarement).

On peut donc travailler avec la probabilité uniforme en choisissant l'univers Ω égal à $\llbracket 1, 6 \rrbracket^2$, mais pas avec $\Omega = \llbracket 2, 12 \rrbracket$.

Une probabilité uniforme peut apparaître, non pas comme la probabilité de l'espace probabilisé étudié, mais sous forme d'une probabilité conditionnelle.

Exemple

Reconnaître une probabilité uniforme dans l'expérience suivante. On dispose de deux dés à six faces – un non pipé et un truqué qui donne 6 avec une probabilité de $\frac{1}{2}$. L'expérience consiste à choisir un des deux dés au hasard, puis à la lancer.

L'univers de cette expérience est $\llbracket 1, 6 \rrbracket$ mais la probabilité $\mathbb{P}$ associée n'est pas la probabilité uniforme : à cause du dé pipé, les valeurs 1 et 6 n'ont pas les mêmes fréquences d'apparition.

~ Cependant, sachant réalisé l'événement A : « le dé non pipé est choisi », toutes les valeurs de 1 à 6 ont alors la même probabilité de sortir. Autrement dit, la probabilité $\mathbb{P}_A$ est uniforme.

SF133 Construire « la bonne » probabilité

Définir une probabilité consiste à donner une application $\mathbb{P}$ définie sur la tribu choisie de l'espace probabilisable vers $[0, 1]$.

Si Ω est ne serait-ce que de cardinal 10 alors la tribu $\mathbb{P}(\Omega)$ contient $2^{10} = 1024$ éléments : il peut donc être vite fastidieux de vouloir construire à la main une application sur un tel ensemble...

Heureusement, sur un univers au plus dénombrable, la situation se simplifie grandement : il suffit d'étudier sur les événements élémentaires, c'est-à-dire sur les singletons de $\mathbb{P}(\Omega)$; il suffit de vérifier que les probabilités élémentaires associées sont toutes positives et que leur somme (éventuellement infinie) vaut 1 pour être certain d'avoir un cadre de travail correct.

Le calcul de la probabilité d'un événement quelconque passe alors par la décomposition de celui-ci en une union disjointe en événements élémentaires.

Exemple

~ On lance une fois un dé truqué dont la probabilité de chaque face est proportionnelle au numéro inscrit sur cette face. Calculer la probabilité d'obtenir un nombre impair ?

~ L'univers Ω est $\llbracket 1, 6 \rrbracket$. Pour i compris entre 1 et 6, on note p_i la probabilité d'obtenir le numéro i . On introduit aussi λ le coefficient de proportionnalité dont il est question dans l'énoncé. Ainsi :

$$\forall i \in \llbracket 1, 6 \rrbracket, p_i = \lambda \times i.$$

~ On remarque immédiatement que λ est nécessairement positif ; par ailleurs :

$$\begin{aligned} \sum_{k=1}^6 p_k = 1 &\iff \lambda \sum_{k=1}^6 k = 1 \\ &\iff 21\lambda = 1 \\ &\iff \lambda = \frac{1}{21}. \end{aligned}$$

~ On connaît donc maintenant toutes les probabilités p_i , l'espace probabilisé est donc bien défini.

~ En notant I l'événement « obtenir un nombre impair », on trouve :

$$\begin{aligned} \mathbb{P}(I) &= \mathbb{P}(\{1; 3; 5\}) \\ &= p_1 + p_3 + p_5 \\ &= \frac{3}{7}. \end{aligned}$$

SF134 Repérer l'indépendance entre expériences

Lorsque l'expérience se décompose en plusieurs étapes où l'on répète exactement le même processus, on peut faire l'hypothèse que ces « sous-expériences » sont indépendantes : par exemple, des lancers de dés successifs ou des tirages avec remise dans une urne...

Cette indépendance entre expériences va se traduire, sans autre vérification, par une indépendance entre événements : plus précisément, si n événements correspondent chacun à n étapes supposées indépendantes de l'expérience, alors ces n événements seront supposés mutuellement indépendants.

Exemple

On lance n fois de suite un dé équilibré. Quelle est la probabilité d'obtenir le même numéro à tous les lancers ?

Pour i compris entre 1 et 6 et j compris entre 1 et n , on introduit l'événement $A_{i,j}$ défini par : « obtenir le numéro i au j -ème lancer ».

Ainsi, l'événement M « obtenir le même numéro à tous les lancers » est :

$$M = \bigcup_{i=1}^6 \left(\bigcap_{j=1}^n A_{i,j} \right).$$

Étant donné qu'un lancer ne peut donner qu'un seul numéro, l'union ci-dessus est disjointe. Donc :

$$\mathbb{P}(M) = \sum_{i=1}^6 \mathbb{P} \left(\bigcap_{j=1}^n A_{i,j} \right).$$

De plus, on peut supposer que les lancers sont tous indépendants ; donc pour i fixé, la famille d'événements $(A_{i,j})_{j \in \llbracket 1, n \rrbracket}$ est mutuellement indépendante. D'où :

$$\mathbb{P}(M) = \sum_{i=1}^6 \prod_{j=1}^n \mathbb{P}(A_{i,j}) = \sum_{i=1}^6 \prod_{j=1}^n \left(\frac{1}{6} \right) = \left(\frac{1}{6} \right)^{n-1}.$$

L'indépendance entre « sous-expériences » peut apparaître seulement après avoir conditionné par la réalisation d'un événement précis.

Exemple

Reconnaître une situation d'indépendance dans l'expérience suivante. On dispose de $n+1$ urnes $U_0, U_1, \dots, U_n$ et on suppose que, pour chaque $i \in \llbracket 0, n \rrbracket$, l'urne U_i contient i boules jaunes et $n-i$ boules vertes. On choisit une urne au hasard, puis on effectue deux tirages successifs avec remise dans cette urne.

Les différents tirages ne sont *a priori* pas indépendants. En effet, lors du premier tirage, si on obtient une boule verte alors on peut imaginer que l'urne choisie en contient une grande proportion et donc qu'il y a plus de chance que le second tirage fasse lui-aussi apparaître une boule verte.

Cependant, si l'on suppose réalisé l'événement « l'urne k est choisie », alors tous les tirages peuvent être supposés indépendants : connaître la couleur de la boule obtenue au premier tirage ne change pas les probabilités pour les tirages suivants.

■ Comment travailler avec l'équiprobabilité ?

Le calcul des probabilités dans le cadre de l'équiprobabilité revient à dénombrer des ensembles finis ; plus précisément, il s'agit de dénombrer l'univers ainsi que les événements qui nous intéressent.

On rappelle que deux ensembles finis ont même cardinal si, et seulement si, il existe une bijection entre eux. Par conséquent, si l'on fait le lien avec une situation connue, alors on trouvera le cardinal de l'événement recherché.

SF135 Exploiter les configurations usuelles de dénombrement

Il y a trois situations très classiques sur lesquelles on retombe souvent en dénombrement et pour lesquelles il faut impérativement connaître les résultats associés.

◇ *Tirage avec remise de p boules dans une urne qui en contient n :*

les résultats d'un tel tirage sont des p -listes et il y en a n^p .

◇ *Tirage sans remise de p boules dans une urne qui en contient n :*

les résultats d'un tel tirage sont des p -listes sans répétition et, si $p \leq n$, il y en a $A_n^p = \frac{n!}{(n-p)!}$; sinon il n'y en a pas.

Le nombre de permutations d'un ensemble à n éléments est un cas particulier de cette situation : il y a $n!$ permutations différentes.

◇ *Tirage simultané de p boules dans une urne qui en contient n :*

les résultats d'un tel tirage sont des combinaisons de p éléments parmi n et, si $p \leq n$, il y en a $\binom{n}{p} = \frac{n!}{p!(n-p)!}$; sinon il n'y en a pas.

En conséquence, si E est un ensemble fini de cardinal n , alors $\mathcal{P}(E)$ est de cardinal 2^n .

Exemple

Un enfant dispose de 10 crayons de couleurs différents et veut colorier un dessin composé de 7 zones sans utiliser deux fois la même couleur. Combien de coloriages différents sont possibles ?

Cette situation correspond à un tirage sans remise : pour chaque zone, l'enfant pioche dans sa trousse un crayon sans le remettre une fois le remplissage de la zone terminé. Il y a donc $A_{10}^7 = \frac{10!}{3!} = 604800$ coloriages possibles.

■ Comment calculer une probabilité ?

SF136 Décomposer l'événement par disjonction de cas

D'un point de vue ensembliste, il s'agit d'écrire l'événement sous la forme d'une réunion.

• Si les différents cas sont incompatibles, alors la probabilité de l'union va directement s'écrire comme une somme.

Exemple

Un logiciel informatique permet de tirer un nombre entier naturel au hasard de manière aléatoire. Pour $i \in \mathbb{N}$, la probabilité de l'événement N_i : « obtenir le nombre i » vaut $\left(\frac{1}{2}\right)^{i+1}$.

On lance une fois ce programme. Calculer la probabilité d'obtenir un nombre impair ?

Pour réaliser l'événement I : « obtenir un résultat impair », le logiciel peut renvoyer le nombre 1 ou renvoyer le nombre 3 ou renvoyer le nombre 5 et ainsi de suite... Ainsi :

$$I = \bigcup_{k=0}^{+\infty} N_{2k+1}.$$

Comme les événements de cette union sont deux à deux incompatibles, on a :

$$\mathbb{P}(I) = \sum_{k=0}^{+\infty} \mathbb{P}(N_{2k+1}) = \sum_{k=0}^{+\infty} \left(\frac{1}{2}\right)^{2k+2} = \frac{1}{4} \sum_{k=0}^{+\infty} \left(\frac{1}{4}\right)^k = \frac{1}{3}.$$

• les différents cas de l'alternative ne s'excluent pas mutuellement, alors il convient d'exploiter la formule du crible de Poincaré.

Exemple

On place six boules au hasard dans 4 compartiments, quelle est la probabilité qu'au moins un des compartiments soit vide ?

Pour i allant de 1 à 4, on note V_i l'événement « le compartiment i est vide ». On recherche donc la probabilité de l'événement $V_1 \cup V_2 \cup V_3 \cup V_4$. Ces quatre événements ne sont pas deux à deux incompatibles (il est possible que les compartiments 1 et 2 soient vides en même temps). Ainsi :

$$\begin{aligned} \mathbb{P}(V_1 \cup V_2 \cup V_3 \cup V_4) &= \mathbb{P}(V_1) + \mathbb{P}(V_2) + \mathbb{P}(V_3) + \mathbb{P}(V_4) \\ &\quad - \mathbb{P}(V_1 \cap V_2) - \mathbb{P}(V_1 \cap V_3) - \mathbb{P}(V_1 \cap V_4) \\ &\quad - \mathbb{P}(V_2 \cap V_3) - \mathbb{P}(V_2 \cap V_4) - \mathbb{P}(V_3 \cap V_4) \\ &\quad + \mathbb{P}(V_1 \cap V_2 \cap V_3) + \mathbb{P}(V_1 \cap V_2 \cap V_4) \\ &\quad + \mathbb{P}(V_1 \cap V_3 \cap V_4) + \mathbb{P}(V_2 \cap V_3 \cap V_4) \\ &\quad - \mathbb{P}(V_1 \cap V_2 \cap V_3 \cap V_4). \end{aligned}$$

Tous les rangements sont équiprobables, on peut donc passer par le dénombrement pour calculer ces probabilités. Un rangement peut se voir comme un tirage avec remise de six éléments dans un ensemble qui en contient quatre : la cardinal de l'univers vaut donc 4^6 .

Le cardinal de V_1 correspond au nombre de tirages où le numéro 1 n'a jamais été tiré. Un tel rangement correspond maintenant à un tirage avec remise de six éléments dans un ensemble qui en contient trois : il vaut donc 3^6 .

De même, le cardinal de $V_1 \cap V_2$ vaut 2^6 , celui de $V_1 \cap V_2 \cap V_3$ est égal à 1 et $V_1 \cap V_2 \cap V_3 \cap V_4$ est l'événement impossible.

Donc :

$$\begin{aligned} \mathbb{P}(V_1 \cup V_2 \cup V_3 \cup V_4) &= 4 \times \left(\frac{3}{4}\right)^6 - 6 \times \left(\frac{2}{4}\right)^6 + 6 \times \left(\frac{1}{4}\right)^6 - 0 \\ &= \frac{4 \times 3^6 - 6 \times 2^6 + 6}{4^6}. \end{aligned}$$

SF137 Décomposer l'événement en une succession d'étapes

D'un point de vue ensembliste, une succession se traduit par une intersection.

• Si les différentes étape de la succession sont indépendantes, alors la probabilité de l'intersection va directement s'écrire comme un produit.

• Si les différentes étape de la succession ne sont pas indépendantes, on se tournera vers la formule des probabilités composées.

Exemple

Un élève se rend au lycée le matin à vélo. Sur son trajet, il rencontre cinq feux tricolores dont le comportement suit les règles suivantes :

- au premier carrefour, le feu est vert avec une probabilité de $\frac{1}{2}$;
- pour $i \in \llbracket 1, 4 \rrbracket$,
 - si le feu i est rouge pour l'élève, alors le feu $i + 1$ sera vert de manière certaine,
 - sinon, la probabilité que le feu $i + 1$ soit lui-aussi vert est égale à $\frac{2}{3}$ de la probabilité que l'élève avait d'avoir un feu vert au carrefour i .

Gourmand, l'élève aime bien s'arrêter au premier carrefour pour sentir les effluves de la boulangerie

attendant ; puis il espère n'avoir que des feux verts pour ne pas arriver en retard en cours.

Quelle est la probabilité que l'élève rencontre une telle configuration des feux tricolore ?

Pour i compris entre 1 et 5, on note V_i l'événement « le feu i est vert ». On cherche donc la probabilité de $\overline{V_1} \cap V_2 \cap V_3 \cap V_4 \cap V_5$.

Ces différents événements ne sont pas indépendants donc on utilise la formule des probabilités composées :

$$\begin{aligned} \mathbb{P}(\overline{V_1} \cap V_2 \cap V_3 \cap V_4 \cap V_5) \\ &= \mathbb{P}(\overline{V_1}) \mathbb{P}_{\overline{V_1}}(V_2) \mathbb{P}_{\overline{V_1} \cap V_2}(V_3) \mathbb{P}_{\overline{V_1} \cap V_2 \cap V_3}(V_4) \mathbb{P}_{\overline{V_1} \cap V_2 \cap V_3 \cap V_4}(V_5) \\ &= \frac{1}{2} \times 1 \times \frac{2}{3} \times \frac{4}{9} \times \frac{8}{27} \\ &= \frac{32}{729}. \end{aligned}$$

SF138 Étudier l'événement contraire

Pour calculer la probabilité d'un événement A , étudier l'événement contraire $\overline{A}$ simplifie parfois les calculs. Une fois la probabilité de $\overline{A}$ trouvée, il suffit de calculer $1 - \mathbb{P}(\overline{A})$ pour trouver $\mathbb{P}(A)$.

Exemple

On lance une infinité de fois une pièce équilibrée pour jouer à « pile ou face ».

Quelle est la probabilité d'obtenir, au cours de cette série infinie, deux fois de suite le même résultat ?

Il n'est pas question d'essayer, ici, de décrire toutes les séries de tirages dans lesquelles deux résultats successifs sont identiques.

À l'inverse, il n'y a que deux éléments de l'univers où deux résultats successifs ne sont jamais identiques.

En notant, pour tout entier naturel i non nul, P_i l'événement « obtenir pile au i -ème lancer » et F_i l'événement « obtenir face au i -ème lancer », les deux séries de lancers sont :

$$\begin{aligned} P_1 \cap F_2 \cap P_3 \cap F_4 \cap P_5 \cap F_6 \cap \dots \\ F_1 \cap P_2 \cap F_3 \cap P_4 \cap F_5 \cap P_6 \cap \dots \end{aligned}$$

Ainsi, si on note D l'événement « obtenir deux fois de suite le même résultat », on a :

$$\begin{aligned} \mathbb{P}(D) &= 1 - \mathbb{P}(\overline{D}) \\ &= 1 - \mathbb{P}((P_1 \cap F_2 \cap P_3 \cap \dots) \sqcup (F_1 \cap P_2 \cap F_3 \cap \dots)) \\ &= 1 - \mathbb{P}(P_1 \cap F_2 \cap P_3 \cap \dots) - \mathbb{P}(F_1 \cap P_2 \cap F_3 \cap \dots). \end{aligned}$$

Enfin, les événements qui apparaissent dans ces deux intersections sont mutuellement indépendants. En transformant ces probabilités en produits, on remarque qu'elles sont toutes deux nulles. Donc, $\mathbb{P}(D) = 1$; c'est-à-dire l'événement D est presque certain.

SF139 Introduire un système complet d'événements

Lorsque la première étape comporte plusieurs cas, on conditionne par tous les résultats de cette première étape pour calculer les probabilités des étapes suivantes.

Exemple

Comme plus haut, un logiciel informatique permet de tirer un nombre entier naturel au hasard de manière aléatoire mais désormais, pour tout $i \in \mathbb{N}$, la probabilité de l'événement N_k : « obtenir le nombre k » vaut $\frac{1}{k! \cdot e}$.

Après avoir tiré un nombre aléatoire k avec ce programme, on tire une boule dans une urne composée de k boules blanches et une boule noire.

Calculer la probabilité d'obtenir la boule noire ?

La famille d'événements $(N_k)_{k \in \mathbb{N}}$ forme un système complet d'événements (en effet, ils couvrent tous les cas possibles et sont deux à deux incompatibles). Ainsi, en notant A l'événement « tirer la boule noire », la formule des probabilités totales donne :

$$\begin{aligned} \mathbb{P}(A) &= \sum_{k=0}^{+\infty} \mathbb{P}(A \cap N_k) \\ &= \sum_{k=0}^{+\infty} \mathbb{P}_{N_k}(A) \mathbb{P}(N_k) \\ &= \sum_{k=0}^{+\infty} \frac{1}{k+1} \frac{1}{k! \cdot e} \\ &= \frac{1}{e} \sum_{k=0}^{+\infty} \frac{1}{(k+1)!} \\ &= \frac{1}{e} \sum_{\ell=1}^{+\infty} \frac{1}{\ell!} \\ &= \frac{e-1}{e}. \end{aligned}$$

SF140 Lier cause et conséquence

La formule de Bayes permet de « retourner le temps » : si l'on imagine qu'un événement B est une conséquence plus ou moins directe d'un événement A , alors connaissant l'influence de A sur B (c'est-à-dire $\mathbb{P}_A(B)$), on peut calculer la probabilité de réalisation de la cause A sachant que la conséquence B est réalisée.

Exemple

On dispose de deux dés à six faces ; le premier est non pipé, alors que le second est truqué et donne le nombre 1 avec une probabilité de $\frac{1}{5}$.

On choisit un dé au hasard, puis on effectue n lancers avec. Le nombre 1 n'apparaît jamais au cours de ces n lancers.

À partir de quelle valeur de n peut-on affirmer avec une confiance de 90% qu'on a pioché le dé non pipé ?

Notons P l'événement « on pioche le dé pipé » et U_n l'événement « on n'obtient pas le nombre 1 au cours des n premiers tirages ». Connaissant la réalisation de l'un de ces deux événements, les lancers successifs peuvent être supposés indépendants. Ainsi :

$$\mathbb{P}_P(U_n) = \left(\frac{4}{5}\right)^n \quad \text{et} \quad \mathbb{P}_{\bar{P}}(U_n) = \left(\frac{5}{6}\right)^n.$$

On cherche maintenant à partir de quelle valeur de n la probabilité de P sachant U_n est inférieur à 0,1.

Pour calculer cette probabilité, on applique la formule de Bayes :

$$\mathbb{P}_{U_n}(P) = \frac{\mathbb{P}(P)\mathbb{P}_P(U_n)}{\mathbb{P}(U_n)}.$$

La calcul de $\mathbb{P}(U_n)$ repose sur l'utilisation de la formule des probabilités totales avec le système complet d'événements $(P, \bar{P})$:

$$\mathbb{P}(U_n) = \mathbb{P}(P)\mathbb{P}_P(U_n) + \mathbb{P}(\bar{P})\mathbb{P}_{\bar{P}}(U_n) = \frac{1}{2} \left(\left(\frac{4}{5}\right)^n + \left(\frac{5}{6}\right)^n \right).$$

D'où :

$$\begin{aligned} \mathbb{P}_{U_n}(P) \leq \frac{1}{10} &\iff \frac{\left(\frac{4}{5}\right)^n}{\left(\frac{4}{5}\right)^n + \left(\frac{5}{6}\right)^n} \leq \frac{1}{10} \\ &\iff \frac{1}{1 + \left(\frac{25}{24}\right)^n} \leq \frac{1}{10} \\ &\iff 1 + \left(\frac{25}{24}\right)^n \geq 10 \\ &\iff n \geq \frac{2\ln(3)}{2\ln(5) - 3\ln(2) - \ln(3)}. \end{aligned}$$

En calculant une valeur approchée de ce réel, on constate qu'il faudra au moins 54 lancers avant de pouvoir affirmer avoir 90% de chances d'utiliser le dé non pipé.

Quelques exercices d'application

Exercice C-74

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

1. Soit $(A, B) \in \mathcal{T}^2$. Montrer que les quatre propositions suivantes sont équivalentes :
 - i. A et B sont indépendants ;
 - ii. A et $\bar{B}$ sont indépendants ;
 - iii. $\bar{A}$ et B sont indépendants ;
 - iv. $\bar{A}$ et $\bar{B}$ sont indépendants.
2. Soit $A \in \mathcal{T}$. Montrer que les deux propositions suivantes sont équivalentes :
 - i. pour tout événement B, les événements A et B sont indépendants ;
 - ii. A est un événement presque certain ou négligeable.
3. On lance deux fois de suite un dé non pipé et on considère les trois événements suivants :
 - (a) A : « le résultat du premier lancer est pair » ;
 - (b) B : « le résultat du second lancer est impair » ;
 - (c) C : « la somme des deux résultats est impaire ».

Étudier l'indépendance de ces trois événements.

Exercice C-75

On considère une suite infinie de lancers indépendants d'une pièce truquée donnant *pile* avec une probabilité $p \in]0, 1[$.

1. Soit $n \geq 2$. Quelle est la probabilité que la séquence *pile-face* apparaisse pour la première fois aux lancers $n - 1$ et n ?
2. Quelle est la probabilité que la séquence *pile-face* apparaisse au moins une fois?
3. Quelle est la probabilité que la séquence *pile-pile* apparaisse sans qu'il n'y ait jamais eu de séquence *pile-face* auparavant?

Exercice C-76

On suppose pour simplifier qu'une année ne compte que 365 jours différents et que toutes ces journées sont équiprobables pour la naissance des bébés.

1. On considère un groupe composé de n personnes ($n \geq 2$). Quelle est la probabilité qu'au moins deux personnes du groupe partagent la même date d'anniversaire?
2. Écrire une fonction Python qui prend en argument un réel α strictement compris entre 0 et 1 et qui renvoie la première valeur de n pour laquelle la probabilité précédente dépasse la valeur α .
3. Combien faut-il de personnes dans le groupe pour que la probabilité d'avoir au moins deux personnes avec le même date d'anniversaire dépasse 0,5? Et pour 0,9999?

Exercice C-77

1. Soit A et B deux événements de probabilité non nulle. Montrer que :

$$\mathbb{P}_A(B) \geq \mathbb{P}(B) \implies \mathbb{P}_B(A) \geq \mathbb{P}(A).$$

2. Soit $n \in \mathbb{N}^*$. Une urne contient $2n$ boules numérotées de 1 à $2n$. On tire deux boules sans remise dans cette urne.

Sachant que la seconde boule porte un numéro strictement supérieur à n , la probabilité que le premier numéro soit inférieur ou égal à n est-elle supérieure ou inférieure à $\frac{1}{2}$?

Exercice C-78

Une urne contient initialement une boule blanche et une boule rouge, indiscernables au toucher. Un joueur tire une boule au hasard :

- si la boule est blanche, alors le joueur gagne ;
- sinon, il remet la boule rouge dans l'urne et rajoute une boule rouge ; puis répète le même processus.

Ainsi, au fur et à mesure des tirages, la probabilité de gagner devient de plus en plus faible.

1. Soit $n \in \mathbb{N}^*$. Quelle est la probabilité que le joueur gagne avant le n -ième tirage?
2. Quelle est la probabilité que le joueur ne gagne jamais?
3. On recommence le même protocole, mais cette fois-ci, au lieu de rajouter une boule rouge, on double le nombre de boules rouges supplémentaires. Reprendre les questions précédentes et à l'aide d'une fonction Python, estimer la probabilité que le joueur ne gagne jamais.

Exercice C-79

Une puce évolue sur trois cases A, B et C.

À l'instant $t = 0$, la puce se trouve sur la case A, puis elle se déplace de façon aléatoire sur ces cases selon la règle suivante :

- si la puce se trouve en A ou B à l'instant $t = n$, alors elle ira sur l'une des deux autres cases avec équiprobabilité à l'instant $t = n + 1$;
- si la puce se trouve en C à l'instant $t = n$, alors elle y restera à l'instant $t = n + 1$.

On suppose que cette situation est modélisée par un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

On note A_n (resp. B_n, C_n) l'événement : « la puce se trouve en A (resp. B, C) à l'instant $t = n$ » et on note a_n, b_n et c_n les probabilités correspondantes.

1. Justifier que pour tout $n \in \mathbb{N}$, on a : $a_{n+1} = \frac{1}{2}b_n$.
2. Donner (sans démonstration) les relations de récurrence analogues entre b_{n+1} puis c_{n+1} et a_n, b_n et c_n .
3. Montrer que la suite $(a_n + b_n)_{n \in \mathbb{N}}$ est géométrique.
4. En déduire la valeur de c_n en fonction de n .
5. Déterminer les expressions de a_n et b_n en fonction de n .
6. Montrer que la puce atteint la case C presque sûrement.
7. Sachant que la puce est en C à l'instant $t = n + 1$, calculer la probabilité qu'elle y ait été pour la première fois à l'instant $t = n$.

Exercice C-80 (lemme de Borel-Cantelli)

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

1. Interpréter l'événement $B = \bigcap_{n=1}^{+\infty} \bigcup_{k=n}^{+\infty} A_k$.
2. Montrer que si la série $\sum \mathbb{P}(A_n)$ converge alors B est négligeable.
3. On suppose que la série $\sum \mathbb{P}(A_n)$ diverge et que les $(A_n)_{n \in \mathbb{N}}$ sont mutuellement indépendants.
 - a. Exprimer $\mathbb{P}\left(\bigcap_{k=n}^m \overline{A_k}\right)$ en fonction des $\mathbb{P}(A_k)$.
 - b. En déduire que la série $\sum \ln(1 - \mathbb{P}(A_n))$ diverge.
 - c. En déduire que B est presque sûr.
4. Application.

Sébastien tape aléatoirement sur un clavier (à 105 touches), quelle est la probabilité qu'il écrive une infinité de fois la phrase « à jamais les premiers » ?

B - Notion de variable aléatoire réelle

B.1 - Définitions et premières propriétés

Définition V-1

Étant donné un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, on appelle **variable aléatoire réelle** toute application X définie sur Ω à valeurs dans $\mathbb{R}$, telle que pour tout $t \in \mathbb{R}$:

$$\{\omega \in \Omega ; X(\omega) \leq t\} \in \mathcal{A}.$$

Remarques

1 ▶ Si I est un intervalle de $\mathbb{R}$ et $t \in \mathbb{R}$, on note :

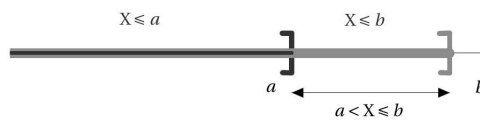
$$(X \in I) = \{\omega \in \Omega \mid X(\omega) \in I\}, \quad (X \leq t) = \{\omega \in \Omega \mid X(\omega) \leq t\} \quad \text{et} \quad (X = t) = \{\omega \in \Omega \mid X(\omega) = t\}.$$

Ainsi, X est une variable aléatoire si, pour tout réel t , l'ensemble $(X \leq t)$ est un événement.

2 ▶ Notons que l'on emploie également la notation $[X \in I]$, etc.

3 ▶ Soit a et b deux réels et X une variable aléatoire réelle, alors $(X \in]a, b])$ est un événement. En effet $(X \leq a) \in \mathcal{A}$ et $(X \leq b) \in \mathcal{A}$ donc :

$$(X \in]a, b]) = (X \leq b) \setminus (X \leq a) = (X \leq b) \cap \overline{(X \leq a)} \in \mathcal{A}.$$



4 ▶ La donnée des probabilités $\mathbb{P}(X \in]a, b])$, pour tous réels a et b avec $a < b$, définit la **loi de probabilité** de la variable aléatoire réelle X .

5 ▶ Soit X et Y deux variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, alors :

- la combinaison linéaire $\lambda X + \mu Y$ où $\lambda, \mu \in \mathbb{R}$,
- le produit $X \cdot Y$,
- le maximum $\max(X, Y)$ et le minimum $\min(X, Y)$

sont encore des variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$. Cela s'étend à un nombre fini de variables aléatoires.

Définition V-2

Soit X une variable aléatoire réelle sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

On appelle **fonction de répartition** de X , l'application :

$$F_X : \mathbb{R} \rightarrow [0, 1], \quad t \mapsto \mathbb{P}(X \leq t).$$

Avec les notations précédentes, on a :

Proposition V-3

1. F_X est croissante sur $\mathbb{R}$.
2. $F_X(x) \xrightarrow{x \rightarrow -\infty} 0$ et $F_X(x) \xrightarrow{x \rightarrow +\infty} 1$.
3. F_X est continue à droite en tout point de $\mathbb{R}$.
4. Pour tout réel x , on a : $\mathbb{P}(X > x) = 1 - F_X(x)$.
5. Pour tous réels a et b avec $a < b$, on a : $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$.

Démonstration

1. Soit x et y deux réels avec $x < y$ alors $(X \leq x) \subset (X \leq y)$ donc :

$$\mathbb{P}(X \leq x) \leq \mathbb{P}(X \leq y) \text{ i.e. } F_X(x) \leq F_X(y),$$

donc F_X est croissante sur $\mathbb{R}$.

2. $\diamond$ La fonction F_X est croissante et majorée par 1 donc elle admet une limite en $+\infty$.

La suite d'événements $((X \leq n))_{n \in \mathbb{N}}$ est croissante et on a $\Omega = \bigcup_{n \in \mathbb{N}} (X \leq n)$, d'où par le théorème de la limite monotone :

$$\mathbb{P}(\Omega) = \mathbb{P}\left(\bigcup_{n \in \mathbb{N}} (X \leq n)\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq n),$$

ce qui donne :

$$1 = \lim_{n \rightarrow +\infty} F_X(n).$$

Comme on sait que F_X admet une limite en $+\infty$, cette limite est égale à 1.

- $\diamond$ La fonction F_X est croissante et minorée par 0 donc elle admet une limite en $-\infty$.

La suite d'événements $((X \leq -n))_{n \in \mathbb{N}}$ est décroissante et on a $\emptyset = \bigcap_{n \in \mathbb{N}} (X \leq -n)$, d'où par le théorème de la limite monotone :

$$\mathbb{P}(\emptyset) = \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} (X \leq -n)\right) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq -n),$$

ce qui donne :

$$0 = \lim_{n \rightarrow +\infty} F_X(-n) = \lim_{n \rightarrow +\infty} F_X(n).$$

Comme on sait que F_X admet une limite en $-\infty$, cette limite est égale à 0.

3. Soit a un réel, puisque la fonction F_X est croissante, elle admet une limite à droite ℓ en a .

La suite d'événements $((X \leq a + \frac{1}{n}))_{n \in \mathbb{N}^*}$ est décroissante et on a :

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

d'où par le théorème de la limite monotone :

$$\mathbb{P}(X \leq a) = \mathbb{P}\left(\bigcap_{n \in \mathbb{N}^*} \left(X \leq a + \frac{1}{n}\right)\right) = \lim_{n \rightarrow +\infty} \mathbb{P}\left(X \leq a + \frac{1}{n}\right),$$

ce qui donne :

$$F_X(a) = \lim_{n \rightarrow +\infty} F_X\left(a + \frac{1}{n}\right).$$

Comme on sait que F_X admet une limite à droite en a , cette limite est égale à $F_X(a)$ ce qui signifie que F_X est continue à droite en a .

4. On a :

$$\mathbb{P}((X > x) \cup (X \leq x)) = \mathbb{P}(\Omega) = 1,$$

soit :

$$\mathbb{P}(X > x) + \mathbb{P}(X \leq x) = 1,$$

puis :

$$\mathbb{P}(X > x) = 1 - \mathbb{P}(X \leq x) = 1 - F_X(x).$$

5. On a $(a < X \leq b) = (X \leq b) \setminus (X \leq a)$ et $(X \leq a) \subset (X \leq b)$, d'où :

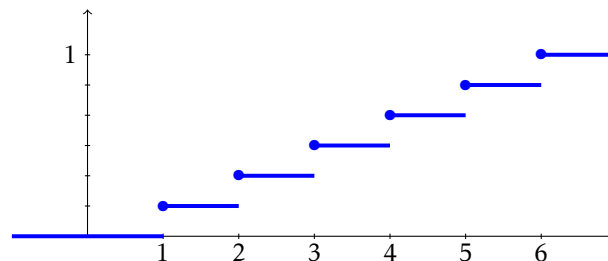
$$\mathbb{P}(a < X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X \leq a),$$

c'est-à-dire :

$$\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a).$$

Exemples

1 ▶ On lance un dé à 6 faces non pipé et on note X le chiffre obtenu, alors $X(\Omega) = \llbracket 1, 6 \rrbracket$ et la représentation graphique de la fonction de répartition est :



2 ▶ On joueur lance des fléchettes sur une cible circulaire de rayon R et atteint la cible à chaque fois.

On suppose que la probabilité d'atteindre une partie de la cible est proportionnelle à son aire c'est-à-dire que si A est une partie de la cible alors la probabilité que le point d'impact soit dans A est $\frac{\text{aire}(A)}{\pi R^2}$.

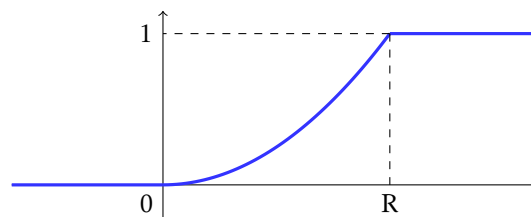
Le joueur envoie une fléchette, on note X la variable aléatoire réelle égale à la distance en cm entre le centre O de la cible et le point d'impact.

Déterminons la fonction de répartition F_X de X .

Tout d'abord, puisque X est une distance, on a $\mathbb{P}(X \leq x) = 0$ pour tout réel négatif x . De plus, on suppose que le joueur atteint toujours la cible donc $\mathbb{P}(X \leq x) = 1$ pour tout réel $x \geq R$. Enfin, dans le cas où $x \in [0, R]$, on a :

$$\mathbb{P}(X \leq x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}.$$

La représentation graphique de F_X est donc :



Remarques

- 1 ▶ Une fonction de répartition caractérise une loi de probabilité : deux variables aléatoires ont même loi si et seulement si elles ont la même fonction de répartition.
- 2 ▶ Notons que des variables aléatoires peuvent suivre la même loi sans être égales !

Par exemple, si l'on lance deux dés à 6 faces non pipés et l'on note X_1 le chiffre donné par le premier et X_2 celui donné par le second. Alors X_1 et X_2 suivent la même loi :

k	1	2	3	4	5	6
$\mathbb{P}(X_1 = k)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
$\mathbb{P}(X_2 = k)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Cependant, on a $\mathbb{P}(X_1 \neq X_2) > 0$.

Exemple

Une urne contient 9 boules blanches et une boule noire, toutes indiscernables au toucher. On tire successivement des boules de cette urne et on note X le nombre de tirages nécessaires pour tirer la boule noire. On note B_k l'événement «on tire une boule blanche au k -ième tirage».

1. Déterminons la loi de X si les tirages sont effectués sans remise.

2. Déterminons la loi de X si les tirages sont effectués avec remise.

B.2 - Le cas des variables aléatoires discrètes

Dans le cas où $X(\Omega)$ est fini ou dénombrable, la définition de la variable aléatoire se simplifie.

Définition V-4

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Une **variable aléatoire discrète** est une application $X : \Omega \rightarrow \mathbb{R}$ telle que :

- ▷ $X(\Omega) = \{x_i ; i \in I\}$ où I est une partie de $\mathbb{N}$;
- ▷ pour tout $i \in I$, $(X = x_i) = \{\omega \in \Omega ; X(\omega) = x_i\}$ est un événement.

Donner la loi d'une variable aléatoire discrète X signifie donner l'ensemble $X(\Omega)$ des valeurs prises par X et, pour chaque $x \in X(\Omega)$, la probabilité $\mathbb{P}(X = x)$.

Exemples

- 1 ▷ On lance une fléchette en visant une cible et on note Ω l'ensemble des points d'impact possibles ; au résultat ω du lancer, on fait correspondre un score $X(\omega)$ égal à 100, 50, 30, 20 ou 10 selon la zone de la cible atteinte, et 0 si la fléchette est en dehors de la cible. On a donc une application $X : \Omega \rightarrow \mathbb{R}$ telle que $X(\Omega) = \{0, 10, 20, 30, 50, 100\}$.
- 2 ▷ On lance deux dés jusqu'à obtenir un double 6 et on note Ω l'ensemble des séries possibles de lancers ; au résultat ω de l'expérience, on associe le nombre total $X(\omega)$ de lancers. On a donc une application $X : \Omega \rightarrow \mathbb{R}$ telle que $X(\Omega) = \mathbb{N}^*$.

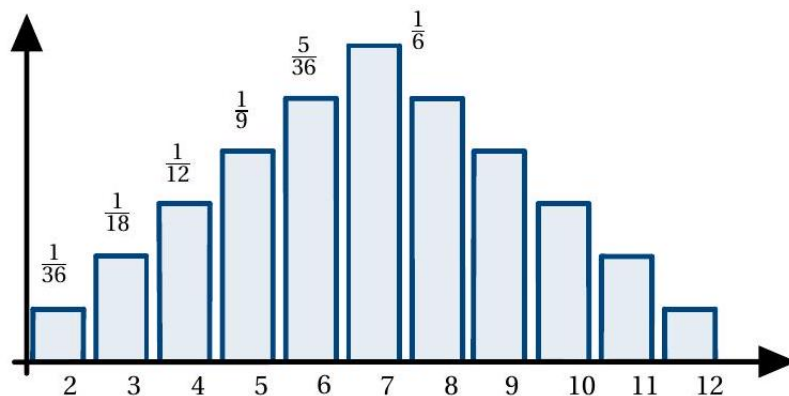
Remarques

- 1 ▷ On donne parfois la loi sous forme d'un tableau.

Par exemple, on lance deux dés à 6 faces non pipés et on note X la somme des chiffres obtenus, alors $X(\Omega) = \llbracket 2, 12 \rrbracket$ et la loi de X est donnée dans le tableau :

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(X = k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

On peut également utiliser un diagramme en bâtons. En abscisse, on place les valeurs possibles de la variable (ici, 2, 3, ..., 12) et en ordonnée les probabilités correspondantes (ici $\mathbb{P}(X = 2)$, $\mathbb{P}(X = 3)$, etc.).



- 2 ▷ À une v.a discrète X est associée le système complet d'événements $((X = x))_{x \in X(\Omega)}$. En particulier :

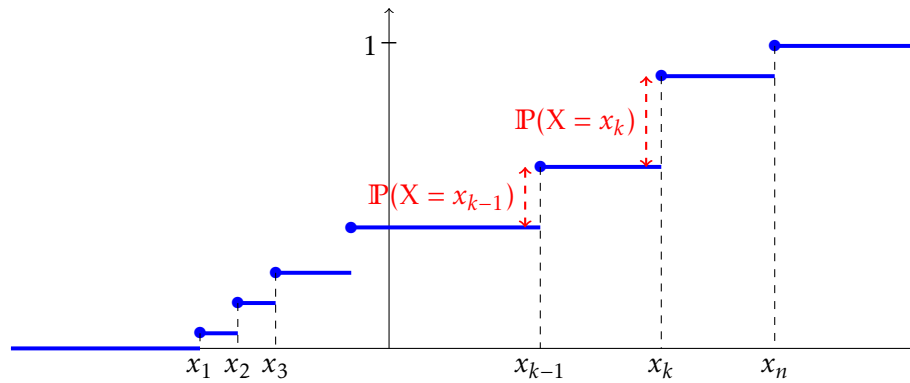
$$\sum_{x \in X(\Omega)} \mathbb{P}(X = x) = \mathbb{P}(\Omega) = 1.$$

- 3 ▶ La courbe d'une fonction de répartition d'une variable aléatoire discrète est en escalier. Les abscisses des sauts donnent les valeurs x_k et la hauteur du saut donne la probabilité $\mathbb{P}(X = x_k)$.

Considérons par exemple le cas où X est une variable aléatoire discrète finie, on note $X(\Omega) = \{x_1, x_2, \dots, x_n\}$ avec $x_1 < x_2 < \dots < x_n$. Avec un léger abus de notation sans incidence pour ce qui suit, on peut même supposer que $\mathbb{P}(X = x_k) \neq 0$ pour tout $k \in \llbracket 1, n \rrbracket$ (ce qui signifie que F_X a une discontinuité en x_k).

Alors, pour tout réel x , on a :

$$F_X(x) = \begin{cases} 0 & \text{si } x < x_1 \\ \mathbb{P}(X = x_1) & \text{si } x_1 \leq x < x_2 \\ \mathbb{P}(X = x_1) + \mathbb{P}(X = x_2) & \text{si } x_2 \leq x < x_3 \\ \mathbb{P}(X = x_1) + \mathbb{P}(X = x_2) + \mathbb{P}(X = x_3) & \text{si } x_3 \leq x < x_4 \\ \vdots & \\ \mathbb{P}(X = x_1) + \dots + \mathbb{P}(X = x_{n-1}) & \text{si } x_{n-1} \leq x < x_n \\ 1 & \text{si } x_n \leq x \end{cases}.$$



Considérons maintenant le cas où X est une variable aléatoire discrète infinie, on note $X(\Omega) = \{x_n ; n \in \mathbb{N}\}$ avec $(x_n)_{n \in \mathbb{N}}$ en supposant là encore que tous les événements $(X = x_k)$ sont de probabilité non nulle.

Alors $F_X(x) = 0$ pour tout réel $x < x_0$ et, pour $x \in [x_k, x_{k+1}[$, on a :

$$F_X(t) = \sum_{i=0}^k \mathbb{P}(X = x_i).$$

Dans le cas particulier où $X(\Omega) \subset \mathbb{N}$, on a :

$$\mathbb{P}(X = k) = \mathbb{P}(X \leq k) - \mathbb{P}(X \leq k-1) = \mathbb{P}(X \geq k) - \mathbb{P}(X \geq k+1).$$

Définition V-5

1. Soit X et Y deux variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

On dit que X et Y sont **indépendantes** lorsque :

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad \mathbb{P}((X = x) \cap (Y = y)) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y).$$

2. Soit $X_1, \dots, X_n$ des variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

On dit que $X_1, \dots, X_n$ sont mutuellement indépendantes lorsque :

$$\forall (x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega), \quad \mathbb{P}\left(\bigcap_{i=1}^n (X_i = x_i)\right) = \prod_{i=1}^n \mathbb{P}(X_i = x_i).$$

Remarque

On constate que si X et Y sont indépendantes alors :

$$\forall (t_1, t_2) \in \mathbb{R}^2, \quad F_X(t_1) F_Y(t_2) = \mathbb{P}([X \leq t_1] \cap [Y \leq t_2]).$$

En effet, soit $(t_1, t_2) \in \mathbb{R}^2$. On a

$$(X \leq t_1) \cap (Y \leq t_2) = \bigcup_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ x \leq t_1, y \leq t_2}} (X = x) \cap (Y = y)$$

L'union est disjointe, par la propriété d'additivité d'une probabilité, puis par la définition de l'indépendance, on a :

$$\begin{aligned} \mathbb{P}((X \leq t_1) \cap (Y \leq t_2)) &= \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ x \leq t_1, y \leq t_2}} \mathbb{P}((X = x) \cap (Y = y)) \\ &= \sum_{\substack{(x,y) \in X(\Omega) \times Y(\Omega) \\ x \leq t_1, y \leq t_2}} \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y) \\ &= \left(\sum_{\substack{x \in X(\Omega) \\ x \leq t_1}} \mathbb{P}(X = x) \right) \cdot \left(\sum_{\substack{y \in Y(\Omega) \\ y \leq t_2}} \mathbb{P}(Y = y) \right) \\ \mathbb{P}((X \leq t_1) \cap (Y \leq t_2)) &= F_X(t_1) F_Y(t_2). \end{aligned}$$

La réciproque est vraie mais plus délicate à prouver.

B.3 - Lois discrètes usuelles

Donner la loi d'une variable aléatoire X discrète revient à la donnée de $X(\Omega)$ (ensemble des valeurs prises par la variable aléatoire) et, pour chaque $k \in X(\Omega)$, $\mathbb{P}(X = k)$.

B.3.a - Loi certaine

On dit qu'une variable aléatoire X suit la **loi certaine** égale à m lorsque X est constante égale à m .

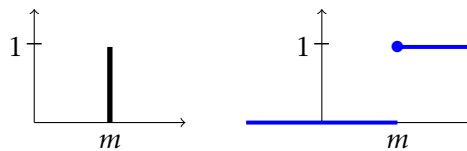
Cela signifie que $X(\Omega) = \{m\}$ et $\mathbb{P}(X = m) = 1$.

C'est le cas lorsque le résultat de l'expérience ne dépend en fait pas du hasard.

Remarque

Cela ne paraît pas être une situation intéressante mais c'est en fait parfois utile pour décrire des lois plus complexes. Par exemple, considérons un jeu dont la mise d'entrée est de m euros et dont on note G le gain alors le « vrai » gain est obtenu en retirant à G la variable certaine m .

Les représentations graphiques de la loi et de la fonction de répartition sont :



B.3.b - Loi uniforme

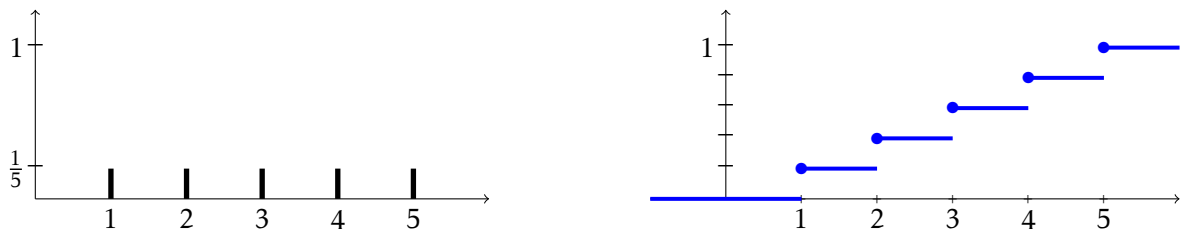
Soit X une variable aléatoire finie, on note $X(\Omega) = \{x_1, \dots, x_n\}$.

On dit que X suit la **loi uniforme** sur $X(\Omega)$ lorsque les événements $(X = x_1), (X = x_2), \dots, (X = x_n)$ sont équiprobables ce qui revient à dire :

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

On note alors $X \hookrightarrow \mathcal{U}(X(\Omega))$.

Les représentations graphiques de la loi et de la fonction de répartition sont, par exemple pour $\mathcal{U}(\llbracket 1, 5 \rrbracket)$:



B.3.c - Loi de Bernoulli

On considère un réel $p \in]0, 1[$.

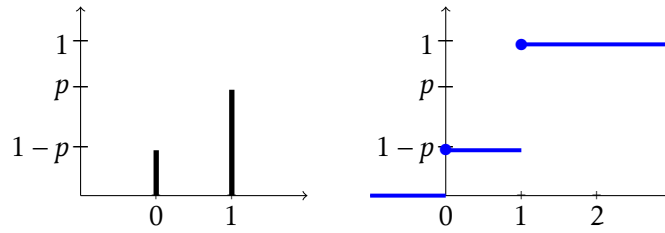
On dit qu'une variable aléatoire X suit la **loi de Bernoulli** de paramètre p lorsque :

$$X(\Omega) = \{0; 1\} \text{ , } \mathbb{P}(X = 1) = p \text{ et } \mathbb{P}(X = 0) = 1 - p.$$

On note alors $X \hookrightarrow \mathcal{B}(p)$.

La loi de Bernoulli régit les expériences aléatoires qui n'ont que deux issues possibles : le « succès » avec une probabilité p et « l'échec » avec la probabilité $q = 1 - p$.

Les représentations graphiques de la loi et de la fonction de répartition de $\mathcal{B}(p)$ sont (par exemple dans le cas où $p > \frac{1}{2}$) :



Remarque

L'indicatrice $\mathbb{1}_A$ d'un événement A est la fonction qui vaut 1 lorsque l'événement A se produit et 0 sinon :

$$\mathbb{1}_A(\omega) = \begin{cases} 0 & \text{si } \omega \notin A \\ 1 & \text{si } \omega \in A \end{cases}.$$

Si $A \in \mathcal{A}$ alors $\mathbb{1}_A$ est une variable aléatoire réelle.

En effet, considérons $x \in \mathbb{R}$. Si $x < 0$ alors $(\mathbb{1}_A \leq x) = \emptyset \in \mathcal{A}$.

Si $0 \leq x < 1$ alors :

$$(\mathbb{1}_A \leq x) = (\mathbb{1}_A = 0) = \{\omega \in \Omega ; \omega \notin A\} = \overline{A} \in \mathcal{A}.$$

Si $x \leq 1$ alors :

$$(\mathbb{1}_A \leq x) = (\mathbb{1}_A = 1) = \Omega \in \mathcal{A}.$$

On a donc $(\mathbb{1}_A \leq x) \in \mathcal{A}$ pour tout réel x donc $\mathbb{1}_A$ est une variable aléatoire réelle sur $(\Omega, \mathcal{A})$.

Enfin, on a vu que $(\mathbb{1}_A = 0) = \overline{A}$ donc $\mathbb{P}(\mathbb{1}_A = 0) = \mathbb{P}(\overline{A})$.

Puisque $\mathbb{1}_A(\Omega) = \{0, 1\}$, il s'ensuit que $\mathbb{P}(\mathbb{1}_A = 1) = 1 - \mathbb{P}(\overline{A}) = \mathbb{P}(A)$.

Donc la variable aléatoire $\mathbb{1}_A$ suit la loi $\mathcal{B}(\mathbb{P}(A))$.

B.3.d - Loi binomiale

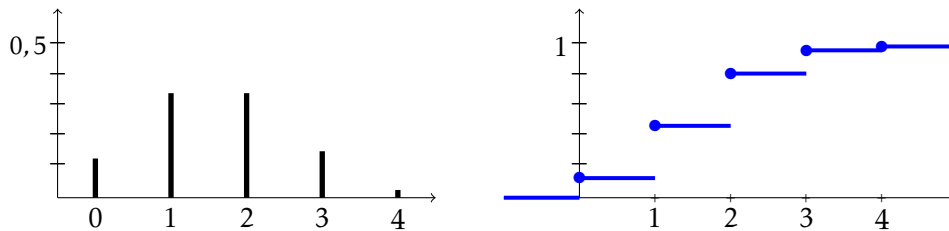
On considère un réel $p \in]0, 1[$, on note $q = 1 - p$.

On dit qu'une variable aléatoire X suit la **loi binomiale** de paramètres n et p lorsque $X(\Omega) = \llbracket 0, n \rrbracket$ et :

$$\forall k \in \llbracket 0, n \rrbracket, \mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k}.$$

On note alors $X \hookrightarrow \mathcal{B}(n, p)$.

Les représentations graphiques de la loi et de la fonction de répartition de $\mathcal{B}(n, p)$ sont, par exemple pour $n = 4$ et $p = 0,4$:



Remarques

- 1► Si $X \hookrightarrow \mathcal{B}(n, p)$ alors $n - X \hookrightarrow \mathcal{B}(n, q)$ car cela revient à intervertir les notions de succès et d'échec à chacune des expériences de Bernoulli.
- 2► La loi binomiale régit les répétitions de n expériences de Bernoulli identiques et mutuellement indépendantes (avec la probabilité du succès égale à p).

Exemple

On lance n fois une pièce équilibrée. Pour tout entier i , on note A_i l'événement «on fait pile au i -ème lancer».

Alors $X = \sum_{i=1}^n \mathbb{1}_{A_i}$ compte le nombre de piles obtenus lors des n lancers donc X suit la loi binomiale $\mathcal{B}(n, \frac{1}{2})$.

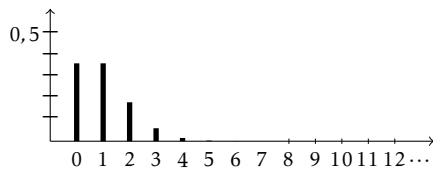
B.3.e - Loi de Poisson

Considérons un réel $\lambda > 0$.

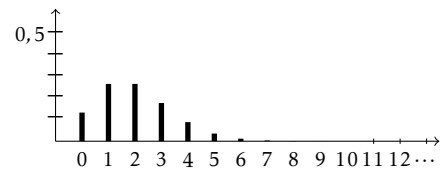
On dit qu'une variable aléatoire réelle discrète X suit la **loi de Poisson** de paramètre λ lorsque $X(\Omega) = \mathbb{N}$ et :

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

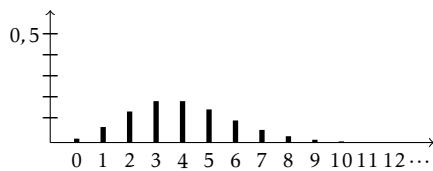
On note $X \hookrightarrow \mathcal{P}(\lambda)$.



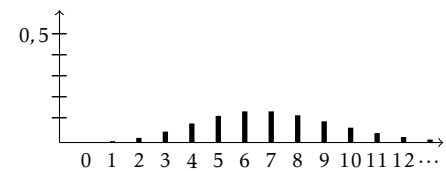
loi pour $\mathcal{P}(1)$



loi pour $\mathcal{P}(2)$



loi pour $\mathcal{P}(4)$



loi pour $\mathcal{P}(7)$

Remarque

Considérons une succession de n épreuves indépendantes, chacune caractérisée par un succès de probabilité $p \in]0, 1[$ (ou un échec de probabilité $q = 1 - p$), alors la variable aléatoire X comptant le nombre de succès suit la loi binomiale $\mathcal{B}(n, p)$ de paramètres n et p .

Cependant, lorsque n est grand et p petit, on peut adopter comme modèle pour X la loi de Poisson $\mathcal{P}(np)$.

Cette approximation justifie le nom également employé pour cette loi de « loi des événements rares » (par exemple pour compter le nombre de coquilles dans un livre, le nombre d'accidents en une journée sur une portion de route donnée,...)

Dans la pratique, on fait cette approximation lorsque $p \leq 0,1$, $n \geq 30$ et $np \leq 10$.

Le résultat justifiant cette approximation est le suivant.

Soit (X_n) une suite de variables aléatoires réelles discrètes telle que pour tout n , $X_n \hookrightarrow \mathcal{B}(n, \frac{\lambda}{n})$ avec $\lambda > 0$. Alors :

$$\forall k \in \mathbb{N} \quad \mathbb{P}(X_n = k) \xrightarrow{n \rightarrow +\infty} \mathbb{P}(Y = k),$$

où Y est une variable aléatoire de loi $\mathcal{P}(\lambda)$.

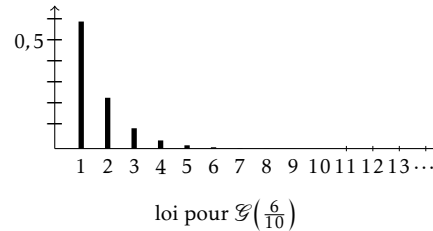
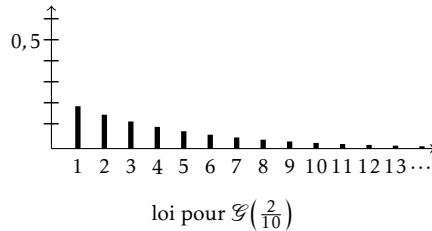
B.3.f - Loi géométrique

Considérons un réel $p \in]0, 1[$.

On dit qu'une variable aléatoire réelle discrète X suit la **loi géométrique** de paramètre p lorsque $X(\Omega) = \mathbb{N}^*$ et :

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

On note $X \hookrightarrow \mathcal{G}(p)$.



Remarques

- 1► On considère une succession infinie d'expériences indépendantes dont chacune a une même probabilité $p \in]0, 1[$ de conduire à un résultat appelé « succès ». On montre alors que l'on obtient au moins un succès presque sûrement.

Dans ces conditions, la variable aléatoire X qui donne le rang du premier succès suit la loi géométrique de paramètre p .

- 2► On dit qu'une variable aléatoire réelle X suit la loi géométrique « décalée » sur $\mathbb{N}$, et on note $X \hookrightarrow \mathcal{G}_0(p)$, lorsque $X(\Omega) = \mathbb{N}$ et :

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

Cela revient à dire que la variable $Y = X + 1$ suit la loi $\mathcal{G}(p)$.

Dans le contexte de modélisation de la remarque précédente, cela revient à compter le nombre d'échecs avant le premier succès.

- 3► Soit X une variable aléatoire réelle suivant la loi $\mathcal{G}(p)$, alors pour tous entiers s et t de $\mathbb{N}$, on a :

$$\mathbb{P}_{(X>s)}(X > s + t) = \mathbb{P}(X > t)$$

(la relation est également vraie avec des inégalités larges). Ce résultat traduit une « absence de mémoire ». Dans le cadre du modèle exposé plus haut, cela revient à dire que la probabilité d'attendre t résultats avant le premier succès ne dépend pas du nombre d'échecs qui ont précédé.

On a également un résultat réciproque. Si X est une variable aléatoire réelle discrète, avec $X(\Omega) = \mathbb{N}^*$, qui vérifie la propriété précédente alors X suit la loi géométrique de paramètre $p = \mathbb{P}(X = 1)$.

C - Espérance et variance d'une variable aléatoire discrète

C.1 - Espérance et formule de transfert

Définition V-6

Soit X une variable aléatoire discrète infinie. On note $X(\Omega) = \{x_k; k \in \mathbb{N}\}$.

On dit que X admet une espérance si la série de terme général $x_k \mathbb{P}(X = x_k)$ est absolument convergente.

Alors on définit l'espérance de X par :

$$\mathbb{E}(X) = \sum_{k=0}^{+\infty} x_k \mathbb{P}(X = x_k).$$

Il ne faut pas oublier de vérifier la convergence absolue. Elle est nécessaire pour justifier que l'espérance est bien définie et ne dépend pas du choix de l'indexation $X(\Omega)$.

Example

On effectue une infinité de tirages successifs, supposés mutuellement indépendants, dans une urne contenant initialement une boule rouge et une boule bleue. À chaque tirage, on note la couleur de la boule et on la remet dans l'urne en y ajoutant une boule bleue.

On note X (respectivement Y) la variable aléatoire égale au rang d'apparition de la première boule bleue (respectivement rouge). Pour que ces deux variables aléatoires soient bien définies, on admet que presque sûrement une boule bleue ou rouge finit par apparaître.

1. Donner la loi de X .

2. La v.a. X admet-elle une espérance?

3. Que peut-on dire pour la variable Y?

⇒ Principales propriétés

1 ▶ Linéarité de l'espérance

Soit X et Y deux variables aléatoires discrètes sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et $(\lambda, \mu) \in \mathbb{R}^2$. Si X et Y admettent une espérance alors $\lambda X + \mu Y$ admet une espérance et :

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

2 ▶ Existence d'une espérance par domination

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

Si $0 \leq |X| \leq Y$ presque sûrement (*i.e.* $\mathbb{P}(|X| \leq Y) = 1$) et si Y admet une espérance alors X admet une espérance et $|\mathbb{E}(X)| \leq \mathbb{E}(Y)$.

3 ▶ Croissance de l'espérance

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

Si X et Y admettent une espérance et si $X \leq Y$ presque sûrement (*i.e.* $\mathbb{P}(X \leq Y) = 1$) alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

4 ▶ Théorème de transfert

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et g une application définie sur $X(\Omega) = \{x_k ; k \in I\}$. Les énoncés suivants sont équivalents :

- i. la variable aléatoire $g(X)$ possède une espérance ;
- ii. la série de terme général $g(x_k) \mathbb{P}(X = x_k)$ est absolument convergente.

Dans ce cas :

$$\mathbb{E}(g(X)) = \sum_{k \in I} g(x_k) \mathbb{P}(X = x_k).$$

Remarques

- 1 ▶ Le moment d'ordre 1 de X est son espérance.
 2 ▶ Dans le cas d'une variable aléatoire discrète, et sous réserve de convergence absolue, la formule de transfert donne :

$$m_r(X) = \sum_{x \in X(\Omega)} x^r \mathbb{P}(X = x).$$

- 3 ▶ Si X admet un moment d'ordre r alors X admet un moment à tout ordre s entre 0 et $r - 1$.

En effet, notons $X(\Omega) = \{x_n; n \in \mathbb{N}\}$ et considérons $s \in \llbracket 0, r - 1 \rrbracket$. On a pour tout réel x :

$$|x| \geq 1 \Rightarrow |x|^s \leq |x|^r \quad \text{et} \quad |x| \leq 1 \Rightarrow |x|^s \leq 1,$$

d'où :

$$\forall n \in \mathbb{N}, |x_n|^s \leq 1 + |x_n|^r.$$

Par hypothèse, les séries $\sum \mathbb{P}(X = x_n)$ et $\sum |x_n|^r \mathbb{P}(X = x_n)$ donc, par majoration, la série $\sum |x_n|^s \mathbb{P}(X = x_n)$ converge, ce qui signifie que X admet un moment d'ordre s .

- 4 ▶ Si X admet un moment d'ordre r alors la variable centrée associée $X - \mathbb{E}(X)$ admet également un moment à l'ordre r .

En effet, il suffit de remarquer que :

$$|x_n - \mathbb{E}(X)|^r \leq \sum_{k=0}^r \binom{r}{k} |x_n|^k |\mathbb{E}(X)|^{r-k}$$

et d'utiliser le fait qu'une combinaison linéaire de séries (absolument) convergentes est (absolument) convergente.

Définition V-8

Soit X une variable aléatoire réelle discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Si X admet un moment d'ordre 2 alors le moment d'ordre 2 de la variable centrée associée est appelé **variance de** X et est noté $\mathbb{V}(X)$ ou $\text{Var}(X)$:

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

Remarques

- 1 ▶ La variance mesure la dispersion (quadratique) autour de la moyenne (l'espérance).
 2 ▶ Pour tout $x \in X(\Omega)$, on a $(x - \mathbb{E}(X))^2 \geq 0$ donc $\mathbb{V}(X) \geq 0$.
 On appelle **écart-type** de X le réel $\sigma(X) = \sqrt{\mathbb{V}(X)}$.
 3 ▶ On a $\mathbb{V}(X) = 0$ si et seulement si X est constante presque sûrement ; dans ce cas, on a $X = \mathbb{E}(X)$ presque sûrement.

En effet, notons $X(\Omega) = \{x_n; n \in \mathbb{N}\}$.

Si $X = m$ presque sûrement alors on a $\mathbb{P}(X = x) = 0$ dès que $x \neq m$ ce qui donne $m = \mathbb{E}(X)$.

On a $\mathbb{P}((X - \mathbb{E}(X))^2) = \mathbb{P}(X = m) = 1$ d'où $\mathbb{V}(X) = 0$.

Réciproquement, si $\mathbb{V}(X) = 0$ alors :

$$\sum_{n=0}^{+\infty} \underbrace{(x_n - \mathbb{E}(X))^2}_{\geq 0} \underbrace{\mathbb{P}(X = x_n)}_{\geq 0} = 0,$$

ce qui donne pour tout $n \in \mathbb{N}$:

$$x_n = \mathbb{E}(X) \text{ ou } \mathbb{P}(X = x_n) = 0.$$

On en déduit que $\mathbb{P}(X = \mathbb{E}(X)) = 1$ donc X est constante, égale à $\mathbb{E}(X)$, presque sûrement.

4 ▶ Pour tous réels a et b , on a : $\mathbb{V}(aX + b) = a^2 \mathbb{V}(X)$.

En effet, on a :

$$\begin{aligned} (aX + b - \mathbb{E}(aX + b))^2 &= (aX + b - (a\mathbb{E}(X) + b))^2 \\ &= (a(X - \mathbb{E}(X)))^2 \\ &= a^2(X - \mathbb{E}(X))^2, \end{aligned}$$

d'où :

$$\mathbb{E}((aX + b - \mathbb{E}(aX + b))^2) = a^2 \mathbb{E}((X - \mathbb{E}(X))^2),$$

c'est-à-dire :

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

5 ▶ Soit X une variable aléatoire discrète admettant un moment d'ordre 2 donc un moment d'ordre 1. Notons $X(\Omega) = \{x_k; k \in \mathbb{N}\}$, on a la convergence (absolue) des séries :

$$\sum_{k \geq 0} x_k \mathbb{P}(X = x_k) \text{ et } \sum_{k \geq 0} x_k^2 \mathbb{P}(X = x_k),$$

On en déduit que, pour tout $(a, b, c) \in \mathbb{R}^3$, $aX^2 + bX + c$ admet une espérance et :

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

Par exemple, $X(X - 1)$ admet une espérance et l'on a :

$$\mathbb{E}(X(X - 1)) = \mathbb{E}(X^2 - X) = \mathbb{E}(X^2) - \mathbb{E}(X).$$

6 ▶ On en déduit la **Formule de Koenig-Huygens** :

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

En effet, on a :

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

d'où :

$$\begin{aligned} \mathbb{E}((X - \mathbb{E}(X))^2) &= \mathbb{E}(X^2 - 2\mathbb{E}(X)X + \mathbb{E}(X)^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2. \end{aligned}$$

7 ▶ On utilise fréquemment une relation du même type :

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

Cela résulte des deux remarques précédentes.

8 ▶ Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2 et non presque sûrement constante.

◦ On dit X est **centrée réduite** si $\mathbb{E}(X) = 0$ et $\mathbb{V}(X) = 1$.

◦ La variable aléatoire $X^* = \frac{X - \mathbb{E}(X)}{\sigma(X)}$ est appelée la **v.a.r.d. centrée réduite associée à X** .

Cette terminologie est cohérente puisque :

$$\mathbb{E}(X^*) = \mathbb{E}\left(\frac{1}{\sigma(X)}X - \frac{1}{\sigma(X)}\mathbb{E}(X)\right) = \frac{1}{\sigma(X)}\mathbb{E}(X) - \frac{1}{\sigma(X)}\mathbb{E}(X) = 0,$$

et :

$$\mathbb{V}(X^*) = \mathbb{V}\left(\frac{1}{\sigma(X)}X - \frac{1}{\sigma(X)}\mathbb{E}(X)\right) = \frac{1}{\sigma^2(X)}\mathbb{V}(X) = 1.$$

⇒ Cas des lois discrètes usuelles

- 1 ▶ Si X est une variable aléatoire certaine, ou presque sûrement constante, égale à m alors : $\mathbb{E}(X) = m$ et $\mathbb{V}(X) = 0$.
- 2 ▶ Si $X \hookrightarrow \mathcal{B}(p)$, alors : $\mathbb{E}(X) = p$ et $\mathbb{V}(X) = p(1-p)$.
- 3 ▶ Si $X \hookrightarrow \mathcal{B}(n, p)$, alors : $\mathbb{E}(X) = np$ et $\mathbb{V}(X) = np(1-p)$.
- 4 ▶ Si $X \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$, alors : $\mathbb{E}(X) = \frac{n+1}{2}$ et $\mathbb{V}(X) = \frac{n^2-1}{12}$.
- 5 ▶ Si $X \hookrightarrow \mathcal{G}(p)$, alors X admet une espérance et une variance et : $\mathbb{E}(X) = \frac{1}{p}$ et $\mathbb{V}(X) = \frac{1-p}{p^2}$.
- 6 ▶ Si $X \hookrightarrow \mathcal{P}(\lambda)$, alors X admet une espérance et une variance et : $\mathbb{E}(X) = \mathbb{V}(X) = \lambda$.

C.3 - Espérance conditionnelle

Définition V-9

Soit X une variable aléatoire sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et A un événement de probabilité non nulle.

1. La **loi de X sachant A** ou **loi de X conditionnellement à A** , la loi de X pour la probabilité $\mathbb{P}_A$.

Dans le cas où A est discrète, cela revient donc à la donnée des $\mathbb{P}_A(X = x)$ pour tout $x \in X(\Omega)$.

2. L'**espérance de X sachant A** est définie, sous réserve d'existence, comme l'espérance de X pour la probabilité $\mathbb{P}_A$, elle est notée $\mathbb{E}(X | A)$.

3. Dans le cas où X est une variable discrète, si la série $\sum_{x \in X(\Omega)} x \mathbb{P}_A(X = x)$ converge absolument alors X admet une espérance sachant A et :

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

Exemple

Soit X une variable aléatoire renvoyant la valeur d'un dé usuel équilibré et $A = \{1, 3, 5\}$.

Alors :

$$\begin{aligned}\mathbb{E}(X | A) &= \sum_{i=1}^6 i \cdot \mathbb{P}_A(X = i) \\ &= 1 \cdot \mathbb{P}_A(X = 1) + 0 + 3 \cdot \mathbb{P}_A(X = 3) + 0 + 5 \cdot \mathbb{P}_A(X = 5) + 0 \\ &= 1 \cdot \frac{1}{3} + 3 \cdot \frac{1}{3} + 5 \cdot \frac{1}{3} \\ &= 3.\end{aligned}$$

et :

$$\begin{aligned}\mathbb{E}(X | \bar{A}) &= \sum_{i=1}^6 i \cdot \mathbb{P}_{\bar{A}}(X = i) \\ &= 0 + 2 \cdot \mathbb{P}_{\bar{A}}(X = 2) + 0 + 4 \cdot \mathbb{P}_{\bar{A}}(X = 4) + 0 + 6 \cdot \mathbb{P}_{\bar{A}}(X = 6) \\ &= 2 \cdot \frac{1}{3} + 4 \cdot \frac{1}{3} + 6 \cdot \frac{1}{3} \\ &= 4.\end{aligned}$$

Théorème V-10 (formule de l'espérance totale)

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

On considère un système complet d'événements $(A_n)_{n \in \mathbb{N}}$ avec : $\forall n \in \mathbb{N}, \mathbb{P}(A_n) \neq 0$.

Alors X admet une espérance si et seulement si :

- ▷ pour tout $n \in \mathbb{N}$, l'espérance conditionnelle $\mathbb{E}(X | A_n)$ existe ;
- ▷ la série $\sum_{n \in \mathbb{N}} \mathbb{E}(|X| | A_n) \mathbb{P}(A_n)$ converge.

Dans ce cas, on a :

$$\mathbb{E}(X) = \sum_{n \in \mathbb{N}} \mathbb{E}(X | A_n) \mathbb{P}(A_n).$$

Démonstration

- Pour comprendre l'idée, considérons le cas où $X(\Omega) = \{x_1, \dots, x_k\}$ est fini et où le système complet d'événements est $A_1, \dots, A_n$. Dans ce cas, toutes les sommes sont finies donc existent et on a :

$$\begin{aligned}\sum_{i=1}^n \mathbb{P}(A_i) \mathbb{E}(X | A_i) &= \sum_{i=1}^n \mathbb{P}(A_i) \sum_{j=1}^k j \mathbb{P}_{A_i}(X = j) \quad (\text{définition de l'espérance conditionnelle}) \\ &= \sum_{j=1}^k \sum_{i=1}^n j \mathbb{P}(A_i) \mathbb{P}_{A_i}(X = j) \quad (\text{inversion de l'ordre de sommation}) \\ &= \sum_{j=1}^k j \left(\sum_{i=1}^n \mathbb{P}(A_i) \mathbb{P}_{A_i}(X = j) \right) \quad (\text{car } j \text{ ne dépend pas de } i) \\ &= \sum_{j=1}^k j \mathbb{P}(X = j) \quad (\text{formule des probabilités totales}) \\ &= \mathbb{E}(X).\end{aligned}$$

► Pour le cas général, on note $X(\Omega) = \{x_k; k \in \mathbb{N}\}$ et on pose pour tout $(n, k) \in \mathbb{N}^2$:

$$u_{n,k} = x_k \mathbb{P}(A_n) \mathbb{P}_{A_n}(X = x_k) = x_k \mathbb{P}(A_n \cap (X = x_k)).$$

Tout d'abord, d'après la formule des probabilités totales avec le s.c.e. formé par les A_n , on a la convergence absolue de la série $\sum u_{n,k}$ et :

$$\sum_{n=0}^{+\infty} |u_{n,k}| = \sum_{n=0}^{+\infty} |x_k| \mathbb{P}(A_n \cap (X = x_k)) = |x_k| \mathbb{P}(X = x_k).$$

$\Rightarrow$ Supposons que X admette une espérance *i.e.* la série $\sum |x_k| \mathbb{P}(X = x_k)$ converge.

Pour tout $(n, k) \in \mathbb{N}^2$, on a :

$$0 \leq |x_k| \mathbb{P}_{A_n}(X = x_k) = \frac{1}{\mathbb{P}(A_n)} |x_k| \mathbb{P}(A_n \cap (X = x_k)) \leq \frac{1}{\mathbb{P}(A_n)} |x_k| \mathbb{P}(X = x_k)$$

donc, par théorème de comparaison pour les séries à termes positifs, la série $\sum_k |x_k| \mathbb{P}_{A_n}(X = x_k)$ converge *i.e.* $\mathbb{E}(X | A_n)$ existe.

De plus, par hypothèse, la série $\sum_k |x_k| \mathbb{P}(X = x_k)$ converge *i.e.* la série double $\sum_k \sum_n |u_{n,k}|$ converge donc, par le théorème de Fubini :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=0}^{+\infty} x_k \mathbb{P}(X = x_k) \\ &= \sum_{k=0}^{+\infty} x_k \sum_{n=0}^{+\infty} \mathbb{P}(A_n) \mathbb{P}_{A_n}(X = x_k) \\ &= \sum_{k=0}^{+\infty} \sum_{n=0}^{+\infty} u_{n,k} \\ &= \sum_{n=0}^{+\infty} \sum_{k=0}^{+\infty} u_{n,k} \\ &= \sum_{n=0}^{+\infty} \mathbb{P}(A_n) \sum_{k=0}^{+\infty} x_k \mathbb{P}_{A_n}(X = x_k) \\ &= \sum_{n=0}^{+\infty} \mathbb{P}(A_n) \mathbb{E}(X | A_n). \end{aligned}$$

$\Leftarrow$ Supposons que les espérances conditionnelles existent et que la série $\sum_{n \in \mathbb{N}} \mathbb{E}(|X| | A_n) \mathbb{P}(A_n)$ converge.

Par hypothèse, l'espérance conditionnelle $\mathbb{E}(X | A_n)$ existe ce qui signifie que la série $\sum_k x_k \mathbb{P}_{A_n}(X = x_k)$ converge absolument donc la série $\sum_k |x_k| \mathbb{P}_{A_n}(X = x_k)$ converge ce qui signifie que $\mathbb{E}(|X| | A_n)$ existe et :

$$\mathbb{E}(|X| | A_n) = \sum_{k=0}^{+\infty} |u_{n,k}|.$$

On a donc la convergence de la série double :

$$\sum_n \sum_k |u_{n,k}|$$

donc le théorème de Fubini donne la convergence de :

$$\sum_k \sum_n |u_{n,k}| \quad i.e. \quad \sum_k |x_k| \mathbb{P}(X = x_k)$$

donc $\mathbb{E}(X)$ existe et le même calcul que pour l'autre implication donne :

$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n) \mathbb{E}(X | A_n).$$

Exemple

Le nombre de clients qui entrent dans un restaurant un soir de semaine est une variable aléatoire X qui suit une loi de Poisson de paramètre λ . Chaque client choisit, indépendamment du choix de ses voisins, de la viande avec probabilité p et du poisson avec la probabilité $1 - p$. On note Y le nombre de clients qui choisissent de la viande.

La famille $\{(X = n); n \in \mathbb{N}\}$ est un système complet d'événements de probabilités non nulles donc la formule des probabilités totales donne pour tout $k \in \mathbb{N}$:

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

La loi de Y sachant $(X = n)$ est la loi $\mathcal{B}(n, p)$:

$$\mathbb{P}_{(X=n)}(Y = k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{si } 0 \leq k \leq n \\ 0 & \text{sinon} \end{cases}.$$

On a donc pour tout $k \in \mathbb{N}$:

$$\begin{aligned} \mathbb{P}(Y = k) &= \sum_{n=k}^{+\infty} e^{-\lambda} \frac{\lambda^n}{n!} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{n=k}^{+\infty} e^{-\lambda} \frac{\lambda^n}{n!} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= e^{-\lambda} p^k \frac{\lambda^k}{k!} \sum_{n=k}^{+\infty} \lambda^{n-k} \frac{1}{(n-k)!} (1-p)^{n-k} \\ &= e^{-\lambda} p^k \frac{\lambda^k}{k!} \sum_{n=0}^{+\infty} \lambda^n \frac{1}{n!} (1-p)^n \\ &= e^{-\lambda} \frac{(p\lambda)^k}{k!} e^{\lambda(1-p)} \\ &= e^{-\lambda p} \frac{(p\lambda)^k}{k!} \end{aligned}$$

donc Y suit la loi $\mathcal{P}(\lambda p)$.

La loi de Y sachant $(X = n)$ est la loi $\mathcal{B}(n, p)$, montrons par le calcul que $\mathbb{E}(Y | X = n) = np$:

$$\begin{aligned}
 \mathbb{E}(Y | X = n) &= \sum_{k=0}^{+\infty} k \mathbb{P}_{(X=n)}(Y = k) \\
 &= \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\
 &= \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} \\
 &= \sum_{k=1}^n n \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
 &= \sum_{\ell=0}^{n-1} n \binom{n-1}{\ell} p^{\ell+1} (1-p)^{n-1-\ell} \\
 &= np \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} p^{\ell} (1-p)^{n-1-\ell} \\
 &= np (p + (1-p))^{n-1} \\
 &= np.
 \end{aligned}$$

Puisque $Y \hookrightarrow \mathcal{P}(p\lambda)$, on a $\mathbb{E}(Y) = p\lambda$. Retrouvons ce résultat par le calcul. Pour tout $n \geq 1$, on a :

$$\mathbb{P}(X = n) \mathbb{E}(Y | X = n) = n p e^{-\lambda} \frac{\lambda^n}{n!} = p \lambda e^{-\lambda} \frac{\lambda^{n-1}}{(n-1)!}$$

or la série $\sum_{n \geq 1} \frac{\lambda^{n-1}}{(n-1)!}$ converge absolument. D'après le théorème de l'espérance totale, Y admet donc une espérance et :

$$\begin{aligned}
 \mathbb{E}(Y) &= \sum_{n=0}^{+\infty} \mathbb{P}(X = n) \mathbb{E}(Y | X = n) \\
 &= \sum_{n=1}^{+\infty} p \lambda e^{-\lambda} \frac{\lambda^{n-1}}{(n-1)!} \\
 &= p \lambda e^{-\lambda} \sum_{n=1}^{+\infty} \frac{\lambda^{n-1}}{(n-1)!} \\
 &= p \lambda e^{-\lambda} e^{\lambda} \\
 &= p \lambda.
 \end{aligned}$$