Chapitre I

Fondements du calcul des probabilités

Or du Hasard il n'est point de science :
S'il en était, on aurait tort
De l'appeler hasard, ni fortune, ni sort,
Toutes choses très incertaines.
(Jean de La Fontaine)

I.1 Probabilité ; conditionnement

Une probabilité est d'abord une *fonction* qui à un *événement* associe un nombre réel compris entre 0 et 1. Cela implique de préciser ce qu'est un événement. Or cela n'a de sens que dans le cadre d'un ensemble d'*épreuves aléatoires* ou *tirages*, qu'on note généralement Ω. Il peut s'agir par exemple de lancers de dés ou de pièces de monnaie, de tirages d'urne, de durées de vie (d'atomes ou d'individus), de tirs sur une cible, etc.. Ces premiers exemples familiers montrent déjà que l'ensemble Ω peut être fini, dénombrable (ce qui signifie : infini indexé par $\mathbb{N}$ ou $\mathbb{N}^*$), ou continu. Ce sera donc a priori un ensemble non vide quelconque.

Lorsque Ω est fini ou *dénombrable* (c'est dire que ses éléments forment une suite), toutes ses parties seront des événements. Tandis qu'en général il est nécessaire de se restreindre à un sous-ensemble de parties de Ω : $\mathcal{T} \subset \mathcal{P}(\Omega)$, qu'on nomme *tribu* (ou *σ-algèbre*). On a naturellement besoin de pouvoir considérer la réunion et la conjonction (l'intersection) de 2 événements, de même que le complémentaire (le contraire) d'un événement ; en outre il faut aussi pouvoir considérer une réunion dénombrable d'événements. Enfin il est naturel de considérer l'événement impossible (vide : $\emptyset$) et l'événement certain (sûr : Ω). D'où la définition suivante.

Définition I.1.1 *Une* tribu *(ou σ-*algèbre*) est une partie $\mathcal{T}$ de $\mathcal{P}(\Omega)$ stable par réunion dénombrable et par passage au complémentaire, et contenant l'ensemble vide $\emptyset$. Le couple $(\Omega, \mathcal{T})$ est appelé* espace probabilisable*. Un* événement *est un élément de $\mathcal{T}$.*

$\mathcal{P}(\Omega)$ désigne l'ensemble de toutes les parties de Ω ; $\mathcal{T}$ est donc un ensemble de parties de Ω. La stabilité par réunion dénombrable s'écrit formellement : pour toute suite d'événements $\{E_n \,|\, n \in \mathbb{N}\} \subset \mathcal{T}$, leur réunion est aussi un événement : $\bigcup\limits_{n\in\mathbb{N}} E_n \in \mathcal{T}$.

La stabilité par passage au complémentaire s'écrit formellement : le complémentaire $E^{\mathbf{c}} := \Omega \setminus E = \{\omega \in \Omega \,|\, \omega \notin E\}$ de tout événement E est aussi un événement : $E^{\mathbf{c}} \in \mathcal{T}$.

Nota Bene Sur Ω fini ou dénombrable, on choisira toujours par défaut la tribu $\mathcal{P}(\Omega)$. Rappel : $\Big[\bigcap\limits_n A_n\Big]^{\mathbf{c}} = \bigcup\limits_n A_n^{\mathbf{c}}$; $\Big[\bigcup\limits_n A_n\Big]^{\mathbf{c}} = \bigcap\limits_n A_n^{\mathbf{c}}$. On vérifie aussitôt les propriétés suivantes :

Proposition I.1.2 *Ω est un événement (certain). La réunion et l'intersection d'un nombre fini d'événements sont des événements. La différence $A \setminus B := A \cap B^c$ et la différence symétrique $A \Delta B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$ de deux événements A et B sont des événements. Toute intersection dénombrable d'événements est un événement.*

Nous pouvons maintenant définir rigoureusement ce qu'est une probabilité ; dans la pratique, ce sera le plus souvent soit une somme (soit finie, soit infinie, c'est-à-dire une série) ou une intégrale (soit propre, dite de Riemann, soit impropre). (Pour rassembler toutes ces possibilités, et d'autres encore, on parle de *mesure*.)

Définition I.1.3 *Une* probabilité *sur l'espace probabilisable $(\Omega, \mathcal{T})$ est une fonction $\mathbb{P}$ de $\mathcal{T}$ dans $[0,1]$ qui vérifie : $\mathbb{P}(\Omega) = 1$, et la propriété d'additivité dénombrable : $\mathbb{P}\Big(\bigsqcup\limits_{n\in\mathbb{N}} A_n\Big) = \sum\limits_{n\in\mathbb{N}} \mathbb{P}(A_n)$ pour tout suite $\{A_n \,|\, n \in \mathbb{N}\}$ d'événements deux à deux disjoints. Le triplet $(\Omega, \mathcal{T}, \mathbb{P})$ est appelé* espace probabilisé *ou* espace de probabilité*. Les événements de probabilité nulle sont dits* négligeables*. Les événements de probabilité 1 sont dits* presque sûrs *(et « $A = B$ p.s. » signifie que $A\Delta B$ est négligeable).*

C'est toujours dans le cadre d'un espace de probabilité, plus ou moins bien précisé, que peut avoir lieu un calcul de probabilité. Il est généralement préférable de bien le préciser. En effet c'est la non-précision de l'espace considéré qui est à l'origine de paradoxes ou d'erreurs courantes sur les probabilités. On vérifie aisément les propriétés suivantes :

Proposition I.1.4 *(i) L'événement impossible est négligeable:* $\mathbb{P}(\emptyset) = 0$.

(ii) $\mathbb{P}$ *est croissante :* $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$, *pour tous* $A, B \in \mathcal{T}$.

(iii) $\mathbb{P}\Big[\bigcup_{n\in\mathbb{N}} A_n\Big] \leq \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$ *pour tout suite* $\{A_n \,|\, n \in \mathbb{N}\}$ *d'événements.*

(iv) $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(B)$, *lorsque* A *et* B *sont deux événements tels que* $B \subset A$. *En particulier,* $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ *pour tout événement* A.

(v) Toute intersection dénombrable d'événements presque sûrs est presque sûre.

(vi) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$, *pour tous* $A, B \in \mathcal{T}$.

(vii) $\mathbb{P}$ *est continue :* $\mathbb{P}\Big[\bigcup_{n\in\mathbb{N}} A_n\Big] = \lim_{n\to\infty} \mathbb{P}(A_n)$ *pour toute suite croissante d'événements* $\{A_n \,|\, n \in \mathbb{N}\}$ *(c'est-à-dire telle que* $A_n \subset A_{n+1}$ *pour tout* n*) ; et de même* $\mathbb{P}\Big[\bigcap_{n\in\mathbb{N}} B_n\Big] = \lim_{n\to\infty} \mathbb{P}(B_n)$ *pour toute suite décroissante d'événements* $\{B_n \,|\, n \in \mathbb{N}\}$ *(c'est dire que* $B_n \supset B_{n+1}$ *pour tout* n*).*

Exercice I.1.5 • *Établir (dans l'ordre) les propriétés de la proposition I.1.4.*

I.1.6 Exemples

1. Probabilité discrète sur un ensemble $\Omega = \{\omega_1, \ldots, \omega_N\}$ fini :

elle est clairement définie par la liste des probabilités des singletons : $p_j := \mathbb{P}(\{\omega_j\})$. Nous avons en effet $\mathbb{P}(A) = \sum_{\omega_j \in A} p_j$ pour toute partie $A \subset \Omega$.

Réciproquement, toute liste $\{p_1, \ldots, p_N\}$ de réels $p_j \geq 0$ tels que $\sum_{j=1}^{N} p_j = 1$ définit bien (par la même formule) une probabilité unique sur Ω.

Exemples concrets : lancers de dés, de pièces de monnaie, tirages de cartes à jouer, tirages de boules dans des urnes, loteries, etc...

2. Probabilité discrète sur $\mathbb{N}$ (ou sur tout autre ensemble dénombrable) : elle est encore définie par la liste des probabilités des singletons : $p_j := \mathbb{P}(\{\omega_j\})$. Nous avons en effet $\mathbb{P}(A) = \sum_{\omega_j \in A} p_j$ pour toute partie $A \subset \mathbb{N}$. La seule différence avec le cas précédent est que la somme peut être une série (c'est-à-dire comporter une infinité de termes).

Réciproquement, toute suite $\{p_j \mid j \in \mathbb{N}\}$ de réels $p_j \geq 0$ tels que $\sum\limits_{j\geq 1} p_j = 1$ définit bien (par la même formule, forcément convergente) une probabilité unique sur $\mathbb{N}$.

3. *Cordes* (paradoxe de Bertrand). On tire une corde au hasard dans un disque de rayon R. Quelle est la probabilité que la longueur ℓ de la corde soit $\geq R$?

a. ℓ varie continûment dans $[0, 2R]$, de sorte que la probabilité cherchée vaut $1/2$.

b. ℓ est déterminée par la distance d de la corde au centre du disque ; d varie continûment dans $[0, R]$, et $\ell = 2\sqrt{R^2 - d^2} \geq R \Leftrightarrow d \leq R\sqrt{3}/2$, de sorte que la probabilité cherchée vaut $\sqrt{3}/2$.

c. ℓ est déterminée par le milieu M de la corde, qui varie continûment dans le disque ; et $\ell \geq R$ a lieu si et seulement si M est dans le disque concentrique de rayon $\sqrt{3}/2$, de sorte que la probabilité cherchée vaut $3/4$.

Explication : la probabilité choisie est très insuffisamment précisée par l'expression « tirage au hasard ». Ici on a considéré successivement la probabilité uniforme sur l'ensemble : des longueurs, des distances au centre, des milieux. Ce sont trois probabilités différentes !

4. Jeu de pile ou façe illimité ; première apparition de « pile », ou d'une quelconque séquence donnée.

Exercice I.1.7 *Une armoire contient 10 paires de chaussures, parmi lesquelles on prélève au hasard 8 chaussures. Quelle est la probabilité d'avoir ainsi k paires de chaussures exactement ?*

Exercice I.1.8 *Une urne contient n boules noires et b boules blanches. Deux joueurs X et Y tirent avec remise une boule dans l'urne, tour à tour, X tirant le premier. Quelle est la probabilité que X soit le premier à tirer une boule noire ? Même question lorsque les tirages sont effectués sans remise. (On ne cherchera à calculer la somme obtenue que dans le premier cas.)*

Exercice I.1.9 *Un joueur X lance 2 dés usuels, et obtient ainsi la somme S.*

a) Calculer la probabilité de $\{S > n\}$, en fonction des différentes valeurs de l'entier n.

b) Un joueur Y relance les 2 dés et obtient une somme T. Quelles sont les probabilités que $S = T$, que $S > T$, que $S \geq T$?

Proposition I.1.10 *(Formule du crible, de Poincaré) Pour tout espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$, tout entier $n \in \mathbb{N}^*$, et tous événements $A_1, \dots, A_n$, nous avons :*

$$\mathbb{P}\Big[A_1 \cup \ldots \cup A_n\Big] = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 \leq \ldots \leq i_k \leq n} \mathbb{P}\Big[A_{i_1} \cap \ldots \cap A_{i_k}\Big].$$

Exercice I.1.11 *Prouver la proposition I.1.10, par récurrence sur n.*

Exercice I.1.12 *Un sac contient n jetons numérotés de 1 à n, qu'on tire tous un par un, sans remise. a) Calculer la probabilité p_n qu'au moins un jeton sorte au rang indiqué par son numéro, sa limite p_∞, et majorer $|p_n - p_\infty|$.*

b) Soit $p_n(k) := \mathbb{P}$(exactement k jetons sortent au rang indiqué par leur numéros), pour $k \in \{0, \dots, n\}$. Déduire de a) ci-dessus une formule pour $p_n(k)$, puis $\lim\limits_{n \to \infty} p_n(k)$ (pour tout $k \in \mathbb{N}$).

I.1.13 Probabilités conditionnelles

Définition I.1.14 (Probabilité conditionnelle) *Fixons un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ et un événement $C \in \mathcal{T}$, non négligeable. La probabilité conditionnelle relative à C (ou « sachant C ») est définie par :*

$$\mathbb{P}(A/C) := \mathbb{P}(A \cap C)/\mathbb{P}(C).$$

On vérifie immédiatement qu'il s'agit encore d'une probabilité sur $(\Omega, \mathcal{T})$.

Exercice I.1.15 *Lancer de 2 dés usuels : $\Omega = \{1, \dots, 6\}^2$, avec $\mathbb{P}$ uniforme. Soient X_1 le chiffre indiqué par le premier dé, S la somme des chiffres indiqués par les 2 dés, et $C := \{S = 5\}$. Dresser le tableau des valeurs de $\mathbb{P}(\cdot/C)$, puis de $\mathbb{P}(X_1 = \cdot/C)$.*

Exercice I.1.16 *Vous attendez un ami de Vancouver, qui voyage jusqu'à Strasbourg avec changement d'avion à New York, Londres et Francfort. La probabilité d'attentat est estimée à p pour chacun des 4 vols, avec indépendance entre les 4 vols. Votre ami n'arrivant pas, quelle est la probabilité que l'attentat dont il a été victime ait eu lieu : a) dans le 1^er^ avion ? b) dans le 2^e^ avion ? c) dans le 3^e^ avion ? d) dans le 4^e^ avion ? Pourquoi la somme ne vaut-elle pas 1 ?*

Pour effectuer un calcul, il est très souvent indispensable de pouvoir « distinguer des cas ». Cela s'exprime par la formule suivante, très élémentaire et très utile à la fois.

Proposition I.1.17 (Formule des probabilités totales) *Fixons un espace de probabilité* $(\Omega, \mathcal{T}, \mathbb{P})$ *et une* partition (presque sûre) *de* Ω *en événements non négligeables : presque sûrement* $\Omega = \bigsqcup_{j=1}^{N} C_j$ *(et* $1 \leq j < k \leq N \Rightarrow C_j \cap C_k = \emptyset$*).*
Alors nous avons $\mathbb{P}(A) = \sum_{j=1}^{N} \mathbb{P}(A/C_j)\,\mathbb{P}(C_j)$ *, pour tout événement* A*. Cela vaut aussi pour toute partition dénombrable : si* $\Omega = \bigsqcup_{j\in\mathbb{N}} C_j$ *presque sûrement, alors* $\mathbb{P}(A) = \sum_{j\in\mathbb{N}} \mathbb{P}(A/C_j)\,\mathbb{P}(C_j)$*.*

On a souvent à inverser un conditionnement. Cela se fait simplement, au moyen de la formule élémentaire suivante, très utile aussi, quoiqu'également de preuve immédiate.

Proposition I.1.18 (Formule de Bayes) *Fixons un espace de probabilité* $(\Omega, \mathcal{T}, \mathbb{P})$*, et une partition presque sûre de* Ω *en événements non négligeables :* $\Omega = \bigsqcup_{j=1}^{N} C_j$ *presque sûrement. Alors nous avons pour tout événement non négligeable* A *et tout* $k \in \{1, \ldots, N\}$ *:*

$$\mathbb{P}(C_k/A) = \mathbb{P}(A/C_k)\,\mathbb{P}(C_k) \Big/ \sum_{j=1}^{N} \mathbb{P}(A/C_j)\,\mathbb{P}(C_j)\,.$$

Exemple : Les candidats à un examen proviennent de 4 lycées K,L,V,W, à raison de 20% pour K, de 25% pour L, de 15% pour V, et de 40% pour W. K enregistre 35% de succès, L 30%, V 50%, W 45%. Alors la probabilité qu'un candidat reçu provienne de K est $\frac{7}{40}$, de L est $\frac{3}{16}$, de V est $\frac{3}{16}$, de W est $\frac{9}{20}$.

Exercice I.1.19 • *Vérifier que la probabilité conditionnelle* $\mathbb{P}(\cdot/C)$ *de la définition I.1.14 est bien une probabilité sur* $(\Omega, \mathcal{T})$*, et établir les propositions I.1.17 et I.1.18. Vérifier que conditionner successivement par deux événements* C *et* C' *d'intersection non négligeable revient à conditionner par leur intersection :* $\mathbb{P}\big(\,\cdot\,/C/C'\big) = \mathbb{P}(\cdot/C \cap C')$*.*

Exercice I.1.20 *Vous allez chez des gens dont vous savez qu'ils ont 2 enfants, dont au moins une fille. a) Quelle est la probabilité que l'autre enfant soit aussi une fille ?*

b) Une fille vous ouvre la porte ; quelle est la probabilité que l'autre enfant soit aussi une fille ?

Exercice I.1.21 *Trois condamnés X, Y, Z sont informés que l'un d'eux, choisi au hasard, va être exécuté, et que les 2 autres vont être libérés. Mais ils ne doivent pas encore savoir qui le hasard a désigné. X demande au geôlier de lui nommer l'un de ses 2 codétenus devant être libéré, arguant que cette information serait innocente, puisqu'il sait que l'un des 2 au moins doit l'être. Le geôlier refuse, arguant que cette information modifierait réellement l'estimation que X peut faire de ses chances. Qui a raison ?*

Exercice I.1.22 *Un livre a une probabilité $p > 0$ de se trouver dans une commode comportant k tiroirs, et des chances égales de se trouver dans chacun des tiroirs.*

a) ([J]) On ouvre les $(k-1)$ premiers tiroirs, sans le trouver ; quelle est la probabilité de le trouver dans le dernier tiroir ?

b) Soit $j \in \{2, \ldots, k-1\}$. On ouvre les $(k-j)$ premiers tiroirs, sans le trouver ; quelle est la probabilité de le trouver dans le dernier tiroir ? dans l'un des j derniers tiroirs ?

Exercice I.1.23 *Le quart d'une population est vacciné contre le choléra. Au cours d'une épidémie, on constate qu'il y a parmi les malades un vacciné pour 4 non-vaccinés, et qu'il y a un malade sur 12 parmi les vaccinés. Quelle est la probabilité qu'un non-vacciné tombe malade ? Le vaccin est-il efficace ?*

I.1.24 Événements indépendants

L'indépendance est au cœur de la théorie des probabilités. Abordons-la doucement, en commençant par le cas des événements. Le cas général des variables aléatoires viendra ensuite (section I.2.20).

Définition I.1.25 *Fixons un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$. Deux événements A et B sont dits* indépendants *lorsque* $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$.

Exemples : 1) « tirer un roi » et « tirer un trèfle », dans un jeu de bridge ou de belote.

2) Lors du lancer de 2 dés : « obtenir un chiffre pair avec le 1^er^ dé » et « obtenir un 6 avec le 2^e^ dé » $\equiv \{X_2 = 6\}$; noter que ces tirages sont indépendants au sens usuel.

3) Lors du lancer de 2 dés : ≪ obtenir une somme paire ≫ et ≪ obtenir 5 avec le 2e dé ≫.

4) Lors du lancer de 2 dés : ≪ obtenir une somme S égale à 7 ≫ $\equiv \{S = 7\}$ et ≪ obtenir un produit égal à 12 ≫ ne sont pas indépendants.

5) Lors du lancer de 2 dés : $\{S = 7\}$ et $\{X_2 = 2\}$, ou bien $\{S = 6\}$ et $\{X_2 = 2\}$.

6) Lors du lancer de 2 dés : ≪ obtenir un produit pair ≫ et ≪ obtenir 5 avec le 2e dé ≫.

Exercice I.1.26 *Montrer que 2 événements A et B sont indépendants si et seulement si A et B^c le sont, et si et seulement si A^c et B^c le sont, ou bien encore (lorsqu'ils sont non négligeables) si et seulement si $\mathbb{P}(A/B) = \mathbb{P}(A)$, et si et seulement si $\mathbb{P}(B/A) = \mathbb{P}(B)$.*

Définition I.1.27 *Fixons un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$. Des événements $A_1, \ldots, A_n$ sont dits* indépendants *lorsque*

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = \mathbb{P}(A_{i_1})\,\mathbb{P}(A_{i_2}) \ldots \mathbb{P}(A_{i_k})$$

pour tous $1 \leq i_1 < \ldots < i_k \leq n$. Les événements d'une suite sont dits indépendants *lorsque toute sous-suite finie est constituée d'événements indépendants.*

Proposition I.1.28 *Les événements $A_1, \ldots, A_n$ sont indépendants si et seulement si* $\mathbb{P}\big(A_1^{\varepsilon_1} \cap A_2^{\varepsilon_2} \cap \ldots \cap A_n^{\varepsilon_n}\big) = \mathbb{P}(A_1^{\varepsilon_1})\,\mathbb{P}(A_2^{\varepsilon_2}) \ldots \mathbb{P}(A_n^{\varepsilon_n})$ *pour tout choix de $\varepsilon_1, \ldots, \varepsilon_n$, avec soit $A_j^{\varepsilon_j} = A_j$ soit $A_j^{\varepsilon_j} = A_j^{\mathbf{c}}$.*

Preuve. Pour le sens direct, on vérifie par récurrence sur $k \in \{0, \ldots, n\}$ que pour tout choix de $\varepsilon_1, \ldots, \varepsilon_k$, tout $\ell \in \{k+1, \ldots, n\}$ et tous $i_{k+1}, \ldots, i_\ell \in \{k+1 \ldots, n\}$, on a :

$$P\Big[A_1^{\varepsilon_1} \cap \ldots \cap A_k^{\varepsilon_k} \cap A_{i_{k+1}} \cap \ldots \cap A_{i_\ell}\Big] = P[A_1^{\varepsilon_1}] \ldots \mathbb{P}[A_k^{\varepsilon_k}] \times \mathbb{P}[A_{i_{k+1}}] \ldots \mathbb{P}[A_{i_\ell}] \,.$$

Pour la réciproque, si $i_1, \ldots, i_k$ sont fixés dans $\{1, \ldots, n\}$, on note $\{j_1, \ldots, j_{n-k}\} := \{1, \ldots, n\} \setminus \{i_1, \ldots, i_k\}$, et on applique la formule des probabilités totales avec la partition $\big\{A_{j_1}^{\varepsilon_{j_1}} \cap \ldots \cap A_{j_{n-k}}^{\varepsilon_{j_{n-k}}} \,\big|\, \varepsilon_{j_1}, \ldots, \varepsilon_{j_{n-k}} = \quad \text{ou } c\big\}$ (c'est bien une partition de Ω, car l'ensemble des intersections des éléments de 2 partitions en forme encore une). $\diamond$