10 Fondements des probabilités

« Un coup de dés jamais n'abolira le hasard »
Stéphane Mallarmé (1914)

Plan de cours

I | Quelques rappels de dénombrement

A – Cardinal d'un ensemble fini

On rappelle que le cardinal d'un ensemble fini E est noté $\operatorname{card}(E)$ ou $|E|$. On notera dans tout ce chapitre $A \sqcup B$ l'union de deux ensembles A et B disjoints (notation courante mais non officielle).
Tous les ensembles considérés dans cette section sont supposés finis. Soient $n, p \in \mathbb{N}^*$.

Proposition 10.1 : cardinal d'une réunion / d'une différence

Soient deux ensembles finis A et B.

(i) $\operatorname{card}(A \cup B) = \operatorname{card}(A) + \operatorname{card}(B) - \operatorname{card}(A \cap B)$. En particulier, $\operatorname{card}(A \sqcup B) = \operatorname{card}(A) + \operatorname{card}(B)$.

(ii) $\operatorname{card}(A \setminus B) = \operatorname{card}(A) - \operatorname{card}(A \cap B)$;

Plus généralement, si $A_1, \dots, A_n$ sont disjoints, alors $\operatorname{card}\left(\bigsqcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \operatorname{card}(A_k)$.

Proposition 10.2 : cardinal d'un produit

Soient p ensembles finis $A_1, \dots, A_p$. Alors, $\operatorname{card}\left(\prod_{k=1}^{p} A_k\right) = \prod_{k=1}^{n} \operatorname{card}(A_k)$.

En particulier, $\operatorname{card}\left(E^p\right) = \operatorname{card}(E)^p$.

Proposition 10.3

Soient deux ensembles finis E et F.

(i) S'il existe une injection de E dans F alors $\operatorname{card}(E) \leqslant \operatorname{card}(F)$.

(ii) S'il existe une surjection de E dans F alors $\operatorname{card}(E) \geqslant \operatorname{card}(F)$.

(iii) S'il existe une bijection de E dans F alors $\operatorname{card}(E) = \operatorname{card}(F)$.

Si $\operatorname{card}(E) = \operatorname{card}(F)$, $f : E \to F$ est injective si et seulement si elle est surjective.

B – Listes, arrangements, combinaisons et permutations

La résolution d'un problème de dénombrement conduit bien souvent à représenter une situation donnée au moyen de graphes, de mots ou de tout autre objet permettant d'énumérer une certaine *liste de possibilités*. Une fois traduit en langage mathématique, la résolution d'un tel problème se ramènera immanquablement au calcul du cardinal d'un ensemble bien identifié. Cette étape de modélisation est essentielle au calcul de certaines probabilités.

Exemples

- Une urne contient des boules noires et des boules blanches. Le tirage successif d'une boule blanche, puis d'une boule noire et enfin d'une boule blanche peut être représenté par le mot BNB. L'ensemble des tirages possibles de trois boules est alors : $\{BBB, NBB, BNB, NNB, NBN, BBN, BNN, NNN\}$.
On constate qu'il y a $8 = 2 \times 2 \times 2$ possibilités.
- On pioche cinq cartes simultanément dans un jeu de 32 cartes. Une « main » possible est : $\{\text{As}\heartsuit, \text{R}\clubsuit, 8\clubsuit, \text{D}\diamondsuit, 9\heartsuit\}$. Notons que contrairement au cas précédent, l'ordre n'a pas d'importance.

Bien entendu, il sera rarement question de proposer une énumération exhaustive des différentes possibilités pour déterminer le cardinal d'un ensemble. Il s'agira plutôt de reformuler un problème donné pour se ramener à des situations familières, s'appuyant (par exemple) sur l'un des trois modèles suivants :

- tirages successifs avec remise (modélisés par des listes)
- tirages successifs sans remise (modélisés par des arrangements)
- tirages simultanés (modélisés par des combinaisons)

Revenons sur ces éléments, en considérant un ensemble fini E à n éléments ainsi qu'un entier naturel p.

1 – Listes et uplets

Définition 10.4

On appelle p-uplet ou p-liste de E toute famille de p éléments de E, c'est-à-dire un élément de E^p.

Il y a exactement n^p p-uplets de E.

Exemple

On tire trois cartes avec remise dans un jeu comportant 32 cartes. Il y a alors 32^3 tirages possibles.

Proposition 10.5

Le nombre d'applications d'un ensemble E de cardinal n vers un ensemble F de cardinal p vaut n^p.

2 – Arrangements

Définition 10.6 : Arrangement

On appelle arrangement de p éléments de E un p-uplet constitué d'éléments de E deux à deux distincts.

Exemple

Si $E = \{1, 2, 3\}$, les arrangements de deux éléments sont : $(1,2), (1,3), (2,1), (2,3), (3,1)$ et $(3,2)$.

Théorème 10.7 : Nombre d'arrangements

On suppose que $1 \leqslant p \leqslant n$.

Le nombre d'arrangements de p éléments de E vaut $\dfrac{n!}{(n-p)!} = n \times (n-1) \times \cdots \times (n-p+1)$.

C'est le nombre de façons de choisir p éléments ordonnés parmi n.

Démonstration

On peut procéder par récurrence ou, plus simplement, de la façon suivante.
Choisir un arrangement $(x_1, \ldots, x_p)$ de E revient à choisir x_1 parmi n éléments, puis x_2 parmi $n-1$ éléments, ..., jusqu'à x_p parmi les $n-p+1$ éléments restants. Au final, $n \times (n-1) \times \cdots \times (n-p+1)$ possibilités. ■

Exercice 1

Une association comportant 27 membres doit élire un président, un secrétaire et un trésorier. Quel est le nombre de possibilités ?

À chaque arrangement $(x_1, \dots, x_p)$ de E, on peut associer de manière unique l'application $f : [\![1, p]\!] \to E$ définie par $f(i) = x_i$. L'application est injective, ce qui conduit au résultat suivant.

Proposition 10.8

Si E est de cardinal p et F de cardinal n, le nombre d'applications injectives de E dans F est $\dfrac{n!}{(n-p)!}$.

3 – Permutations

Définition 10.9 : Permutation

On appelle permutation de E un arrangement de E à $n = \text{card}(E)$ éléments.

Exemple

Si $E = \{1, 2, 3\}$, les permutations de E sont $(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2)$ et $(3, 2, 1)$.

Proposition 10.10 : Nombre de permutations

Le nombre de permutations de l'ensemble E est $n!$.

À chaque permutation $(x_1, \dots, x_n)$ de E, on peut associer de manière unique l'application $f : [\![1, n]\!] \to E$ définie par $f(i) = x_i$. L'application étant injective, elle est bijective ; d'où le résultat suivant.

Proposition 10.11

Il y a $n!$ bijections de E vers E.

Exercice 2

- Combien existe-t-il d'anagrammes du mot *SABORD* ? du mot *BACHIBOUZOUK* ?
- Combien peut-on fabriquer de colliers distincts avec 5 perles de couleurs différentes ?

4 – Combinaisons

Définition 10.12 : Combinaison

On appelle combinaison de p éléments de E un sous-ensemble de E contenant p éléments.

Attention à la différence de nature entre un arrangement (une liste d'éléments de E) et une combinaison (une partie de E). L'ordre n'important pas dans la constitution d'une combinaison, les combinaisons seront bien moins nombreuses que les arrangements.

Exemple

Si $E = \{1, 2, 3\}$, les combinaisons de deux éléments sont : $\{1, 2\}, \{1, 3\}$ et $\{2, 3\}$.

Théorème 10.13 : Nombre de combinaisons

On suppose que $1 \leqslant p \leqslant n$.

Le nombre de combinaisons de p éléments de E vaut $\dbinom{n}{p} = \dfrac{n!}{p!(n-p)!}$.

Démonstration

Considérons une combinaison $\{x_1,\ldots,x_p\}$ de p éléments de E. On peut alors lui associer $p!$ arrangements distincts. Réciproquement, à un arrangement donné, on ne peut lui associer qu'une seule combinaison. Si on note C_n^p le nombre de combinaisons de E à p éléments, $\dfrac{n!}{(n-p)!}=p!C_n^p$ et donc : $C_n^p=\dfrac{n!}{p!(n-p)!}=\dbinom{n}{p}$.

Exercice 3

On dispose d'un jeu classique de 32 cartes et on en distribue 8 à 4 joueurs.

- Combien y a-t-il de jeux possibles par joueur ?
- Combien y a-t-il de jeux contenant 6 cartes rouges ?

Proposition 10.14 : Propriétés des coefficients binomiaux

Soient $n,p\in\mathbb{N}$, supposés non nuls dans les formules (iii) et (iv).

$$\text{(i)}\quad \binom{n}{0}=\binom{n}{n}=1,\binom{n}{1}=n \qquad \text{(ii)}\quad \binom{n}{p}=\binom{n}{n-p} \qquad \text{(iii)}\quad p\binom{n}{p}=n\binom{n-1}{p-1}$$

$$\text{(iv)}\quad \binom{n-1}{p-1}+\binom{n-1}{p}=\binom{n}{p} \qquad \text{(v)}\quad \sum_{k=0}^{n}\binom{n}{k}=2^n$$

Démonstration (preuves combinatoires)

(i) Il y a n choix possibles quand on tire 1 élément parmi n donc $\dbinom{n}{1}=n$.

(ii) Choisir p éléments parmi n revient à choisir les $n-p$ autres donc $\dbinom{n}{p}=\dbinom{n}{n-p}$.

(iii) Formule du capitaine – Il y a $p\dbinom{n}{p}$ façon de choisir p éléments parmi n puis un élément parmi ceux-là. Cela revient exactement à choisir un élement parmi les n puis à « compléter » en choisissant $p-1$ éléments parmi les $n-1$ restants. Ce qui nous donne bien la formule annoncée !

(iv) Formule de Pascal – Soient E un ensemble à n éléments et $x\in E$. Les parties de E à p éléments sont de deux types : celles qui contiennent x et sont constituées de $p-1$ autres éléments choisis parmi les $n-1$ restants ; celles qui ne contiennent pas x et qui sont constituées de p éléments choisis parmi les $n-1$ restants. Elles forment deux ensembles disjoints donc $\binom{n-1}{p-1}+\binom{n-1}{p}=\binom{n}{p}$.

(v) Toute partie de E étant une combinaison de E, les ensembles de combinaisons de p éléments de E constituent une partition de $\mathscr{P}(E)$, ce qu'on peut écrire sous la forme $\mathscr{P}(E)=\displaystyle\bigsqcup_{p=1}^{n}\mathscr{C}^p(E)$ où $\mathscr{C}^p(E)$ désigne l'ensemble des combinaisons de p éléments. D'où l'égalité demandée. ■

Pour $p>n$, $\dbinom{n}{p}=0$: on ne peut construire d'ensemble à p éléments avec seulement n éléments.

Exercice 4 – Formule de Vandermonde

Soient $m,n,p\in\mathbb{N}$. Montrer à l'aide d'une preuve combinatoire et d'une preuve non combinatoire que :

$$\sum_{k=0}^{p}\binom{n}{k}\binom{m}{p-k}=\binom{n+m}{p}$$

Proposition 10.15 : Formule du binôme

Pour tous $a,b\in\mathbb{C}$ et $n\in\mathbb{N}$, $(a+b)^n=\displaystyle\sum_{k=0}^{n}\binom{n}{k}a^kb^{n-k}$.

Corollaire 10.16 : Cardinal de l'ensemble des parties

Si E est un ensemble fini de cardinal n, $\operatorname{card}(\mathscr{P}(E)) = 2^n$.

Exercice 5

Soient $n, p \in \mathbb{N}^*$. Déterminer le nombre :

(i) de p-listes strictement croissantes d'éléments de $[\![1,n]\!]$.

(ii) de p-listes croissantes d'éléments de $[\![1,n]\!]$.

(iii) de listes $(n_1,\ldots,n_p)$ d'entiers positifs ou nuls tels que $n_1+\cdots+n_p = n$.

(iv) de listes $(n_1,\ldots,n_p)$ d'entiers strictement positifs tels que $n_1+\cdots+n_p = n$.

C – Indicatrice

On rappelle que si A est une partie de E, on définit la fonction indicatrice (ou caractéristique) de A sur E par :

$$\forall x \in E, \quad \mathbb{1}_A(x) = \begin{cases} 1 \text{ si } x \in A \\ 0 \text{ si } x \notin A \end{cases}$$

Les opérations sur les ensembles peuvent se traduire par des opérations sur les indicatrices.

Proposition 10.17

Soient A et B deux parties d'un ensemble E, A étant supposée finie dans la dernière assertion. Alors,

(i) $A = B \iff \mathbb{1}_A = \mathbb{1}_B$ (ii) $A \subset B \iff \mathbb{1}_A \leqslant \mathbb{1}_B$ (iii) $\mathbb{1}_{\overline{A}} = 1 - \mathbb{1}_A$

(iv) $\mathbb{1}_{A\cap B} = \mathbb{1}_A \cdot \mathbb{1}_B$ (v) $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \cdot \mathbb{1}_B$ (vi) $\operatorname{card}(A) = \displaystyle\sum_{x\in E} \mathbb{1}_A(x)$

Exemple

$(A_i)_{1\leqslant i\leqslant n}$ est une partition de E si et seulement si $\mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_n} = 1$.

Exercice 6 – Formule du crible de Poincaré

Soient $A_1,\ldots,A_n$ des ensembles finis. Prouver à l'aide d'indicatrices que :

$$\operatorname{card}\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k+1} \left[\sum_{1\leqslant i_1<\cdots<i_k\leqslant n} \operatorname{card}\left(A_{i_1}\cap\cdots\cap A_{i_k}\right)\right]$$

II | Des tribus aux probabilités

A – Expérience aléatoire, univers

On appelle expérience aléatoire une expérience dont toutes les issues possibles sont connues *a priori* mais dont le résultat peut varier lorsqu'on la répète (lancer d'un dé, tirage d'une boule dans une urne, ...). On appelle alors univers et on note en général Ω l'ensemble des résultats possibles – ou issues – d'une telle expérience. Ces résultats peuvent prendre des formes variées :

- lancer d'un dé : on peut s'intéresser au chiffre obtenu ($\Omega = [\![1,6]\!]$), à sa parité ($\Omega = \{P, I\}$), etc.
- lancer de deux dés discernables : $\Omega = [\![1,6]\!]^2$;
- tirage de trois cartes dans un jeu de 32 : Ω est alors l'ensemble des combinaisons de 3 cartes parmi 32.

Une expérience aléatoire peut également conduire à un univers infini, dénombrable ou non dénombrable :

- lancer d'une pièce jusqu'à obtenir pile : $\Omega = \{P, FP, FFP, FFFP, \ldots\} \cup \{\omega_\infty\}$ (dénombrable) où ω_∞ désigne le mot infini dont toutes les lettres valent F ;

- lancer infini d'une pièce : $\Omega = \{P, F\}^{\mathbb{N}}$ (non dénombrable (!)) ;
- durée de vie d'une ampoule : $\Omega = \mathbb{R}_+^*$ (non dénombrable) ;
- jeu de fléchettes : $\Omega = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leqslant R^2\}$ (non dénombrable).

Il s'avère parfois difficile d'expliciter l'univers associé à une expérience aléatoire (mutation de gènes d'une population de drosophiles, files d'attente, etc.) mais cela ne constituera pas pour nous un réel obstacle. S'il peut être instructif de préciser l'univers manipulé dans certains exercices de probabilités élémentaires (essentiellement ceux qui se rapportent à des questions combinatoires), l'univers sera invisibilisé dans la plupart des problèmes rencontrés, au profit de son image par une variable aléatoire, seule trace digne d'intérêt.

B – Tribu et événements

Le programme de première année se limite aux expériences aléatoires associées à des univers finis. Dans ce cadre, on peut écrire $\Omega = \{\omega_1, \dots, \omega_n\}$ et appeler événement toute partie de Ω ; $\mathscr{P}(\Omega)$ représentant alors l'ensemble des événements associés à une expérience aléatoire donnée. Tout événement s'écrit comme réunion finie d'événements élémentaires, c'est-à-dire d'événements de la forme $\{\omega\}$ avec $\omega \in \Omega$.

Exemple

On lance un dé et on obtient 5. $\Omega = [\![1, 6]\!]$.

- L'événement « face impaire » décrit par $A = \{1, 3, 5\}$ est donc réalisé.
- L'événement « score supérieur ou égal à 3 » décrit par $A = \{3, 4, 5, 6\}$ est réalisé.

Rappelons quelques éléments de vocabulaire probabiliste, simples traductions de termes ensemblistes.

Langage probabiliste	Langage ensembliste
Résultat possible	$\omega \in \Omega$
Événement	$A \subset \Omega$
Événement certain	Ω
Événement impossible	$\varnothing$
Événement contraire	$\overline{A}$
Événement « A ou B »	$A \cup B$
Événement « A et B »	$A \cap B$
Événements incompatibles	$A \cap B = \varnothing$
Système complet d'événements	$\Omega = \displaystyle\bigsqcup_{k=1}^{n} A_k$

Vocabulaire associé aux événements

Munir Ω d'une probabilité revient à définir sur $\mathscr{P}(\Omega)$ une application $\mathbf{P}$ à valeurs dans $[0, 1]$ vérifiant $\mathbf{P}(\Omega) = 1$ et une propriété d'additivité : $\mathbf{P}(A \sqcup B) = \mathbf{P}(A) + \mathbf{P}(B)$ pour tous événements A et B incompatibles. On remarque en première année que définir une telle application revient à se donner une famille de réels positifs $(p_k)_{1 \leqslant k \leqslant n}$, de somme égale à 1. $\mathbf{P}$ est en effet entièrement définie par la donnée de $\mathbf{P}(\{\omega_k\}) = p_k$, la probabilité de toute partie de $\mathscr{P}(\Omega)$ s'obtenant alors en écrivant cette partie comme réunion disjointe d'événements élémentaires.

Lorsque l'univers Ω est dénombrable, même combat ! La description de l'univers sous la forme $\Omega = \{\omega_n \mid n \in \mathbb{N}\}$ nous permet d'écrire toute partie de Ω comme réunion (cette fois-ci dénombrable) d'événements élémentaires. Il ne nous reste plus qu'à introduire une probabilité comme application $\mathbf{P} : \mathscr{P}(\Omega) \to [0, 1]$ vérifiant :

$$\mathbf{P}(\Omega) = 1 \quad \text{et} \quad \mathbf{P}\left(\bigsqcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbf{P}(A_n)$$

La deuxième propriété, dite d'additivité dénombrable ou σ-additivité, s'avère indispensable si l'on souhaite sortir du cadre fini et effectuer des passages à la limite (probabilité d'« obtenir une infinité de *pile* » par exemple). Se doter d'une telle probabilité revient encore à se donner une famille de réels positifs $(p_k)_{k\in\mathbb{N}}$, de somme égale à 1. La probabilité de toute partie de Ω, c'est-à-dire de tout événement, se retrouve alors au moyen de :

$$\forall A \in \mathscr{P}(\Omega), \quad \mathbf{P}(A) = \sum_{\omega_k \in A} \mathbf{P}(\{\omega_k\})$$

En revanche, lorsque l'univers Ω n'est pas dénombrable, un fait nouveau et déroutant se produit, tout droit issu de la théorie de la mesure : toutes les parties de Ω ne peuvent être considérées comme des événements car il est impossible de définir une probabilité véritablement intéressante sur $\mathscr{P}(\Omega)$.

Exemple – le jeu de fléchettes et l'axiome du choix

> On lance une fléchette sur une cible linéaire modélisée par le segment $\Omega = [0,1]$. On cherche à définir une probabilité sur Ω conforme, de façon intuitive, au point de vue fréquentiste :
>
> - La probabilité d'atteindre un point particulier, par exemple celui d'abscisse $1/\sqrt{2}$, devrait être nulle.
> - La probabilité d'atteindre un point d'abscisse située entre 0 et $1/2$ devrait valoir $1/2$. Plus généralement, la probabilité d'atteindre un point du segment $[a,b]$, pour $0 \leqslant a \leqslant b \leqslant 1$, devrait valoir $b-a$.
>
> L'axiome du choix permet malheureusement d'infirmer l'existence d'une application $\mathbf{P} : \mathscr{P}([0,1]) \to [0,1]$ supposée σ-additive et vérifiant, pour tous $a \leqslant b$, $\mathbf{P}([a,b]) = b-a$. (résultat non trivial [1] et hors programme)

Ne perdons cependant pas espoir de définir une probabilité sur un univers non dénombrable. Si $\mathscr{P}(\Omega)$ est trop riche, et c'est le cas lorsque Ω n'est pas dénombrable, il suffit de se débarrasser de ses éléments les plus exotiques. L'ablation sera indolore, de telles parties de Ω ne peuvent se décrire au moyen d'une liste de parties élémentaires et d'opérations ensemblistes : inaccessibles, elles sont en quelque sorte invisibles à nos yeux.

Fixons-nous un cahier des charges pour définir une partie $\mathscr{A}$ de $\mathscr{P}(\Omega)$ qui regrouperait les parties de Ω qui seraient par la suite qualifiées d'*événements*. Quelques considérations s'imposent :

- Si A est un événement, on s'attend à ce que $\overline{A}$ soit encore un événement.
- De même, si nous disposons d'une suite (A_n) d'événements, on souhaiterait que $\bigcup_{n=0}^{+\infty} A_n$ soit encore un événement, donc un élément de $\mathscr{A}$. Cette propriété devrait s'étendre au cas des intersections dénombrables.

Ce qui nous conduit tout droit à la définition suivante :

Définition 10.18 : Tribu ou σ-algèbre

Soit Ω un ensemble non vide. On appelle tribu sur Ω, ou σ-algèbre, toute partie $\mathscr{A}$ de $\mathscr{P}(\Omega)$ qui vérifie :

(i) $\Omega \in \mathscr{A}$;

(ii) Si $A \in \mathscr{A}$ alors $\overline{A} \in \mathscr{A}$ *(stabilité par passage au complémentaire)*;

(iii) Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathscr{A}$ alors $\bigcup_{n=0}^{+\infty} A_n \in \mathscr{A}$ *(stabilité par réunion dénombrable)*.

Le couple $(\Omega, \mathscr{A})$ est alors appelé espace probabilisable; tout élément de $\mathscr{A}$ est appelé événement de Ω.

L'association des propriétés (ii) et (iii) montre qu'une tribu est stable par intersection dénombrable : si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathscr{A}$, alors $\bigcap_{n=0}^{+\infty} A_n \in \mathscr{A}$.

Exemples

> Si Ω est un ensemble non vide, voici deux exemples de tribus naturelles :
>
> $$\mathscr{A} = \{\varnothing, \Omega\} \ \textit{(tribu grossière)} \quad ; \quad \mathscr{A} = \mathscr{P}(\Omega) \ \textit{(tribu exhaustive)}$$

1. Il est néanmoins possible de définir une telle probabilité sur un ensemble plus restreint que $\mathscr{P}([0,1])$, appelé « tribu borélienne ». Cette partie de $\mathscr{P}([0,1])$ n'est rien d'autre que la tribu (voir définition ci-après) engendrée par les intervalles ouverts inclus dans $[0,1]$.

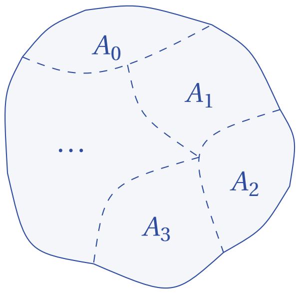
En pratique,

- si Ω est au plus dénombrable, on choisira $\mathscr{A} = \mathscr{P}(\Omega)$. La notion de tribu n'est alors plus très pertinente ;
- si Ω est non dénombrable, l'énoncé fournira le cadre adapté en admettant bien souvent l'existence d'une tribu convenable autre que $\mathscr{P}(\Omega)$. Si jamais l'énoncé s'en soucie…

Définition 10.19 : Système complet d'événements

Soit $(\Omega, \mathscr{A})$ un espace probabilisable. On appelle système complet d'événements toute famille finie ou dénombrable $(A_i)_{i\in I}$ d'événements vérifiant :

(i) pour tous $i, j \in I$ distincts, $A_i \cap A_j = \varnothing$ (ii) $\bigcup_{i\in I} A_i = \Omega$



Représentation d'un système complet d'événements

Exemples

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

- Si A est un événement, $\left(A, \overline{A}\right)$ est un s.c.e. ;
- Si l'univers $\Omega = \{\omega_1, \dots, \omega_n\}$ est fini, $(\{\omega_1\}, \dots, \{\omega_n\})$ est un système complet d'événements.
- Plus généralement, pour tout univers au plus dénombrable $\Omega = \{\omega_i \mid i \in \mathbb{N}\}$, la famille $(\{\omega_i\})_{i\in\mathbb{N}}$ est un système complet d'événements.

C – Probabilité

1 – Définition et premières propriétés

Définition 10.20 : Probabilité

Soit $(\Omega, \mathscr{A})$ un espace probabilisable, où Ω désigne un univers quelconque (fini, dénombrable ou non dénombrable). On appelle probabilité sur $(\Omega, \mathscr{A})$ toute application $\mathbf{P} : \mathscr{A} \to [0,1]$ vérifiant :

(i) $\mathbf{P}(\Omega) = 1$

(ii) Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux incompatibles,

$$\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbf{P}(A_n) \qquad \text{(propriété de } \sigma\text{-additivité)}$$

Le triplet $(\Omega, \mathscr{A}, \mathbf{P})$ est alors appelé espace probabilisé. Il est dit discret lorsque Ω est au plus dénombrable.

- La somme qui apparaît dans la propriété (ii) est celle d'une série convergente. C'est ce que suppose implicitement l'axiome de σ-additivité.
- Si on considère une suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux incompatibles telle que $A_n = \varnothing$ pour $n > p$, la propriété (ii) se traduit par :
$$\mathbf{P}\left(A_1 \cup \dots \cup A_p\right) = \mathbf{P}(A_1) + \dots + \mathbf{P}(A_p)$$
En particulier, pour tout couple (A, B) d'événements incompatibles, $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$.

On retrouve alors les propriétés établies dans le cadre fini.

Proposition 10.21

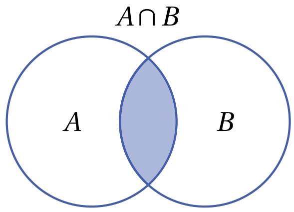
Soient $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé et $A, B \in \mathscr{A}$.

- $\mathbf{P}(\varnothing) = 0$ et $\mathbf{P}(\overline{A}) = 1 - \mathbf{P}(A)$.
- Si $A \subset B$ alors $\mathbf{P}(A) \leqslant \mathbf{P}(B)$. *(croissance de la probabilité)*
- $\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)$.

Démonstration

Démontrons seulement la dernière propriété.
$A \cup B = (A \setminus B) \sqcup (A \cap B) \sqcup (B \setminus A)$ donc :

$$\begin{aligned}\mathbf{P}(A \cup B) &= \mathbf{P}(A \setminus B) + \mathbf{P}(A \cap B) + \mathbf{P}(B \setminus A)\\ &= [\mathbf{P}(A) - \mathbf{P}(A \cap B)] + \mathbf{P}(A \cap B) + [\mathbf{P}(B) - \mathbf{P}(A \cap B)]\\ &= \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B)\end{aligned}$$



Proposition 10.22

Soient $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé, I un ensemble au plus dénombrable, et $(A_i)_{i \in I}$ un système complet d'événements. Alors, $\displaystyle\sum_{i \in I} \mathbf{P}(A_i) = 1$.

Définition 10.23

Soient $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé et A un événement.

(i) Si $\mathbf{P}(A) = 0$, l'événement A est dit négligeable ou quasi-impossible.

(ii) Si $\mathbf{P}(A) = 1$, l'événement A est dit presque sûr ou quasi-certain.

2 – Distribution de probabilités discrète

Définition 10.24

Soit Ω un ensemble. On appelle distribution de probabilités discrète sur Ω toute famille de réels positifs indexée par Ω et de somme égale à 1.

Cette famille étant sommable, son support est nécessairement au plus dénombrable (cf. chapitre *Procédés sommatoires discrets*). Dans la suite de ce paragraphe, on considérera donc uniquement des espace probabilisés $(\Omega, \mathscr{A}, \mathbf{P})$ discrets. Dans ce cadre spécifique, et comme mentionné précédemment, une probabilité est entièrement déterminée par les valeurs prises sur chaque événement élémentaire. C'est le principe mis en œuvre dans l'exercice suivant et l'objet du théorème qui suit.

Exercice 7

On dispose d'un dé pipé dont on connaît la table de probabilités.

Face obtenue	1	2	3	4	5	6
Probabilité	$\frac{1}{12}$	$\frac{1}{8}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{8}$

Montrer que cette table définit une unique probabilité sur $\Omega = [\![1, 6]\!]$.
Quelle la probabilité qu'on obtienne un chiffre supérieur à 4 lors d'un tirage ? un chiffre pair ? impair ?

Théorème 10.25

Soit Ω un ensemble au plus dénombrable.

- Soit $\mathbf{P}$ une probabilité définie sur $(\Omega, \mathscr{P}(\Omega))$. On pose, pour tout $\omega \in \Omega$, $p_\omega = \mathbf{P}(\{\omega\})$.
  Alors, $(p_\omega)_{\omega \in \Omega}$ est une distribution de probabilités.
- Si $(p_\omega)_{\omega \in \Omega}$ est une distribution de probabilités, il existe une unique probabilité $\mathbf{P}$ définie sur $(\Omega, \mathscr{P}(\Omega))$ telle que pour tout $\omega \in \Omega$, $p_\omega = \mathbf{P}(\{\omega\})$.

On établit ici une bijection entre l'ensemble des probabilités sur $(\Omega, \mathscr{P}(\Omega))$ et l'ensemble des distributions de probabilités discrètes.

Démonstration

- Pour tout $\omega \in \Omega$, $p_\omega \in [0,1]$. Par σ-additivité, (p_ω) est sommable et $\mathbf{P}(\Omega) = \mathbf{P}\left(\bigsqcup_{\omega\in\Omega}\{\omega\}\right) = \sum_{\omega\in\Omega} p_\omega = 1$.
- Réciproquement, donnons-nous une distribution de probabilités $(p_\omega)_{\omega\in\Omega}$. Soit $\mathbf{P}$ une application σ-additive définie sur $\mathscr{P}(\Omega)$ vérifiant $p_\omega = \mathbf{P}(\{\omega\})$ pour tout $\omega \in \Omega$. Nécessairement,

$$\forall A \in \mathscr{P}(\Omega), \quad \mathbf{P}(A) = \mathbf{P}\left(\bigsqcup_{\omega\in A}\{\omega\}\right) = \sum_{\omega\in A} p_\omega$$

L'unicité étant assurée, il reste à vérifier que l'on a bien défini une probabilité. On vérifie sans peine que pour tout $A \in \mathscr{P}(\Omega)$, $\mathbf{P}(A) \in [0,1]$ puis que $\mathbf{P}(\Omega) = 1$. Soit maintenant une suite $(A_n)_{n\in\mathbb{N}}$ d'événements deux à deux incompatibles, de réunion notée B. La famille $(p_\omega)_{\omega\in B}$ étant sommable, une sommation par paquets donne :

$$\mathbf{P}(B) = \mathbf{P}\left(\bigsqcup_{n=0}^{+\infty} A_n\right) = \sum_{\omega\in B} p_\omega = \sum_{n=0}^{+\infty}\sum_{\omega\in A_n} p_\omega = \sum_{n=0}^{+\infty}\mathbf{P}(A_n)$$

■

Choisir les réels p_ω revient à choisir un modèle probabiliste (associé par exemple au lancer d'un dé pipé).

Exemple

Considérons $\mathbb{N}$ muni de sa tribu naturelle $\mathscr{P}(\mathbb{N})$. On définit une probabilité sur $\mathbb{N}$ en posant pour tout $n \in \mathbb{N}$, $\mathbf{P}(\{n\}) = \left(\frac{1}{2}\right)^{n+1} = p_n$. La série *à termes positifs* $\sum p_n$ est en effet *convergente*, de somme $\frac{1}{2}\cdot\frac{1}{1-\frac{1}{2}} = 1$.

Dans le cas fini, une probabilité sort du lot : la probabilité uniforme. C'est elle que l'on considère dans le cas de dés non pipés/honnêtes/équilibrés, des pièces non truquées ou de tirages supposés équiprobables.

Définition 10.26 : Probabilité uniforme

Soit Ω un ensemble fini. On appelle probabilité uniforme sur Ω l'unique probabilité qui prend la même valeur pour chaque événement élémentaire.

Pour une telle probabilité, en posant $n = \mathrm{card}(\Omega)$, $\mathbf{P}(\{\omega\}) = \dfrac{1}{\mathrm{card}(\Omega)} = \dfrac{1}{n}$. Si A est un événement,

$$\mathbf{P}(A) = \sum_{\omega\in A}\frac{1}{n} = \frac{\mathrm{card}(A)}{\mathrm{card}(\Omega)} = \frac{\text{nombre de cas où A se réalise}}{\text{nombre de cas possibles}} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

Exercice 8

On considère une urne contenant 3 boules noires, 4 blanches et 5 rouges.

(a) On tire simultanément 3 boules. Quelle est la probabilité qu'elles soient toutes de la même couleur ? de couleurs différentes ?

(b) Reprendre les questions précédentes pour un tirage successif de trois boules avec/sans remise.

Notons qu'il est impossible de définir une probabilité uniforme sur un univers dénombrable.

3 – Théorèmes limites et applications

Proposition 10.27 : Continuité croissante

Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante d'événements (au sens de l'inclusion) sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Alors,

$$\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty}\mathbf{P}(A_n)$$

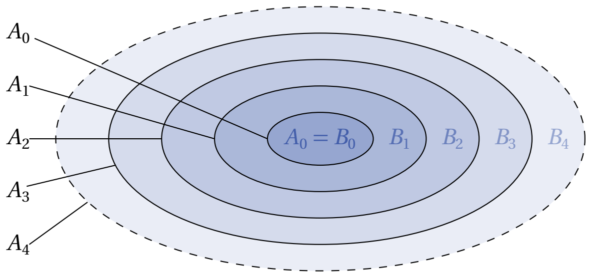
Démonstration

- Tout d'abord, par croissance de la probabilité,

$$\forall n \in \mathbb{N}, \quad A_n \subset A_{n+1} \quad \Longrightarrow \quad \forall n \in \mathbb{N}, \quad \mathbf{P}(A_n) \leqslant \mathbf{P}(A_{n+1})$$

La suite $\left(p_n\right)_{n\in\mathbb{N}}$ définie par $p_n = \mathbf{P}(A_n)$ étant croissante et majorée (par 1), elle converge. Ainsi, la limite figurant dans l'énoncé a bien un sens.

- La suite $(A_n)_{n\in\mathbb{N}}$ étant croissante, $\displaystyle\bigcup_{k=0}^{n} A_k = A_n$ donc $\mathbf{P}\left(\displaystyle\bigcup_{k=0}^{n} A_k\right) = \mathbf{P}(A_n)$. On ne peut cependant pas passer à la limite aussi facilement. On va devoir pour cela utiliser l'axiome de σ-additivité.



Les événements B_n sont deux à deux incompatibles

Les événements A_n n'étant pas deux à deux incompatibles, on utilise le subterfuge :

$$A_{n+1} = A_n \sqcup (A_{n+1} \setminus A_n)$$

qui nous amène à poser :

$$B_0 = A_0 \ \text{ et pour tout } n \geqslant 1,\ B_n = A_n \setminus A_{n-1}$$

- Ainsi construits, les événements B_n sont deux à deux incompatibles et $\displaystyle\bigcup_{n=0}^{+\infty} A_n = \bigsqcup_{n=0}^{+\infty} B_n$.

- Par σ-additivité, $\mathbf{P}\left(\displaystyle\bigcup_{n=0}^{+\infty} A_n\right) = \mathbf{P}\left(\displaystyle\bigsqcup_{n=0}^{+\infty} B_n\right) = \displaystyle\sum_{n=0}^{+\infty} \mathbf{P}(B_n)$. De plus,

$$\forall n \in \mathbb{N}, \quad \sum_{k=0}^{n} \mathbf{P}(B_k) = \left[\sum_{k=1}^{n} (\mathbf{P}(A_k) - \mathbf{P}(A_{k-1})) + \mathbf{P}(A_0)\right] = \mathbf{P}(A_n)$$

Ainsi, $\mathbf{P}\left(\displaystyle\bigcup_{n=0}^{+\infty} A_n\right) = \displaystyle\sum_{n=0}^{+\infty} \mathbf{P}(B_n) = \lim_{n\to+\infty} \mathbf{P}(A_n)$. ■

De même, si $(A_n)_{n\in\mathbb{N}}$ est une suite décroissante d'événements sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$, alors :

$$\mathbf{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} \mathbf{P}(A_n)$$

Il suffit d'appliquer la proposition précédente à la suite croissante d'événements $\left(\overline{A_n}\right)_{n\in\mathbb{N}}$ pour justifier ce résultat qualifié, lui, de *continuité décroissante*. En pratique, on aura plutôt recours au corollaire suivant.

Corollaire 10.28

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

$$\lim_{p\to+\infty} \mathbf{P}\left(\bigcup_{n=0}^{p} A_n\right) = \mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) \quad \text{et} \quad \lim_{p\to+\infty} \mathbf{P}\left(\bigcap_{n=0}^{p} A_n\right) = \mathbf{P}\left(\bigcap_{n=0}^{+\infty} A_n\right)$$

Démonstration

On applique la proposition précédente aux suites respectivement croissante et décroissante d'événements (B_p) et (C_p) définies par $B_p = \displaystyle\bigcup_{n=0}^{p} A_n$ et $C_p = \displaystyle\bigcap_{n=0}^{p} A_n$. ■

Exemple

On lance une pièce de monnaie équilibrée jusqu'à obtenir *pile*. Le jeu s'arrête presque sûrement dans le sens où nous allons prouver que la probabilité de n'obtenir que des *face* est nulle. Notons pour cela F_n l'événement « obtenir *face* au $n^{\text{ième}}$ lancer ». Par indépendance des tirages,

$$\mathbf{P}\left(\bigcap_{k=1}^{n} F_k\right) = \prod_{k=1}^{n} \mathbf{P}(F_k) = \frac{1}{2^n}$$

Il est tentant de passer à la limite ! Et c'est justement ce que nous permet de faire le résultat précédent :

$$\mathbf{P}\left(\bigcap_{n=1}^{+\infty} F_n\right) = \lim_{n\to+\infty} \mathbf{P}\left(\bigcap_{k=1}^{n} F_k\right) = \lim_{n\to+\infty} \frac{1}{2^n} = 0$$

Bien que sa réalisation soit possible, l'événement « obtenir seulement des *face* » est de probabilité nulle : il est négligeable ou quasi-impossible.

Proposition 10.29 : Sous-additivité finie

Soit $(A_1, \dots, A_n)$ une famille d'événements sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Alors,

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

Démonstration

Montrons par récurrence sur n que $\mathbf{P}\left(\bigcup_{k=0}^{n} A_k\right) \leqslant \sum_{k=0}^{n} \mathbf{P}(A_k)$

- La formule est évidente pour $n = 0$. Elle est également vraie pour $n = 1$:

$$\mathbf{P}(A_0 \cup A_1) = \mathbf{P}(A_0) + \mathbf{P}(A_1) - \mathbf{P}(A_0 \cap A_1) \leqslant \mathbf{P}(A_0) + \mathbf{P}(A_1)$$

- Supposons la propriété vraie à un certain rang n. Comme $\bigcup_{k=0}^{n+1} A_k = \left(\bigcup_{k=0}^{n} A_k\right) \cup A_{n+1}$,

$$\mathbf{P}\left(\bigcup_{k=0}^{n+1} A_k\right) \leqslant \mathbf{P}\left(\bigcup_{k=0}^{n} A_k\right) + \mathbf{P}(A_{n+1}) \leqslant \sum_{k=0}^{n} \mathbf{P}(A_k) + \mathbf{P}(A_{n+1}) = \sum_{k=0}^{n+1} \mathbf{P}(A_k)$$

La propriété est ainsi démontrée au rang $n+1$. ■

Un simple passage à la limite nous conduit directement à l'inégalité suivante.

Proposition 10.30 : Sous-additivité dénombrable

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

$$\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) \leqslant \sum_{n=0}^{+\infty} \mathbf{P}(A_n) \quad \text{(égalité dans } [0, +\infty])$$

La propriété de sous-additivité dénombrable s'applique facilement pour justifier que la réunion dénombrable d'événements négligeables reste négligeable.

Corollaire 10.31

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

Démonstration

Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements négligeables.

D'après l'inégalité précédente, $\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) \leqslant \sum_{n=0}^{+\infty} \mathbf{P}(A_n) = 0$. D'où $\mathbf{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = 0$. ■

De même, la réunion d'événements presque certains est encore un événement presque certain.

Définition 10.32 : Système quasi-complet d'événements

Soit $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé. On appelle système complet d'événements toute famille au plus dénombrable $(A_i)_{i\in I}$ d'événements vérifiant :

(i) pour tous $i, j \in I$ distincts, $A_i \cap A_j = \varnothing$ (ii) $\mathbf{P}\left(\bigsqcup_{i\in I} A_i\right) = 1$

La condition $\mathbf{P}\left(\bigsqcup_{i\in I} A_i\right) = 1$ est plus lâche que $\bigsqcup_{i\in I} A_i = \Omega$: la réunion n'est plus une partition de Ω si elle omet certains événements négligeables ; ce qui n'empêchera pas d'appliquer la formule des probabilités totales.

Exemple

Revenons à notre exemple de lancer de pièce de monnaie équilibrée jusqu'à obtenir *pile*. Notons A_n l'événement « obtenir le premier *pile* au $n^{\text{ième}}$ lancer ». En d'autres termes, $A_n = F_1 \cap \cdots \cap F_{n-1} \cap \overline{F_n}$.
$(A_n)_{n\geqslant 1}$ n'est pas un système complet d'événements à cause du tirage infini de *pile*.
C'est en revanche un système quasi-complet d'événements puisque $\mathbf{P}\left(\bigsqcup_{n=1}^{+\infty} A_n\right) = 1$.

III | Conditionnement et indépendance

A – Probabilité conditionnelle

Lorsque l'on dispose d'informations sur le résultat d'une expérience donnée, il est possible d'affiner nos prédictions. Considérons par exemple le lancer d'un dé non pipé. On associe à cette expérience l'univers $\Omega = [\![1, 6]\!]$ et on note A l'événement « le résultat est impair », B l'événement « le résultat est 3 ».

- Quelle est la probabilité d'obtenir 3 ? Bien évidemment, $\mathbf{P}(B) = 1/6$.
- Maintenant, sachant que le résultat est impair, quelle est la probabilité d'avoir obtenu 3 ? Tout se passe comme si l'on avait « déformé » notre univers des possibles en se restreignant à $\Omega' = A = \{1, 3, 5\}$.
  On trouve alors une probabilité de $1/3$.

$\mathbf{P}(A) = 1/2$ et $\mathbf{P}(B) = 1/6$. La probabilité que B soit réalisé sachant que A l'est vaut :

$$\frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} = \frac{\text{nombre de cas favorables pour } A \cap B}{\text{nombre de cas favorables pour } A} = \frac{1/6}{1/2} = \frac{1}{3}$$

On dit que l'événement A a conditionné l'ensemble des issues possibles, c'est-à-dire l'univers Ω.

Théorème / Définition 10.33 : Probabilité conditionnelle

Soient $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé et un A événement tel que $\mathbf{P}(A) \neq 0$. L'application

$$\mathbf{P}_A : \left| \begin{array}{ccc} \mathscr{A} & \longrightarrow & \mathbb{R} \\ B & \longmapsto & \dfrac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} \end{array} \right.$$

est une probabilité sur Ω. On l'appelle probabilité conditionnelle relative à A (ou sachant A).
Pour tout événement B, $\mathbf{P}_A(B)$ – notée encore $\mathbf{P}(B|A)$ – est appelée probabilité de B sachant A.

Démonstration

Montrons que $\mathbf{P}_A$ définit bien une probabilité sur Ω.

- Pour tout $B \in \mathscr{A}$, $A \cap B \subset A$ donc par croissance de $\mathbf{P}$, $\dfrac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)} \in [0,1]$.
- $\mathbf{P}_A(\Omega) = \dfrac{\mathbf{P}(A \cap \Omega)}{\mathbf{P}(A)} = 1$.
- Soit $(B_n)_{n \in \mathbb{N}}$ une suite d'événements deux à deux incompatibles.

$$\mathbf{P}_A\left(\bigsqcup_{n=0}^{+\infty} B_n\right) = \frac{\mathbf{P}\left(A \bigcap \left(\bigsqcup_{n=0}^{+\infty} B_n\right)\right)}{\mathbf{P}(A)} = \frac{\mathbf{P}\left(\bigsqcup_{n=0}^{+\infty} (A \cap B_n)\right)}{\mathbf{P}(A)} = \frac{\sum_{n=0}^{+\infty} \mathbf{P}(A \cap B_n)}{\mathbf{P}(A)} = \sum_{n=0}^{+\infty} \frac{\mathbf{P}(A \cap B_n)}{\mathbf{P}(A)} = \sum_{n=0}^{+\infty} \mathbf{P}_A(B_n)$$

car les événements $A \cap B_n$ sont deux à deux incompatibles. ■

Il importe de bien faire la distinction entre $\mathbf{P}_A(B)$, probabilité que l'événement B se réalise sachant que A est réalisé, et $\mathbf{P}(A \cap B)$, probabilité que les deux événements A et B se réalisent simultanément.

De plus, si l'utilisation de la notation $\mathbf{P}(B|A)$ au lieu de $\mathbf{P}_A(B)$ peut s'avérer agréable, il s'agit de comprendre que l'écriture $B|A$ prise isolément n'a aucun sens : on ne conditionne pas un événement par un autre : l'objet « $B|A$ » est indéfini, ce n'est pas un événement.

Revenons maintenant à quelques résultats déjà abordés en première année. Nous n'en donnerons une démonstration que lorsque celle-ci diffère du fait de l'éventuel caractère infini de l'univers considéré.

B – Les trois grandes formules de probabilités conditionnelles

1 – Formule des probabilités composées

Lemme 10.34

Soient A et B deux événements d'un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Alors, si $\mathbf{P}(A) \neq 0$,

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}_A(B) = \mathbf{P}(A) \cdot \mathbf{P}(B|A)$$

On notera qu'en général, $\mathbf{P}(A \cap B) \neq \mathbf{P}(A) \cdot \mathbf{P}(B)$.

Théorème 10.35 : Formule des probabilités composées

Soient $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé et $A_1, A_2, \ldots, A_n$ des événements tels que $\mathbf{P}(A_1 \cap \cdots \cap A_{n-1}) \neq 0$. Alors,

$$\mathbf{P}(A_1 \cap \cdots \cap A_n) = \mathbf{P}(A_1)\mathbf{P}_{A_1}(A_2)\mathbf{P}_{A_1 \cap A_2}(A_3) \cdots \mathbf{P}_{A_1 \cap \cdots \cap A_{n-1}}(A_n) = \mathbf{P}(A_1)\mathbf{P}(A_2|A_1) \cdots \mathbf{P}(A_n|A_1 \cap \cdots \cap A_{n-1})$$

Exemple

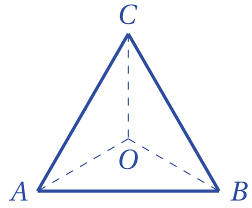
Une urne contient 4 boules blanches et 3 boules noires. On tire au hasard successivement et sans remise quatre boules dans cette urne. Quelle est la probabilité d'obtenir deux boules blanches puis deux noires ?
On munit pour cela $\Omega = \{b_1, b_2, b_3, b_4, n_1, n_2, n_3\}$ de la probabilité uniforme. On note B_i l'événement « obtenir une boule blanche au i^{e} tirage », N_i l'événement « obtenir une boule noire au i^{e} tirage ».
On calcule $\mathbf{P}(B_1 \cap B_2 \cap N_3 \cap N_4)$ à l'aide de la formule des probabilités composées.

$$\mathbf{P}(B_1) = \frac{4}{7}\,; \quad \mathbf{P}(B_2|B_1) = \frac{1}{2}\,; \quad \mathbf{P}(N_3|B_2 \cap B_1) = \frac{3}{5} \quad \text{et} \quad \mathbf{P}(N_4|N_3 \cap B_2 \cap B_1) = \frac{1}{2}$$

On trouve ainsi $\mathbf{P}(B_1 \cap B_2 \cap N_3 \cap N_4) = \dfrac{4}{7} \times \dfrac{1}{2} \times \dfrac{3}{5} \times \dfrac{1}{2} = \dfrac{3}{35}$.

Exercice 9

Une puce se déplace par sauts successifs sur les sommets et le centre de gravité d'un triangle équilatéral. Au temps $t = 0$, elle est en O. À chaque instant, elle saute du point où elle se trouve en l'un des autres points de façon équiprobable.



(i) Calculer la probabilité qu'elle revienne en O pour la première fois au temps $t = n$.

(ii) Calculer la probabilité de l'événement « la puce revient en O ». Commenter.

2 – Formule des probabilités totales

Théorème 10.36 : Formule des probabilités totales

Soient $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ un système complet d'événements.
Pour tout événement B, la série de terme général $\mathbf{P}(B\cap A_n)$ est convergente et :

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

Démonstration

Si $(A_n)_{n\in\mathbb{N}}$ est un système complet d'événement, quel que soit l'événement B,

$$B = B\cap\Omega = B\cap\left(\bigsqcup_{n=0}^{+\infty} A_n\right) = \bigsqcup_{n=0}^{+\infty}(B\cap A_n)$$

Par σ-additivité, $\mathbf{P}(B) = \mathbf{P}\left(\bigsqcup_{n=0}^{+\infty}(B\cap A_n)\right) = \sum_{n=0}^{+\infty}\mathbf{P}(B\cap A_n) = \sum_{n=0}^{+\infty}\mathbf{P}(B|A_n)\mathbf{P}(A_n)$. ■

(i) Pour que cette formule ait toujours un sens, on conviendra que $\mathbf{P}(B|A_n)\mathbf{P}(A_n) = 0$ lorsque $\mathbf{P}(A_n) = 0$.

(ii) La formule est encore valable lorsque la suite $(A_n)_{n\in\mathbb{N}}$ décrit un système quasi-complet d'événements.

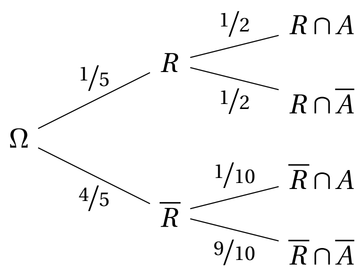
Cas particulier : si A est un événement, alors $(A, \overline{A})$ est un système complet d'événements.
Pour peu que $\mathbf{P}(A) \in]0;1[$, pour tout $B\in\mathscr{A}$, $\mathbf{P}(B) = \mathbf{P}(B|A)\mathbf{P}(A) + \mathbf{P}(B|\overline{A})\mathbf{P}(\overline{A})$.

Exemple

Une compagnie d'assurance distingue deux types d'assurés : les conducteurs à risque représentent 20% des assurés et les conducteurs prudents 80%. Les premiers ont une probabilité de 0.5 d'avoir un accident par an alors que les seconds seulement 0.1. Quelle est la probabilité qu'un nouvel assuré provoque un accident dans l'année qui suit la signature du contrat ? dans les deux ans suivant la signature du contrat ?

Notons R l'événement « le conducteur est à risque » et A l'événement « le conducteur provoque un accident dans l'année ». La première probabilité que l'on nous demande de calculer est $\mathbf{P}(A)$.
D'après la formule des probabilités totales,

$$\begin{aligned}\mathbf{P}(A) &= \mathbf{P}(A|R)\mathbf{P}(R) + \mathbf{P}(A|\overline{R})\mathbf{P}(\overline{R})\\ &= \frac{1}{2}\cdot\frac{1}{5} + \frac{1}{10}\cdot\frac{4}{5} = \frac{9}{50}\end{aligned}$$



Rappelons qu'il est important d'accompagner son raisonnement de tout schéma explicatif mais que celui-ci ne peut constituer une preuve.

Pour la deuxième question, mieux vaut calculer la probabilité pour un nouvel assurant de ne pas provoquer d'accident dans les deux premières années, à savoir $\left(\frac{9}{50}\right)^2$. On trouve ainsi $1-\left(\frac{9}{50}\right)^2 = \frac{2419}{2500} \approx 97\%$.

Exercice 10

Des boules en nombre infini sont placées successivement et indépendamment dans trois boîtes.

(i) Pour $k \geqslant 2$, on note A_k l'événement « deux des trois boîtes sont non vides pour la première fois lorsqu'on place la k–ième boule ». Calculer $\mathbf{P}(A_k)$ puis $\displaystyle\sum_{k=2}^{+\infty} \mathbf{P}(A_k)$. Interpréter.

(ii) Pour $i \geqslant 3$, on note B_i l'événement « les trois boîtes sont non vides pour la première fois lorsqu'on place la i–ème boule ». Calculer $\mathbf{P}(B_i|A_k)$ pour $k \geqslant 2$ et $i \geqslant 3$. En déduire $\mathbf{P}(B_i)$ puis $\displaystyle\sum_{i=3}^{+\infty} \mathbf{P}(B_i)$. Interpréter.

3 – Formule de Bayes

Théorème 10.37 : Formule de Bayes

Soient $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé.

- Si (A, B) est un couple d'événements vérifiant $\mathbf{P}(A) \cdot \mathbf{P}(B) \neq 0$, alors, $\mathbf{P}(A|B) = \dfrac{\mathbf{P}(A)}{\mathbf{P}(B)} \mathbf{P}(B|A)$.
- Si $(A_n)_{n\in\mathbb{N}}$ est un système complet d'événements et $B \in \mathscr{A}$, alors,

$$\forall k \in \mathbb{N}, \quad \mathbf{P}(A_k|B) = \frac{\mathbf{P}(B|A_k)\mathbf{P}(A_k)}{\displaystyle\sum_{n=0}^{+\infty} \mathbf{P}(B|A_n)\mathbf{P}(A_n)}$$

La formule de Bayes permet de « remonter aux causes » en inversant le conditionnement.

Exemple

Un magasin vend des téléviseurs provenant de deux entreprises (40% pour l'entreprise A et 60% pour l'entreprise B). 5% des appareils provenant de l'entreprise A présentent un défaut, 3% pour l'entreprise B. J'achète un appareil défectueux, quelle est la probabilité qu'il provienne de l'entreprise A?
On note F_A l'événement « l'appareil a été fabriqué par l'entreprise A » et D l'événement « il est défectueux ».

$$\mathbf{P}(F_A|D) = \frac{\mathbf{P}(F_A)\mathbf{P}(D|F_A)}{\mathbf{P}(D)} = \frac{\mathbf{P}(F_A)\mathbf{P}(D|F_A)}{\mathbf{P}(D|F_A)\mathbf{P}(F_A) + \mathbf{P}(D|F_B)\mathbf{P}(F_B)} = \frac{4}{10} \cdot \frac{5}{100} \cdot \frac{500}{19} = \frac{10}{19}$$

Exercice 11

Un magasin possède n caisses. Les clients se répartissent de façon indépendante et équiprobable entre les différentes caisses. On suppose que la probabilité qu'il y ait k clients dans le magasin est $p_k = \mathrm{e}^{-\lambda} \dfrac{\lambda^k}{k!}$. La caisse n°1 a reçu m clients un jour donné. Donner la probabilité qu'il y ait eu dans le magasin $n \cdot m$ clients.

C – Indépendance

Intuitivement, deux événements sont indépendants si la réalisation de l'un n'est pas conditionnée par la réalisation de l'autre. En clair, si $\mathbf{P}(B|A) = \mathbf{P}(B)$ lorsque $\mathbf{P}(A) > 0$. Cette idée se traduit par la définition suivante.

Définition 10.38 : Indépendance mutuelle d'événements

Soit $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisé.

- Deux événements A et B sont dits indépendants si $\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$.
- Soit $(A_i)_{i\in I}$ une famille d'événements. Ces événements sont dits (mutuellement) indépendants si pour toute partie finie $J \subset I$, $\mathbf{P}\left(\displaystyle\bigcap_{j\in J} A_j\right) = \displaystyle\prod_{j\in J} \mathbf{P}(A_j)$.

On retiendra que l'indépendance de deux événements est une propriété qui se vérifie par un calcul de probabilités… mais pas toujours ! L'indépendance d'une famille d'événements peut être supposée d'emblée par l'énoncé, mais c'est alors tout le modèle probabiliste retenu (univers, tribu, probabilité) qui s'adapte à cette hypothèse préalable. Cela ne va pas sans poser de problème de fond : rien ne nous garantit, par exemple, l'existence d'un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ adapté au jeu de *pile ou face*, modèle pour lequel l'indépendance de l'ensemble des lancers devrait être assurée. Cette question non triviale sera momentanément laissée de côté[2]. Par le calcul ou comme hypothèse préalable, dans tous les cas, on suivra le chemin tracé par l'énoncé !

Deux remarques supplémentaires :

- si les événements $A_1, \ldots, A_n$ sont mutuellement indépendants, $\mathbf{P}\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathbf{P}(A_i)$. En revanche, cette égalité ne garantit pas la mutuelle indépendance. Pour s'en convaincre, considérer le cas où $A_n = \varnothing$.
- La mutuelle indépendance assure l'indépendance deux à deux des événements : $\mathbf{P}(A_i \cap A_j) = \mathbf{P}(A_i)\mathbf{P}(A_j)$ si $i \neq j$. La réciproque est à nouveau fausse. Considérer pour cela les événements $A = \{1,2\}$, $B = \{1,3\}$ et $C = \{1,4\}$ et la probabilité uniforme $\mathbf{P}$ définie sur $\Omega = \{1,2,3,4\}$.

Proposition 10.39

Si les événements A et B sont indépendants, alors A et $\overline{B}$, $\overline{A}$ et $\overline{B}$, $\overline{A}$ et B le sont également.

Démonstration

Par exemple, $\mathbf{P}(A \cap \overline{B}) = \mathbf{P}(A \setminus B) = \mathbf{P}(A) - \mathbf{P}(A \cap B) = \mathbf{P}(A) - \mathbf{P}(A)\mathbf{P}(B) = \mathbf{P}(A)\mathbf{P}(\overline{B})$ par indép. de A et B. ■

Par extension, si $A_1, \ldots, A_n$ sont mutuellement indépendants, il en va de même pour $B_1, \ldots, B_n$ où $B_i \in \{A_i, \overline{A_i}\}$.

Exercice 12

Déterminer la probabilité d'une réunion finie d'événements mutuellement indépendants.

IV | Variables aléatoires discrètes

A – Généralités

Bien souvent, les événements dont on cherche à calculer la probabilité sont construits à partir de valeurs numériques prises par certaines grandeurs associées à une expérience aléatoire, comme « le nombre de *pile* obtenus au cours des tirages vaut 5 », ou « la somme des valeurs des deux dés vaut 7 ». L'introduction de variables aléatoires pour décrire les résultats d'une expérience aura comme double objectif d'aboutir à des rédactions épurées mais aussi, pour ne pas dire surtout, d'enrichir un point de vue jusqu'à présent purement ensembliste par l'introduction de fonctions.

Dans toute la suite, $(\Omega, \mathscr{A})$ désignera un espace probabilisable quelconque et E un ensemble non vide.

Définition 10.40 : Variable aléatoire discrète

Soient $(\Omega, \mathscr{A})$ un espace probabilisable et E un ensemble non vide. On appelle variable aléatoire discrète toute application $X : \Omega \to E$ telle que :

- $X(\Omega)$ est un ensemble fini ou dénombrable ;
- Pour tout $x \in X(\Omega)$, $X^{-1}(\{x\}) \in \mathscr{A}$.

L'événement $X^{-1}(\{x\})$ est traditionnellement noté $(X = x)$ ou $\{X = x\}$.

La définition nous assure de pouvoir définir la probabilité de l'événement $(X = x)$. Plusieurs remarques :

- Malgré son nom, une variable aléatoire n'est pas une *variable* (c'est une fonction) et elle n'est pas *aléatoire*.
- Lorsque $E = \mathbb{R}$ (la situation la plus courante), la variable aléatoire discrète est qualifiée de *réelle*.
- $X(\Omega)$ est l'image (directe) de Ω par X. Par définition d'une v.a.d., on écrira couramment $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$.
- Si Ω est au plus dénombrable, il en va de même pour $X(\Omega)$. Dans ce cas, on constate qu'en munissant Ω de sa tribu naturelle $\mathscr{P}(\Omega)$, toute fonction de Ω dans E est une variable aléatoire ; il n'y a alors rien à vérifier.

2. Nous y répondrons efficacement au moyen d'un théorème d'existence spécifique aux variables aléatoires.

- Lorsque Ω n'est pas dénombrable (exemple du *pile ou face*), prouver que X est une variable aléatoire n'est pas toujours chose aisée. À noter : les variables à densité ne figurent pas au programme de MP.

Si $A \subset E$, $X^{-1}(A)$ désigne l'image réciproque de A par X :

$$X^{-1}(A) \underset{\text{déf.}}{=} \{\omega \in \Omega \mid X(\omega) \in A\} \underset{\text{notation}}{=} (X \in A)$$

La notation $X^{-1}(A)$ ne sous-entend aucunement la bijectivité de X. $(X \in A)$ est clairement un événement :

Proposition 10.41

Soient $(\Omega, \mathscr{A})$ un espace probabilisable et X une variable aléatoire discrète sur cet espace.
Si $A \subset X(\Omega)$, alors $(X \in A)$ est un événement.

Démonstration

A est au plus dénombrable et $(X \in A) = \displaystyle\bigcup_{x \in A} (X = x) \in \mathscr{A}$ comme réunion au plus dénombrable d'événements.

Dans le cas d'une variable aléatoire réelle, on note également :

$$(X \leqslant x) = X^{-1}(]-\infty, x]) = \{\omega \in \Omega \mid X(\omega) \leqslant x\}\,; \quad (X < x) = X^{-1}(]-\infty, x[) = \{\omega \in \Omega \mid X(\omega) < x\}$$

$$(X \geqslant x) = X^{-1}([x, +\infty[) = \{\omega \in \Omega \mid X(\omega) \geqslant x\}\,; \quad (X > x) = X^{-1}(]x, +\infty[) = \{\omega \in \Omega \mid X(\omega) > x\}$$

Exercice 13

On lance simultanément deux dés discernables. On note X la somme des valeurs obtenues et Y le maximum des deux valeurs. Expliciter les événements suivants :

$$(X > 10)\,; \quad (X \geqslant 11)\,; \quad (X \leqslant 1)\,; \quad (Y = 3)\,; \quad (Y \in \{1, 2\})$$

Théorème 10.42

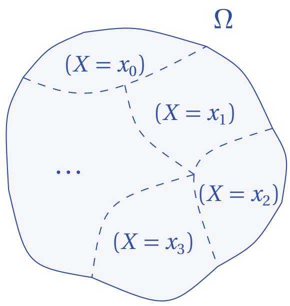
Soient $(\Omega, \mathscr{A})$ un espace probabilisable et X une variable aléatoire discrète sur cet espace.
Avec les notations précédentes, la famille $((X = x))_{x \in X(\Omega)}$ est un système complet d'événements.

Démonstration

Soient $\omega \in \Omega$ et $x \in X(\Omega)$. Alors, $\omega \in (X = x) \Longleftrightarrow X(\omega) = x$.
Ainsi, tout élément ω appartient à un, et un seul, événement $(X = x)$:

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

Lorsque Ω est au plus dénombrable, on pourra écrire $X(\Omega) = \{x_n \mid n \in \mathbb{N}\}$ et donc $\Omega = \bigsqcup_{n \in \mathbb{N}} (X = x_n)$.



$((X = x_n))_{n \in \mathbb{N}}$ *est un s.c.e.*

Exercice 14

Décrire le système complet d'événements associé à la variable Y introduite dans l'exercice précédent.

B – Loi d'une variable aléatoire

On munit désormais l'espace probabilisable $(\Omega, \mathscr{A})$ d'une probabilité $\mathbf{P}$.

Définition 10.43 : Loi d'une variable aléatoire

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. On appelle loi de probabilité de X (ou plus simplement loi de X) l'application $\mathbf{P}_X : A \in \mathscr{P}(X(\Omega)) \mapsto \mathbf{P}(X \in A)$.

La donnée d'une variable X à valeurs dans E confère à $X(\Omega)$ une structure d'espace probabilisé (et même par extension à l'ensemble E).

Théorème 10.44

Soit X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. $\mathbf{P}_X$ est une probabilité sur $X(\Omega)$.

Démonstration

- Pour tout $A \subset X(\Omega)$, $\mathbf{P}_X(A) = \mathbf{P}(X \in A) \in [0,1]$.
- $\mathbf{P}_X(X(\Omega)) = \mathbf{P}(X \in X(\Omega)) = \mathbf{P}(\Omega) = 1$.
- Pour une suite (A_n) d'événements incompatibles,

$$\mathbf{P}_X\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \mathbf{P}\left(\bigsqcup_{n\in\mathbb{N}} (X \in A_n)\right) = \sum_{n=0}^{+\infty} \mathbf{P}(X \in A_n) = \sum_{n=0}^{+\infty} \mathbf{P}_X(A_n)$$

La loi de X est entièrement déterminée par la distribution de probabilités $(\mathbf{P}(X = x))_{x\in X(\Omega)}$, appelée distribution de probabilités de X. On appelle support de la loi l'ensemble des éléments x de E tels que $\mathbf{P}(X = x) \neq 0$.

Dans le cas fini, la loi d'une variable peut être représentée par un tableau ou par un diagramme en bâtons.

Réciproquement, étant donné une certaine distribution de probabilités $(p_x)_{x\in E}$ définie sur un ensemble E au plus dénombrable, il est toujours possible de définir une variable aléatoire X de loi donnée par $(p_x)_{x\in E}$.

Théorème 10.45 : Existence d'une variable aléatoire de distribution de probabilités donnée

Soit $(p_x)_{x\in E}$ une distribution de probabilités définie sur un ensemble E au plus dénombrable.
Il existe un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et une variable aléatoire discrète X de loi donnée par $(p_x)_{x\in E}$.

Démonstration

Il suffit de considérer l'espace probabilisé $(E, \mathscr{P}(E), \mathbf{P})$ où $\mathbf{P}$ est définie par $\mathbf{P}(\{x\}) = p_x$.
En prenant $X = \mathrm{id}_E$, pour tout $x \in E$, $\mathbf{P}(X = x) = p_x$. ■

Rappelons quelques notations utiles :

- Si une variable X suit une certaine loi de probabilité $\mathscr{L}$, on notera $X \sim \mathscr{L}$.
- Si deux variables X et Y suivent une même loi de probabilité, on notera $X \sim Y$.

Deux variables aléatoires X et Y suivent la même loi si $X(\Omega) = Y(\Omega)$ et pour tout $x \in X(\Omega)$, $\mathbf{P}(X = x) = \mathbf{P}(Y = x)$. Attention, l'égalité en loi ne présuppose pas que les variables X et Y sont définies sur le même univers, ni donc *a fortiori* qu'elles sont égales.

Exemples

- Soit X une variable aléatoire suivant une *loi de Rademacher*, c'est-à-dire que $\mathbf{P}(X = 1) = \mathbf{P}(X = -1) = \dfrac{1}{2}$. Alors, $X \sim -X$.
- On considère une pièce équilibrée que l'on lance 10 fois. Le nombre X de piles obtenus et le nombre Y de faces obtenues suivent la même loi mais ne sont pas identiques.

Terminons ce premier tour d'horizon avec une dernière définition relative au conditionnement.

Définition 10.46 : Loi conditionnelle d'une variable aléatoire

Soient X une variable aléatoire discrète sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et un événement A de probabilité non nulle. On appelle loi conditionnelle de X sachant A la loi de X pour la probabilité $\mathbf{P}_A$, c'est-à-dire l'application $B \subset X(\Omega) \mapsto \mathbf{P}_A(X \in B)$.

Exercice 15

On choisit un entier X aléatoirement entre 1 et $2n$. Donner la loi de X sachant $(X \leqslant n)$.

C – Opérations sur les variables aléatoires

Il est courant de définir une variable aléatoire fonction d'autres variables aléatoires. En particulier, la donnée d'une application $f : E \to F$ entre deux ensembles E et F non vides et d'une variable aléatoire discrète X définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ à valeurs dans E permet de définir une nouvelle variable aléatoire $f \circ X$, traditionnellement notée $f(X)$. Il reste à vérifier que l'on a bien ainsi défini une variable aléatoire.

Proposition 10.47

Soient $(\Omega, \mathscr{A})$ un espace probabilisable, X une variable aléatoire discrète à valeurs dans E, ainsi qu'une application $f : E \to F$. Alors, $f(X)$ est une variable aléatoire discrète.

Démonstration

Posons $Y = f(X)$. $Y(\Omega) = f(X(\Omega))$ est au plus dénombrable et :

$$\forall y \in Y(\Omega), \quad (Y = y) = \bigcup_{\substack{x \in X(\Omega) \\ f(x) = y}} (X = x) \in \mathscr{A}$$

Proposition 10.48 : Loi de $f(X)$

Soient $(\Omega, \mathscr{A}, \mathbf{P})$ un espace probabilisable, X une variable aléatoire discrète à valeurs dans E, ainsi qu'une application $f : E \to F$. Alors, la loi de $f(X)$ est entièrement déterminée par f et la loi de X et :

$$\forall A \subset f(X(\Omega)), \quad \mathbf{P}_{f(X)}(A) = \mathbf{P}(f(X) \in A) = \sum_{\substack{x \in X(\Omega) \\ f(x) \in A}} \mathbf{P}(X = x)$$

Démonstration

Soit $A \subset f(X(\Omega))$. Alors,

$$(f(X) \in A) = (f(X) \in A) \cap \Omega = (f(X) \in A) \cap \bigsqcup_{x \in X(\Omega)} (X = x) = \bigsqcup_{x \in X(\Omega)} (f(x) \in A) \cap (X = x) = \bigsqcup_{\substack{x \in X(\Omega) \\ f(x) \in A)}} (X = x)$$

En particulier, si deux variables X et Y suivent la même loi, il en va de même pour $f(X)$ et $f(Y)$.

Notons enfin que les considérations précédentes s'étendent aux fonctions de plusieurs variables en considérant des n-uplets de variables aléatoires. En effet, si pour tout $i \in [\![1, n]\!]$, X_i est une variable discrète à valeurs dans un ensemble E_i, on vérifie aisément que $X = (X_1, \dots, X_n)$ est elle-même une variable discrète à valeurs dans $E_1 \times \cdots \times E_n$, appelée *vecteur aléatoire*. De ce fait, la somme et le produit de variables aléatoires discrètes réelles sont encore des variables aléatoires. Ainsi, X^2, $X + Y$, XY ou encore $\min(X, Y)$ sont des variables aléatoires. La loi d'une somme de n variables aléatoires s'exprime facilement :

$$\mathbf{P}(X_1 + \cdots + X_n = x) = \sum_{\substack{(x_1, \dots, x_n) \in \mathbb{R}^n \\ x_1 + \cdots + x_n = x}} \mathbf{P}(X_1 = x_1, \dots, X_n = x_n)$$

En revanche, le calcul peut s'avérer délicat sans autre hypothèse, comme l'indépendance des variables aléatoires qui fera l'objet de la section suivante.

Exercice 16

Déterminer la loi de X^2 où X est une variable aléatoire dont la table de probabilité est :

x_i	-2	-1	0	1	2
$\mathbf{P}(X = x_i)$	0.10	0.35	0.15	0.25	0.15

Exercice 17

Soit X une variable aléatoire discrète sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ telle que $X(\Omega) = \mathbb{N}^*$ et,

$$\forall k \in \mathbb{N}^*, \quad \mathbf{P}(X = k) = \frac{\alpha}{k(k+1)(k+2)}$$

1. Déterminer la valeur de α.
2. Montrer que $Y = X + 1$ est une variable aléatoire réelle discrète et déterminer sa loi.
3. Montrer que $Z = X^2 - 4X + 4$ est une variable aléatoire réelle discrète et déterminer sa loi.

D – Lois usuelles

Toutes les lois qui suivent sont directement définies à partir de la distribution de probabilités associée. X désignera une variable aléatoire discrète définie sur $(\Omega, \mathscr{A}, \mathbf{P})$ et à valeurs dans E.

1 – Loi certaine

Définition 10.49 : Loi certaine

On dit que la variable aléatoire X suit la loi certaine ou qu'elle est presque sûrement constante s'il existe $x \in E$ tel que $\mathbf{P}(X = x) = 1$.

2 – Loi uniforme

Définition 10.50 : Loi uniforme

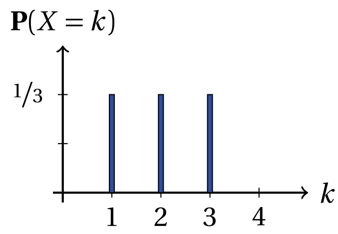
On dit que la variable aléatoire X suit la loi uniforme sur un ensemble fini E si :

$$X(\Omega) = E \quad \text{et} \quad \forall x \in E, \quad \mathbf{P}(X = x) = \frac{1}{|E|}$$

Notation : $X \sim \mathscr{U}(E)$.

Situations types :

- Lancer d'un dé équilibré ;
- Tirage aléatoire d'une boule dans une urne.



LOI DE PROBABILITÉ POUR $E = [\![1, 3]\!]$

3 – Loi de Bernoulli

Définition 10.51 : Loi de Bernoulli

On dit que la variable aléatoire X suit la loi de Bernoulli de paramètre $p \in [0, 1]$ si :

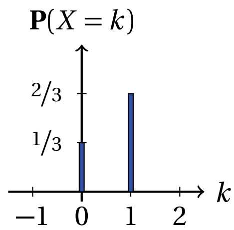
$$X(\Omega) = \{0, 1\} \quad \text{et} \quad \mathbf{P}(X = 1) = p$$

Notation : $X \sim \mathscr{B}(p)$ ou $X \sim \mathscr{B}(1, p)$.

On notera qu'alors $\mathbf{P}(X = 0) = q = 1 - p$.

Situations types :

- Lancer d'une pièce truquée (lancer unique) ;
- Tirage d'une boule dans une urne avec a boules noires et b boules blanches.
  Si on note X le nombre de boules noires tirées, $X \sim \mathscr{B}\left(\frac{a}{a+b}\right)$.



LOI DE PROBABILITÉ POUR $p = 2/3$

Exemple fondamental : si A est un événement, $\mathbb{1}_A \sim \mathscr{B}(\mathbf{P}(A))$.

4 – Loi binomiale

Définition 10.52 : Loi binomiale

On dit que la variable aléatoire X suit la loi binomiale de paramètre $(n,p)\in\mathbb{N}^*\times[0,1]$ si :

$$X(\Omega)=[\![0,n]\!] \quad \text{et} \quad \forall k\in[\![0,n]\!],\quad \mathbf{P}(X=k)=\binom{n}{k}p^k q^{n-k} \text{ où } q=1-p$$

Notation : $X\sim\mathscr{B}(n,p)$.

Vérifions que l'on a bien défini une loi de probabilité :

$$\forall k\in[\![0,n]\!],\quad \mathbf{P}(X=k)=\binom{n}{k}p^k q^{n-k}\geqslant 0 \quad \text{et} \quad \sum_{k=0}^{n}\mathbf{P}(X=k)=\sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}=(p+q)^n=1$$

Pour $n=1$, on retrouve le cas particulier d'une loi de Bernoulli.

Une loi binomiale de paramètres (n,p) apparaît lorsque l'on compte le nombre de succès lors de la répétition de n expériences aléatoires *indépendantes* conduisant à une situation de succès (de probabilité p) ou d'échec (de probabilité $q=1-p$). En notant X le nombre de succès, $X\sim\mathscr{B}(n,p)$.

Démonstration

Soient $n\in\mathbb{N}^*$ et $p\in[0,1]$. $X(\Omega)=[\![0,n]\!]$. Considérons une série de n lancers de pièce qui amène pile avec une probabilité $p\in]0,1]$ et face avec une probabilité $1-p$. Notons A_i l'événement « pile apparaît au i^{e} lancer ». Pour tout $k\in[\![1,n]\!]$, en notant $\mathscr{P}_k([\![1,n]\!])$ l'ensemble des parties à k éléments de $[\![1,n]\!]$.

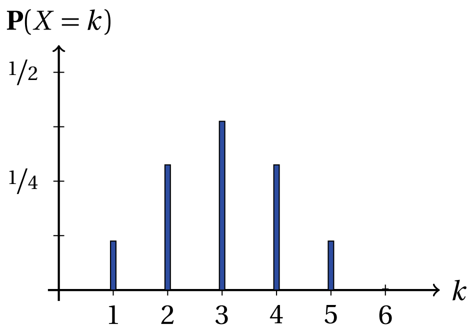
$$(X=k)=\bigsqcup_{I\in\mathscr{P}_k([\![1,n]\!])}\left(\bigcap_{i\in I}A_i\cap\bigcap_{i\notin I}\overline{A_i}\right)$$

Ainsi, par indépendance, $\mathbf{P}(X=k)=\displaystyle\sum_{I\in\mathscr{P}_k([\![1,n]\!])}\mathbf{P}\left(\bigcap_{i\in I}A_i\cap\bigcap_{i\notin I}\overline{A_i}\right)=\sum_{I\in\mathscr{P}_k([\![1,n]\!])}p^k q^{n-k}=\binom{n}{k}p^k q^{n-k}$. ■

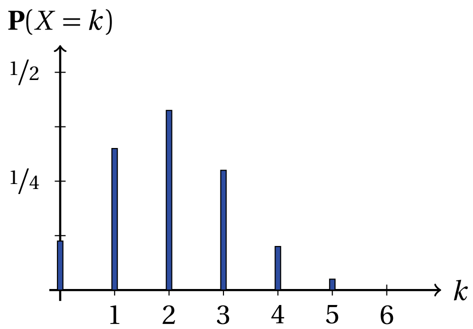
On se débarrassera de ce formalisme inutile en constatant dans la section suivante que la somme de n variables indépendantes suivant toutes des lois de Bernoulli de paramètre p suit une loi binomiale de paramètre (n,p).

Situations types :

- Lancers successifs d'une pièce éventuellement truquée ;
- Tirages successifs *avec remise* de n boules dans une urne avec a boules noires et b boules blanches. Si on note X le nombre de boules noires tirées, $X\sim\mathscr{B}\left(n,\frac{a}{a+b}\right)$.



REPRÉSENTATION D'UNE LOI BINOMIALE DE PARAMÈTRES $\left(6,\frac{1}{2}\right)$



REPRÉSENTATION D'UNE LOI BINOMIALE DE PARAMÈTRES $\left(6,\frac{1}{3}\right)$

5 – Loi géométrique

Définition 10.53 : Loi géométrique

On dit que la variable aléatoire X suit une loi géométrique de paramètre $p \in]0,1]$ sur $\mathbb{N}^*$ si :

$$X(\Omega) = \mathbb{N}^* \quad \text{et} \quad \forall n \in \mathbb{N}^*, \quad \mathbf{P}(X = n) = q^{n-1}p \text{ où } q = 1 - p$$

Notation : $X \sim \mathscr{G}(p)$.

Vérifions que l'on a bien défini une loi de probabilité :

$$\forall n \in \mathbb{N}^*, \quad \mathbf{P}(X = n) = q^{n-1}p \geqslant 0 \quad \text{et} \quad \sum_{n=1}^{+\infty} \mathbf{P}(X = n) = p \sum_{n=1}^{+\infty} q^{n-1} = p \sum_{n=0}^{+\infty} q^n = \frac{p}{1-q} = 1$$

Une loi géométrique apparaît lorsque l'on s'intéresse à une succession d'épreuves de Bernoulli indépendantes jusqu'à obtenir un premier succès. On parle alors de temps d'attente du premier succès.

Démonstration

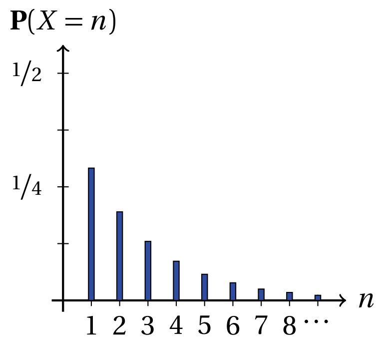
Considérons une série infinie de lancers de pièce qui amène pile avec une probabilité $p \in]0,1]$ et face avec une probabilité $1-p$. Notons A_k l'événement « pile apparaît au k^{e} lancer ». Par indépendance des événements,

$$\forall k \geqslant 2 \quad \mathbf{P}(X = k) = \mathbf{P}\left(\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_{k-1}} \cap A_k\right) = \mathbf{P}(\overline{A_1}) \times \cdots \times \mathbf{P}(\overline{A_{k-1}}) \times \mathbf{P}(A_k) = q^{k-1}p$$

De plus, $\mathbf{P}(X = 1) = p$. ■

Situations types :

- Premier pile obtenu au cours de lancers successifs d'une pièce truquée ;
- Tirages successifs *avec remise* de boules dans une urne avec des boules noires et blanches jusqu'à obtenir une boule noire.



Représentation d'une loi géométrique de paramètre $1/3$

Nous avons jusqu'à présent interprété une loi géométrique comme le temps d'attente du 1er succès dans une suite d'épreuves de Bernoulli indépendantes. Si un observateur commence son observation au temps n et qu'il constate que l'événement ne s'est pas encore produit, quelle est la probabilité que l'événement se produise au temps $n+k$? Elle est en fait égale à la probabilité que l'événement se produise au temps k, comme si rien ne s'était produit jusque là. La loi géométrique est qualifiée de « loi sans mémoire » ; c'est même la seule loi discrète vérifiant cette propriété.

Proposition 10.54 : Caractérisation comme loi sans mémoire

Soit X une variable aléatoire sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ telle que $X(\Omega) = \mathbb{N}^*$.
X suit une loi géométrique sur $\mathbb{N}^*$ si et seulement si :

$$\forall (k,n) \in \mathbb{N}^2, \quad \mathbf{P}(X > n + k | X > n) = \mathbf{P}(X > k)$$

Démonstration

Remarquons tout d'abord que si $X \sim \mathscr{G}(p)$, $\mathbf{P}(X > n) = 1 - \sum_{k=1}^{n} \mathbf{P}(X = k) = 1 - \sum_{k=1}^{n} pq^{k-1} = q^n$. L'interprétation est simple : la réalisation de l'événement à un temps supérieur à n nécessite n premiers échecs consécutifs.

$\boxed{\Longrightarrow}$ Si $X \sim \mathscr{G}(p)$ avec $p \in]0,1[$, pour tout $(k,n) \in \mathbb{N}^2$, $\mathbf{P}(X > n+k | X > n) = \dfrac{\mathbf{P}(X > n+k)}{\mathbf{P}(X > n)} = \dfrac{q^{n+k}}{q^n} = \mathbf{P}(X > k)$.

$\boxed{\Longleftarrow}$ Supposons maintenant que pour tout $(k,n)\in\mathbb{N}$, $\mathbf{P}(X>n+k|X>n)=\mathbf{P}(X>k)$.
Ainsi, $\mathbf{P}(X>n+1)=\mathbf{P}(X>n+1|X>n)\mathbf{P}(X>n)=\mathbf{P}(X>1)\mathbf{P}(X>n)$.
En posant $q=\mathbf{P}(X>1)$, on voit que $(\mathbf{P}(X>n))_{n\in\mathbb{N}^*}$ est une suite géométrique de raison q.

$$\forall n\in\mathbb{N}^*,\quad \mathbf{P}(X>n)=q\mathbf{P}(X>n-1)=\cdots=q^{n-1}\mathbf{P}(X>1)=q^n$$

Ainsi, $\mathbf{P}(X=n)=\mathbf{P}(X\leqslant n)-\mathbf{P}(X\leqslant n-1)=\mathbf{P}(X>n-1)-\mathbf{P}(X>n)=q^{n-1}-q^n=pq^{n-1}$; $X\sim\mathscr{G}(p)$. ■

6 – Loi de Poisson

Contrairement aux lois précédentes, il s'avère difficile de donner un modèle simple qui conduise à une loi de Poisson. Celle-ci apparaît plus naturellement comme une loi limite. Cette loi a été introduite en 1838 par Poisson et sert souvent à modéliser l'apparition d'événements rares.

Définition 10.55 : Loi de Poisson

On dit que la variable aléatoire X suit une loi de Poisson de paramètre $\lambda\in\mathbb{R}_+^*$ si :

$$X(\Omega)=\mathbb{N}\quad \text{et}\quad \forall n\in\mathbb{N},\quad \mathbf{P}(X=n)=\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}$$

Notation : $X\sim\mathscr{P}(\lambda)$.

Vérifions que l'on a bien défini une loi de probabilité :

$$\forall n\in\mathbb{N},\quad \mathbf{P}(X=n)=\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}\geqslant 0\quad\text{et}\quad \sum_{n=0}^{+\infty}\mathbf{P}(X=n)=\sum_{n=0}^{+\infty}\frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}=\mathrm{e}^{-\lambda}\sum_{n=0}^{+\infty}\frac{\lambda^n}{n!}=\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda}=1$$

Théorème 10.56 : Approximation d'une loi binomiale par une loi de Poisson

Si pour tout $n\in\mathbb{N}^*$, $X_n\sim\mathscr{B}(n,p_n)$ et si $\lim\limits_{n\to+\infty} np_n=\lambda$, alors,

$$\forall k\in\mathbb{N},\quad \mathbf{P}(X_n=k)\xrightarrow[n\to+\infty]{}\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$$

Démonstration

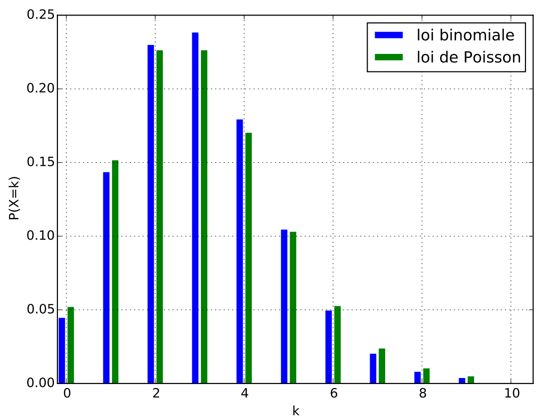
Pour tout $k\in\mathbb{N}$, $\mathbf{P}(X_n=k)=\binom{n}{k}p_n^k(1-p_n)^{n-k}=\dfrac{n(n-1)\times\cdots\times(n-k+1)}{k!}p_n^k(1-p_n)^{n-k}$.

Donc $\mathbf{P}(X_n=k)\underset{n\to+\infty}{\sim}\dfrac{(np_n)^k}{k!}\mathrm{e}^{(n-k)\ln(1-p_n)}$. Or $(n-k)\ln(1-p_n)\underset{n\to+\infty}{\sim}-np_n$ puisque $p_n\xrightarrow[n\to+\infty]{}0$.

Ainsi, $\mathbf{P}(X_n=k)\xrightarrow[n\to+\infty]{}\mathrm{e}^{-\lambda}\dfrac{\lambda^k}{k!}$. ■

Ce résultat affirme que si l'on réalise un grand nombre d'épreuves de Bernoulli indépendantes de paramètre vérifiant $p_n\underset{n\to+\infty}{\sim}\lambda/n$, le nombre de succès suit approximativement une loi de Poisson de paramètre λ. La qualité de l'approximation reste en revanche à quantifier.

```python
import numpy as np
import matplotlib.pyplot as plt

p = 0.1
n = 30

def binom(n, k):
    return np.math.factorial(n)/(np.math.factorial(k)*np.math.factorial(n-k))

def B(k):
    return binom(n, k)*p**k*(1-p)**(n-k)

def P(k):
    return np.exp(-n*p)*(n*p)**k/np.math.factorial(k)

for k in range(10):
    plt.plot([k-0.1, k-0.1], [0, B(k)], 'b', linewidth = 6)
    plt.plot([k+0.1, k+0.1], [0, \P(k)], 'g', linewidth = 6)
plt.axis([-0.2,10.5,0,0.25])
plt.grid(True)
plt.xlabel('k')
plt.ylabel('\P(X=k)')
plt.legend(['loi binomiale', 'loi de Poisson'])
plt.show()
```



Approximation d'une loi binomiale par une loi de Poisson

V | Familles de variables aléatoires

A – Couples de variables aléatoires

Théorème / Définition 10.57 : Couple de variables, loi conjointe, lois marginales

Soient X, Y deux variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et à valeurs respectivement dans E et F. Alors,

- $(X, Y) : \Omega \to E \times F$ est une variable aléatoire discrète.
- La loi de (X, Y) est appelée loi conjointe de X et Y.
- Les lois de X et Y sont appelées lois marginales de (X, Y).

Démonstration

- $X(\Omega)$ et $Y(\Omega)$ sont au plus dénombrables. Il en va de même pour $(X, Y)(\Omega) \subset X(\Omega) \times Y(\Omega)$.
- Pour tout $(x, y) \in (X, Y)(\Omega)$, $\big((X, Y) = (x, y)\big) = (X = x) \cap (Y = y) \in \mathscr{A}$. ■

- On notera qu'en général, $(X, Y)(\Omega) \neq X(\Omega) \times Y(\Omega)$ malgré l'inclusion évidente.
- On pourra utiliser les notations suivantes : $(X = x) \cap (Y = y) = \big((X, Y) = (x, y)\big) = \big(X = x, Y = y\big)$.
- Donner la loi conjointe de (X, Y) revient à calculer $\mathbf{P}(X = x, Y = y)$ pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$.
- $\big((X = x, Y = y)\big)_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est un système complet d'événements.

Exercice 18

On lance deux dés. On note X le maximum des valeurs obtenues, Y la valeur du deuxième dé.
Que vaut $X(\Omega)$? $Y(\Omega)$? $(X, Y)(\Omega)$?

Les lois marginales se déduisent de la loi conjointe (la réciproque est fausse !). En effet, d'après la formule des probabilités totales appliquée aux systèmes complets d'événements $((X = x))_{x \in X(\Omega)}$ et $\big((Y = y)\big)_{y \in Y(\Omega)}$,

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

Dans le cas fini, la loi conjointe peut être représentée par une table de probabilités. Le résultat précédent exprime l'idée qu'il suffit de sommer les éléments le long des lignes et des colonnes pour retrouver les lois marginales.

Exemple

On considère une urne avec 4 boules B_1, B_2, B_3, B_4 et les tirages de 2 boules successifs avec remise. On note X_1 le numéro de la boule au 1^er^ tirage et X_2 le numéro de la boule au 2^ème^ tirage. On pose enfin $Y = \max(X_1, X_2)$ et $Z = (X_1, Y)$. Quelle est la loi conjointe de X_1 et Y ? $X_1(\Omega) = Y(\Omega) = [\![1, 4]\!]$ et de plus,

X_1 / Y	1	2	3	4
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
2	0	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$
3	0	0	$\frac{3}{16}$	$\frac{1}{16}$
4	0	0	0	$\frac{1}{4}$

$$\forall k \in \{1, \dots, 4\}, \quad \mathbf{P}(X_1 = k) = \frac{1}{4}$$

$$\mathbf{P}(Y = 1) = \frac{1}{16}\ ; \quad \mathbf{P}(Y = 2) = \frac{3}{16}\ ;$$
$$\mathbf{P}(Y = 3) = \frac{5}{16}\ ; \quad \mathbf{P}(Y = 4) = \frac{7}{16}.$$

Exemple

Une urne contient deux boules B et $n \in \mathbb{N}^*$ boules N. On tire toutes les boules une par une.
On note X le rang du tirage de la 1e boule blanche tirée, Y celui de la 2ème.
Quelle est la loi du couple (X, Y) ? la loi de X ? la loi de Y ? la loi de $Y - X$?

- Ω est l'ensemble des arrangements des $n+2$ boules parmi $n+2$ et $X(\Omega) = [\![1, n+1]\!]$, $Y(\Omega) = [\![2, n+2]\!]$.
- $(X = i) \cap (Y = j) = \varnothing$ pour $i \geqslant j$, et pour $i < j$, $\mathbf{P}((X = i) \cap (Y = j)) = \dfrac{2}{(n+1)(n+2)}$.
- On en tire facilement les lois de X et Y :

$$\mathbf{P}(X = i) = \sum_{j=2}^{n+2} \mathbf{P}((X = i) \cap (Y = j)) = \sum_{j=i+1}^{n+2} \mathbf{P}((X = i) \cap (Y = j)) = \frac{2(n+2-i)}{(n+1)(n+2)}$$
$$\mathbf{P}\big(Y = j\big) = \sum_{i=1}^{n+1} \mathbf{P}((X = i) \cap (Y = j)) = \sum_{i=1}^{j-1} \mathbf{P}((X = i) \cap (Y = j)) = \frac{2(j-1)}{(n+1)(n+2)}$$

- $(Y - X)(\Omega) = [\![1, n+1]\!]$; $((X = i))$ étant un système complet d'événements,

$$\forall k \in [\![1, n+1]\!], \quad \mathbf{P}(Y - X = k) = \mathbf{P}(Y = X + k) = \sum_{i=1}^{n+1} \mathbf{P}(Y = X + k | X = i)\mathbf{P}(X = i) = \frac{2(n+2-k)}{(n+1)(n+2)}$$

Exercice 19

Dans une succession de *pile ou face* pour laquelle la probabilité d'obtenir pile est $p \in]0, 1[$ et d'obtenir face est $q = 1 - p$, on note X le rang d'apparition du premier pile et Y celui du second.
Déterminer la loi conjointe de (X, Y) puis les lois marginales.

Définition 10.58 : Loi conditionnelle

Soient X, Y deux variables aléatoires discrètes définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.
On appelle loi conditionnelle de X sachant $(Y = y)$ pour $\mathbf{P}(Y = y) \neq 0$, la loi définie par :

$$\forall x \in X(\Omega), \quad \mathbf{P}_{(Y=y)}((X = x)) = \frac{\mathbf{P}\big((X = x) \cap (Y = y)\big)}{\mathbf{P}\big(Y = y\big)}$$

On définit de même la loi conditionnelle de Y sachant $(X = x)$ pour $\mathbf{P}(X = x) \neq 0$.

Théorème / Définition 10.59 : Indépendance de deux variables aléatoires

Deux variables aléatoires discrètes X et Y définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ sont dites indépendantes si l'une des deux assertions équivalentes suivantes est vérifiée :

(i) Pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, $\mathbf{P}\big((X = x) \cap (Y = y)\big) = \mathbf{P}(X = x) \times \mathbf{P}\big(Y = y\big)$.

(ii) Pour tout $(A, B) \subset X(\Omega) \times Y(\Omega)$, $\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A) \times \mathbf{P}(Y \in B)$.

Notation (inutile) : $X \perp\!\!\!\perp Y$.

Démonstration

Supposons que pour tous $x \in X(\Omega)$ et $y \in Y(\Omega)$, $\mathbf{P}\big(X = x, Y = y\big) = \mathbf{P}(X = x) \times \mathbf{P}\big(Y = y\big)$.
Soit $(A, B) \subset X(\Omega) \times Y(\Omega)$. Alors, par sommabilité,

$$\begin{aligned} \mathbf{P}(X \in A) \cdot \mathbf{P}(Y \in B) &= \sum_{x \in A} \mathbf{P}(X = x) \cdot \sum_{y \in B} \mathbf{P}(Y = y) = \sum_{(x,y) \in A \times B} \mathbf{P}(X = x) \cdot \mathbf{P}(Y = y) \\ &\underset{\text{ind.}}{=} \sum_{(x,y) \in A \times B} \mathbf{P}(X = x, Y = y) = \mathbf{P}\left(\bigsqcup_{(x,y) \in A \times B} (X = x, Y = y) \right) = \mathbf{P}(X \in A, Y \in B) \end{aligned}$$

La réciproque est immédiate. ■

En cas d'indépendance, on retrouve la loi conjointe à partir des lois marginales. Plus précisément, le résultat précédent montre que la distribution de probabilités associée à (X,Y) est le produit des distributions de probabilités des variables X et Y.

Exemple – somme de deux variables indépendantes (important !)

On exprime facilement la loi de la somme de deux variables aléatoires indépendantes X et Y à partir des lois marginales. Dans le cas de variables prenant des valeurs entières, on trouve :

$$\forall k \in (X+Y)(\Omega), \quad \mathbf{P}(X+Y=k) = \sum_{\substack{(i,j)\in\mathbb{N}^2 \\ i+j=k}} \mathbf{P}(X=i)\mathbf{P}(Y=j) = \sum_{i=0}^{k} \mathbf{P}(X=i)\mathbf{P}(Y=k-i)$$

Corollaire 10.60

Soient X et Y deux variables aléatoires sur $(\Omega,\mathscr{A},\mathbf{P})$ et deux applications $f : X(\Omega) \to E$ et $g : Y(\Omega) \to F$.
Si X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ sont indépendantes.

Démonstration

Soient $e \in E$ et $f \in F$. Alors, $\mathbf{P}\big(f(X)=e, g(Y)=f\big) = \mathbf{P}\big(X \in f^{-1}(\{e\}), Y \in g^{-1}(\{f\}\big)$ et donc, par indépendance, $\mathbf{P}\big(f(X)=e, g(Y)=f\big) = \mathbf{P}\big(X \in f^{-1}(\{e\})\big)\cdot\mathbf{P}\big(Y \in g^{-1}(\{f\}\big) = \mathbf{P}(X=e)\cdot\mathbf{P}(Y=f)$. ■

Exemple

Si X et Y sont indépendantes, alors X^2 et Y^2 le sont également.

B – Familles finies de variables aléatoires

On étend sans difficulté les définitions et résultats précédents aux n-uplets de variables aléatoires :

- si pour tout $i \in [\![1,n]\!]$, X_i est une variable discrète sur $(\Omega,\mathscr{A},\mathbf{P})$ à valeurs dans un ensemble E_i, alors $X=(X_1,\dots,X_n)$ est une variable discrète à valeurs dans $E_1\times\cdots\times E_n$;
- la loi de X est appelée loi conjointe quand les lois des variables X_i sont qualifiées de lois marginales ;
- chaque loi marginale se retrouve à partir de la loi conjointe. En posant $F_i = E_1\times\cdots\times E_{i-1}\times E_{i+1}\times\cdots\times E_n$,

$$\forall i \in [\![1,n]\!], \quad \forall y \in X_i(\Omega), \quad \mathbf{P}(X_i=y) = \sum_{(x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)\in F_i} \mathbf{P}(X_1=x_1,\dots,X_i=y,\dots,X_n=x_n)$$

Théorème / Définition 10.61 : Variables aléatoires mutuellement indépendantes

Soit $(X_i)_{1\leqslant i\leqslant n}$ une famille finie de variables aléatoires définies sur $(\Omega,\mathscr{A},\mathbf{P})$. Les variables sont dites (mutuellement) indépendantes si l'une des deux assertions équivalentes suivantes est vérifiée :

(i) Pour tout $(x_1,\dots,x_n)\in X_1(\Omega)\times\cdots\times X_n(\Omega)$, $\mathbf{P}(X_1=x_1,\dots,X_n=x_n) = \displaystyle\prod_{i=1}^{n}\mathbf{P}(X_i=x_i)$.

(ii) Pour tout $(A_1,\dots,A_n)\subset X_1(\Omega)\times\cdots\times X_n(\Omega)$, les événements $(X_1\in A_1),\dots,(X_n\in A_n)$ sont mutuellement indépendants.

Démonstration

Soit $(A_1,\dots,A_n)\subset X_1(\Omega)\times\cdots\times X_n(\Omega)$. À l'image de l'indépendance de deux variables, par sommabilité :

$$\begin{aligned}\prod_{i=1}^{n}\mathbf{P}(X_i\in A_i) = \prod_{i=1}^{n}\sum_{x_i\in A_i}\mathbf{P}(X_i=x_i) &= \sum_{(x_1,\dots,x_n)\in A_1\times\cdots\times A_n}\prod_{i=1}^{n}\mathbf{P}(X_i=x_i) \\ &\underset{\text{ind.}}{=} \sum_{(x_1,\dots,x_n)\in A_1\times\cdots\times A_n}\mathbf{P}(X_1=x_1,\dots,X_n=x_n) = \mathbf{P}(X_1\in A_1,\dots,X_n\in A_n)\end{aligned}$$

Prouvons maintenant que toute partie I de $[\![1,n]\!]$, $\mathbf{P}\left(\displaystyle\bigcap_{i\in I}(X_i\in A_i)\right) = \displaystyle\prod_{i\in I}\mathbf{P}(X_i\in A_i)$.

On pose pour cela $B_i = A_i$ si $i \in I$ et $B_i = X_i(\Omega)$ sinon. En adaptant l'égalité précédemment démontrée,

$$\mathbf{P}\left(\bigcap_{i\in I}(X_i \in A_i)\right) = \mathbf{P}\left(\bigcap_{i=1}^{n}(X_i \in B_i)\right) = \prod_{i=1}^{n}\mathbf{P}(X_i \in B_i) = \prod_{i\in I}\mathbf{P}(X_i \in A_i)$$

On prendra garde au fait que si la mutuelle indépendance implique l'indépendance deux à deux, la réciproque est, comme pour les événements, fausse !

Exemple

Les événements $F_1, \dots, F_n$ sont indépendants ssi les indicatrices $\mathbb{1}_{F_1}, \dots, \mathbb{1}_{F_n}$ sont des variables indépendantes.

Énonçons un dernier résultat qui rendra de précieux services et qui se fonde sur une intuition peu originale : si X, Y et Z sont mutuellement indépendants, il en va de même pour $X + Y$ et Z.

Théorème 10.62 : Lemme des coalitions

Soient $X_1, \dots, X_n$ des variables indépendantes sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$, E et F deux ensembles, $m \in [\![1, n]\!]$ et $f : X_1(\Omega) \times \dots \times X_m(\Omega) \to E$, $g : X_{m+1}(\Omega) \times \dots \times X_n(\Omega) \to F$.
Alors, $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont indépendantes.

Démonstration

On montre que les variables aléatoires $X = (X_1, \dots, X_m)$ et $Y = (X_{m+1}, \dots, X_n)$ sont indépendantes. ■

Nous serons couramment amenés à déterminer la loi d'une somme de n variables aléatoires réelles :

$$\mathbf{P}(X_1 + \dots + X_n = x) = \sum_{\substack{(x_1,\dots,x_n)\in\mathbb{R}^n \\ x_1+\dots+x_n=x}} \mathbf{P}(X_1 = x_1, \dots, X_n = x_n)$$

Lorsque les variables aléatoires sont indépendantes,

$$\mathbf{P}(X_1 + \dots + X_n = x) = \sum_{\substack{(x_1,\dots,x_n)\in\mathbb{R}^n \\ x_1+\dots+x_n=x}} \mathbf{P}(X_1 = x_1, \dots, X_n = x_n) = \sum_{\substack{(x_1,\dots,x_n)\in\mathbb{R}^n \\ x_1+\dots+x_n=x}} \mathbf{P}(X_1 = x_1) \times \dots \times \mathbf{P}(X_n = x_n)$$

Donnons maintenant deux exemples fondamentaux de sommes de variables aléatoires indépendantes.

Lemme 10.63 : Formule de Vandermonde (HP)

On considère trois entiers naturels m, n, p tels que $p \leqslant n + m$. Alors, $\displaystyle\sum_{k=0}^{p}\binom{n}{k}\binom{m}{p-k} = \binom{n+m}{p}$.

Démonstration

Donnons une preuve non combinatoire de ce résultat. Posons pour cela $P = (X+1)^n$ et $Q = (X+1)^m$.

- Le terme de degré p du polynôme $PQ = (X+1)^{n+m}$ a pour expression $\displaystyle\binom{n+m}{p}X^p$.
- En utilisant la formule du produit de deux polynômes, $\displaystyle\sum_{i+j=p}\binom{n}{i}\binom{m}{j}X^{i+j} = \sum_{k=0}^{p}\binom{n}{k}\binom{m}{p-k}X^p$.

D'où l'égalité par identification. ■

Théorème 10.64 : Somme de variables indépendantes suivant une loi binomiale

On suppose toutes les variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

(i) Si $X \sim \mathscr{B}(n_1, p)$ et $Y \sim \mathscr{B}(n_2, p)$ sont indépendantes, alors $X + Y \sim \mathscr{B}(n_1 + n_2, p)$.

(ii) Si $X_1 \sim \mathscr{B}(n_1, p), \dots, X_m \sim \mathscr{B}(n_m, p)$ sont indépendantes, alors $X_1 + \dots + X_m \sim \mathscr{B}(n_1 + \dots + n_m, p)$.

Démonstration

(i) Soient $X \sim \mathscr{B}(n_1, p)$ et $Y \sim \mathscr{B}(n_2, p)$ indépendantes. $(X+Y)(\Omega) = [\![0, n_1+n_2]\!]$ et pour $k \in [\![0, n_1+n_2]\!]$,

$$\mathbf{P}(X+Y=k) = \sum_{i=0}^{k} \mathbf{P}(X=i)\mathbf{P}(Y=k-i) = \sum_{i=0}^{k} \binom{n_1}{i}\binom{n_2}{k-i} p^k q^{n_1+n_2+k} = \binom{n_1+n_2}{k} p^k q^{n_1+n_2-k}$$

(ii) On raisonne par récurrence en utilisant l'indépendance des variables $X_1+\cdots+X_n$ et de X_{n+1}. ■

Ainsi, la somme de n variables de Bernoulli indépendantes de même paramètre suit une loi binomiale.

Théorème 10.65 : Somme de variables indépendantes suivant une loi de Poisson (HP)

On suppose toutes les variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$.

(i) Si $X \sim \mathscr{P}(\lambda)$ et $Y \sim \mathscr{P}(\mu)$ sont indépendantes, alors $X+Y \sim \mathscr{P}(\lambda+\mu)$.

(ii) Si $X_1 \sim \mathscr{P}(\lambda_1), \ldots, X_n \sim \mathscr{P}(\lambda_n)$ sont indépendantes, alors $X_1+\cdots+X_n \sim \mathscr{P}(\lambda_1+\cdots+\lambda_n)$.

Démonstration

Supposons $X \sim \mathscr{P}(\lambda)$ et $Y \sim \mathscr{P}(\mu)$ indépendantes. Alors, $(X+Y)(\Omega) = \mathbb{N}$ et pour tout $n \in \mathbb{N}$,

$$\mathbf{P}(X+Y=n) = \sum_{i=0}^{n} \mathbf{P}(X=i)\mathbf{P}(Y=n-i) = \sum_{i=0}^{n} \mathrm{e}^{-\lambda}\frac{\lambda^i}{i!}\mathrm{e}^{-\mu}\frac{\mu^{n-i}}{(n-i)!} = \frac{\mathrm{e}^{-(\lambda+\mu)}}{n!}\sum_{i=0}^{n}\binom{n}{i}\lambda^i\mu^{n-i} = \mathrm{e}^{-(\lambda+\mu)}\frac{(\lambda+\mu)^n}{n!}$$

C – Familles infinies de variables aléatoires

On considère une famille infinie $(X_i)_{i\in I}$ de variables aléatoires sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$. Alors,

- les variables X_i sont dites indépendantes si pour toute partie finie J de I, la famille $(X_i)_{i\in J}$ est indépendante.
- si les variables X_i suivent de plus toutes la même loi, on dira que $(X_i)_{i\in I}$ est une famille de variables indépendantes et identiquement distribuées ou, en abrégé, une famille de variables i.i.d.
- la suite $(X_n)_{n\in\mathbb{N}}$ est qualifiée de processus à temps discret, quand la famille $(X_t)_{t\in\mathbb{R}_+}$ est qualifiée de processus à temps continu.

L'existence d'espaces probabilisés portant des familles infinies de variables aléatoires discrètes indépendantes est un enjeu majeur dans la modélisation de phénomènes aléatoires comme le jeu du *pile ou face* infini. Le théorème suivant, dont la démonstration est inhérente à la construction de la mesure de Lebesgue, est admis.

Théorème 10.66 : Théorème d'extension de Kolmogorov

Soit $(\mathscr{L}_n)_{n\geqslant 1}$ une suite de lois discrètes sur des ensembles E_n. Il existe un espace probabilisé $(\Omega, \mathscr{A}, \mathbf{P})$ et une suite $(X_n)_{n\geqslant 1}$ de variables aléatoires indépendantes telles que pour tout $n \geqslant 1$, $X_n \sim \mathscr{L}_n$.

Ce théorème valide, dans un énoncé, toute formulation de la forme : « On considère une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes et identiquement distribuées, de loi [...] ».

Exemple – loi géométrique, le retour

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables de Bernoulli de paramètre $p \in]0,1]$ indépendantes. On pose :

$$T = \min\left\{n \in \mathbb{N}^* \cup \{+\infty\} \mid X_n = 1\right\}$$

- T est bien une variable aléatoire. Il suffit de vérifier que $T(\Omega)$ est dénombrable et que pour tout $n \in \mathbb{N}^*$, $(T=n) = (X_n=1) \cap \bigcap_{k=1}^{n-1}(X_k=0) \in \mathscr{A}$ et $(T=+\infty) = \bigcap_{k=1}^{+\infty}(X_k=0) \in \mathscr{A}$.
- Pour tout $n \in \mathbb{N}^*$, $\mathbf{P}(T=n) = \mathbf{P}(X_1=\cdots=X_{n-1}=0, X_n=1) = q^{n-1}p$ par indépendance.
- De plus, $\mathbf{P}(T=+\infty) = 1-\mathbf{P}(T\in\mathbb{N}^*) = 1-\sum_{k=1}^{+\infty}\mathbf{P}(T=k) = 0$. Concluons : $T \sim \mathscr{G}(p)$.