JEAN-PIERRE LECOUTRE

ÉCOSUP

Statis-tique et proba-bilités

7E ÉDITION

DUNOD

Éditorial : Guillaume Clapeau
Fabrication : Nelly Roushdi
Conception de couverture : Studio Dunod
Mise en pages : Lumina

11, rue Paul Bert, 92240 Malakoff
www.dunod.com

ISBN 978-2-10-084802-7

Table des matières

Avant-propos

Ce manuel de cours est destiné principalement aux étudiants de la Licence économie et gestion mais peut être utile à toute personne souhaitant connaître et surtout utiliser les principales méthodes de la statistique inférentielle. Il correspond au programme de probabilités et statistique généralement enseigné dans les deux premières années de Licence (L1 et L2). Cette 7e édition s'est enrichie d'exercices nouveaux. Le niveau mathématique requis est celui de la première année de Licence, avec quelques notions (séries, intégrales multiples...) souvent enseignées seulement en deuxième année.

Si une grande partie de l'ouvrage est consacrée à la théorie des probabilités, l'ordre des termes retenu dans le titre veut signifier qu'il ne s'agit que d'un outil au service de la statistique. Ce n'est qu'un passage obligé pour donner des bases rigoureuses à la méthode statistique. On peut le concevoir comme un ensemble de règles grammaticales, parfois difficiles et fastidieuses à retenir, mais qui permettent de rédiger des textes clairs, rigoureux et sans ambiguités, même si l'on n'a pas conscience qu'ils ont été écrits dans le respect de ces règles. La partie statistique correspond aux deux derniers chapitres d'estimation et de tests d'hypothèses.

Les fondements théoriques de la statistique étant parfois délicats, nous avons choisi de présenter sans démonstration les principales propriétés nécessaires à une utilisation judicieuse des méthodes statistiques, en les illustrant systématiquement d'exemples. De même, afin de ne pas alourdir les énoncés de théorèmes, les conditions techniques de leur validité ne sont pas présentées dans leur détail, parfois fastidieux, et qui risque de masquer l'essentiel qui est la propriété énoncée. Notre souci constant a été de faciliter la compréhension, pour pouvoir passer aisément au stade de l'utilisation, sans cependant pour cela sacrifier à la rigueur. La traduction anglaise des termes les plus usuels figure entre parenthèses.

Chaque chapitre se conclut par des exercices corrigés permettant de contrôler l'acquisition des notions essentielles qui y ont été introduites. Faire de nombreux exercices est certainement le meilleur moyen d'arriver à la compréhension de certaines notions quelquefois difficiles. Rappelons cette maxime chinoise : *J'entends et j'oublie. Je vois et je retiens. Je fais et je comprends.* En fin de chapitre se trouvent également quelques compléments ; soit de notions mathématiques utilisées dans celui-ci, la combinatoire par exemple, soit de propriétés comme l'exhaustivité, très importantes et utiles, mais hors du programme d'une Licence d'économie ou de gestion. Avec ces compléments, cet ouvrage peut convenir aussi aux étudiants des écoles de management.

Notations

Ω	Ensemble fondamental
$\mathcal{P}(\Omega)$	Ensemble des parties de Ω
$\overline{A}$, A^c	Complémentaire de A
$\mathcal{A}$	Algèbre ou tribu de parties de Ω
$card\ A$	Cardinal de A
$\binom{n}{p}$	Coefficient binômial
$[x]$	Partie entière de x
$\ln x$	Logarithme népérien de x
$\mathbf{1}_A$	Indicatrice de A
$Cov\ (X,Y)$	Covariance de X et Y
f.r.	Fonction de répartition
v.a.	Variable aléatoire
φ	Densité de la loi $N(0,1)$
Φ	F.r. de la loi loi $N(0,1)$
$\mathbb{C}$	Ensemble des nombres complexes
${}^tA$	Matrice transposée de A
I_n	Matrice unité d'ordre n
$X \rightsquigarrow P$	La v.a. X suit la loi de probabilité P
$\mathcal{B}(n, p)$	Loi binômiale de paramètres n et p
$\mathcal{P}(\lambda)$	Loi de Poisson de paramètre λ
$N\ (m,\sigma)$	Loi normale dans $\mathbb{R}$, d'espérance m et d'écart type σ
$\mathcal{N}_n(\mu, \Sigma)$	Loi normale dans $\mathbb{R}^n$, de vecteur espérance μ et de matrice variances-covariances Σ
T_n	Loi de Student à n degrés de liberté
χ^2_n	Loi du khi-deux à n degrés de liberté
$F(n,m)$	Loi de Fisher-Snedecor à n et m degrés de liberté
emv	Estimateur du maximum de vraisemblance

Introduction

La statistique a une origine très ancienne, se réduisant initialement à une collecte d'observations, notamment le dénombrement des hommes (recensement). On mentionne des opérations de recensement il y a plus de 4 000 ans en Chine, en Mésopotamie ou en Égypte et la Bible en cite plusieurs, dans le Livre des Nombres par exemple. Cependant, le terme statistique est apparu assez récemment, vers le milieu du XVII^{e} siècle ; il vient du latin *statisticus,* relatif à l'état *(status),* et est employé alors dans un sens purement descriptif de recueil ou de collection de faits chiffrés, les **statistiques**. Le mot employé au singulier avec l'article défini, la **statistique**, évoque la méthode utilisée ensuite pour étendre des résultats et dégager des lois (**l'inférence**). Il s'agit donc dans ce sens d'un moyen scientifique d'analyse et de compréhension du phénomène étudié, s'appliquant très largement à l'économie et à toutes les sciences sociales et de la nature.

Cette discipline concerne donc tous ceux qui ont à relever, présenter, analyser ou utiliser une information dont la masse peut être volumineuse. On peut la définir comme un ensemble de méthodes dont le but est de traiter des données, les **statistiques**, relatives à un certain domaine d'étude. Elle traite également de la manière de recueillir ces données, auprès de qui et sous quelle forme (théorie des sondages). Son objectif peut se résumer de la façon suivante : dégager, à partir de données observées sur quelques individus d'une population, des résultats valables pour l'ensemble de la population.

Cela consistera par exemple à remplacer des données nombreuses par des indicateurs (résumés) les plus pertinents possibles : résumé clair avec le minimum de perte d'information, permettant de dégager plus facilement un diagnostic. Il s'agit alors de la **statistique descriptive** qui recouvre les moyens de présenter ces données et d'en décrire les principales caractéristiques, en les résumant sous forme de tableaux ou de graphiques. Il s'agira ensuite de les interpréter. La description statistique se propose de mettre en évidence certaines permanences ou **lois statistiques**, qui peuvent éventuellement conduire à des **prévisions** (élément essentiel de l'étude des séries chronologiques). Une règle qui transforme un ensemble de données en une ou plusieurs valeurs numériques se nomme une **statistique**, le terme étant cette fois utilisé avec l'article indéfini.

Le début de la méthodologie statistique peut se situer au XVII^{e} siècle qui verra également l'éclosion d'un outil fondamental pour une formalisation tout à fait rigoureuse, la **théorie des probabilités**, qui est l'analyse mathématique des phénomènes dans lesquels le hasard intervient. Le **calcul des probabilités** a commencé avec Blaise Pascal, Pierre Fermat, Christian Huygens et Jacques Bernoulli par l'analyse des jeux dits de hasard. Le mot **hasard** est d'ailleurs emprunté à l'arabe *az-zahr* (jeu de dés, *alea* en latin) au XII^{e} siècle d'où est venue cette expression jeu de hasard au XVI^{e} siècle. La théorie des probabilités servira ensuite d'outil de base à un ensemble de méthodes ou de règles objectives permettant d'utiliser des données pour fixer la précision avec laquelle on estime certains paramètres (théorie de l'estimation) ou on teste certaines hypothèses (théorie des tests) : la **Statistique mathématique** (ou inférentielle). Ceci permet d'obtenir une mesure objective de la distance entre un modèle statistique, traduit par une famille P_θ de lois de probabilité indexée par un paramètre θ parcourant un ensemble donné Θ, et un ensemble de données observées.

Tout ceci peut se synthétiser au moyen du schéma suivant :

Données → Importance de la manière de les collecter
(théorie des sondages)
→ Présentation des données recueillies
(statistique descriptive)

⇓

Modélisation → Catalogue de modèles probabilistes disponibles
et outils nécessaires à la déduction
(théorie des probabilités)

⇒ **Statistique mathématique :**
un modèle statistique paramétrique $(P_\theta ; \theta \in \Theta)$
induction ou inférence statistique
- estimation : quelle est le valeur de θ ?
- test : est-ce que $\theta = \theta_0$ ou $\theta = \theta_1$?

Il reste à préciser dans quel cadre cette formalisation à l'aide de modèles aléatoires sera nécessaire. Toute démarche scientifique nécessite la réalisation de certaines expériences que l'on peut regrouper en deux grandes catégories :

- Pour certaines d'entre elles, si elles sont renouvelées dans des conditions totalement identiques, elles produiront le même résultat, qui devient donc prévisible. Il s'agit de **phénomènes déterministes**, où les faits sont régis par des lois universelles physiques (par exemple l'augmentation de la pression d'un gaz provoque une diminution de son volume, ce que traduit la loi de Mariotte : Pression × Volume = constante ; l'eau portée à 100 degrés Celsius se transforme en vapeur...). Le résultat est entièrement déterminé par les conditions de l'expérience : on peut prévoir le phénomène qui va se produire.
- Par contre, d'autres expériences ont toujours un résultat imprévisible (lancer d'un dé ou d'une pièce de monnaie) : effectuées dans des conditions totalement identiques elles donneront des résultats différents. Le résultat est non prévisible et on dit qu'il est dû **au hasard**, cette expression étant utilisée pour la première fois par Fénelon en 1695, le mot hasard étant compris maintenant au sens absolu et philosophique comme « sans évolution prévisible », à opposer à déterministe. Dans son *Essai philosophique sur les probabilités* (1814), Laplace considère en effet que le déterminisme ne laisse aucune place au hasard : l'état de l'univers à un instant donné détermine son état

à tout autre instant ultérieur. Ainsi, quand on jette une pièce de monnaie en l'air, les lois de la mécanique classique déterminent, en principe, avec certitude si elle retombera sur pile ou face. Le résultat n'est pas dû au hasard, mais à la manière dont elle a été lancée en l'air et à la façon dont elle va retomber sur une certaine surface ; mais la trajectoire décrite par cette pièce avant de retomber sur pile est tellement complexe qu'il n'est pas possible de prévoir son issue. Le phénomène ne relève pas du déterminisme entendu au sens de la possibilité de prédiction, par le calcul ou la loi mathématique.

Dans un mémoire de 1774, Laplace énonce que « le hasard n'a aucune réalité en lui-même : ce n'est qu'un terme propre à désigner notre ignorance... La notion de probabilité tient à cette ignorance ». Retenir un modèle probabiliste est donc simplement un aveu de notre ignorance, de notre incapacité à fournir un modèle physique décrivant une réalité trop complexe. On parle alors d'**épreuve** ou d'**expérience aléatoire** et le résultat obtenu sera un **événement**. Les outils appropriés dans ce cadre sont ceux de la **statistique mathématique**, la base de cette discipline étant la **théorie des probabilités**, que nous devrons donc étudier dans les six premiers chapitres de cet ouvrage, comme préalable aux deux chapitres d'estimation et de tests d'hypothèses.

REMARQUE Dans *Science et méthode* publié en 1908, Henri Poincaré exprime que hasard et déterminisme sont rendus compatibles par l'imprédictibilité à long terme. Les relations entre hasard et déterminisme ont été dans les années 1980 l'objet d'une controverse animée entre les mathématiciens René Thom et Ilya Prigogine. L'étude récente des systèmes dynamiques montre que l'on ne peut pas confondre déterminisme et prédictibilité. En effet, une légère perturbation des conditions initiales d'un tel système mathématiquement déterministe peut empêcher de prévoir son évolution future.

Chapitre 1 Notion de probabilité

Introduction

Au cours de ce chapitre, nous allons donner la définition d'un certain nombre de termes du vocabulaire utilisé dans un contexte non déterministe et indiquer comment construire le modèle adéquat. La notion essentielle introduite étant bien sûr celle de probabilité, avec la notion d'indépendance d'événements qui lui est associée et qui joue un rôle très important en statistique. La représentation formelle du modèle probabiliste sous-jacent est presque toujours absente dans un problème concret de statistique. Cependant, cette formalisation rigoureuse est indispensable pour obtenir les outils théoriques nécessaires à la résolution d'un tel problème statistique.

Objectifs

Montrer que le modèle probabiliste est choisi en fonction du but que l'on poursuit.

Définir de façon théorique (et non intuitive) la notion d'indépendance.

Concepts clés

Probabilité

Probabilité conditionnelle

Indépendance

1 Modèle probabiliste

1.1 Ensemble fondamental

Avant toute formalisation, le résultat d'une expérience aléatoire s'appelle **événement**. La quantification des « chances » qu'un tel événement a de se réaliser correspond à la notion intuitive de **probabilité**. Pour réaliser cette quantification, il est nécessaire de décrire au préalable, très précisément, l'ensemble des résultats possibles, appelés **événements élémentaires**. Cet ensemble expérimental s'appelle **ensemble fondamental** (ou univers) et est noté traditionnellement Ω.

Exemple 1.1

Jet d'un dé à six faces numérotées : $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Exemple 1.2

On tire une boule dans une urne contenant une boule noire, deux blanches et cinq rouges et l'ensemble fondamental retenu est $\Omega = \{\text{noire, blanche, rouge}\}$.

Chaque élément $\omega \in \Omega$ représente donc un événement élémentaire, et toute partie $A \subset \Omega$ (ou $A \in \mathcal{P}(\Omega)$) sera un événement. Parfois on dit que Ω est l'ensemble des **éventualités possibles** et les événements élémentaires sont alors les singletons, c'est-à-dire les ensembles réduits à un seul élément $\{\omega\}$, qui sont effectivement en toute rigueur des événements, puisque appartenant à $\mathcal{P}(\Omega)$, ce qui n'est pas le cas du point ω.

Exemple 1.3

À l'expérience du jet de dé on associe $\Omega = \{1, 2, 3, 4, 5, 6\}$; l'événement $A = \{1, 2\}$ traduit, c'est-à-dire représente symboliquement, le résultat « obtenir un résultat inférieur ou égal à 2 ».

L'ensemble Ω dépend évidemment de l'expérience considérée, mais aussi du choix de celui qui construit le modèle, et par là présente donc un certain arbitraire.

Exemple 1.4

Dans l'expérience du jet de dé on peut choisir également $\Omega = \{pair, impair\}$ ou $\Omega = \{\{1, 2, 3, \}, \{4, 5, 6\}\}$.

Exemple 1.5

Si on tire une carte d'un jeu de 32 cartes, on peut retenir comme ensembles fondamentaux $\Omega = \{7, 8, 9, 10, V, D, R, As\}$ ou $\Omega = \{$*trèfle, carreau, cœur, pique*$\}$ ou $\Omega = \{rouge, noir\}$.

Cet ensemble Ω peut être fini ou infini, continu ou discret.

Exemple 1.6

On lance une pièce jusqu'à obtenir pile, l'événement retenu étant le nombre de jets effectués :

$$\Omega = \{1, 2, \ldots, n, \ldots\} = \mathbb{N}^*$$

ensemble infini dénombrable.

Exemple 1.7

On observe la durée de vie d'une lampe :

$$\Omega = [0, +\infty[= \mathbb{R}_+$$

ensemble infini non dénombrable.

L'ensemble retenu est bien sûr une abstraction et peut parfois contenir des événements qui ne se produiront jamais dans la réalité.

Exemple 1.8

On mesure la taille d'un individu choisi au hasard dans une population et on retient $\Omega = \mathbb{R}_+$; cet ensemble contient de très grandes tailles qui n'existent bien sûr pas dans la réalité, mais en raison de la difficulté de fixer une valeur maximale de la taille pour définir l'ensemble fondamental, c'est le choix qui paraît le moins arbitraire.

1.2 Algèbre et tribu d'événements

Un événement étant un élément de $\mathcal{P}(\Omega)$ obéit à la théorie des ensembles. Nous allons indiquer dans le tableau ci-après comment certaines notions ensemblistes s'expriment, ou se traduisent, en termes d'événements.

Ensemble	Événement
On a observé le résultat ω et $\omega \in A$	L'événement A est réalisé.
$A = B$	Les événements A et B sont identiques.
$A \subset B$	L'événement A implique l'événement B.
$\emptyset$	Événement impossible.
Ω	Événement certain.
$A \cup B$	Un au moins des deux événements est réalisé.
$A \cap B$	Les deux événements A et B sont réalisés.
$A \cap B = \emptyset$	Les événements A et B sont incompatibles.
$\overline{A} = \Omega - A$ ou A^c	L'événement A n'est pas réalisé.

Le couple $(\Omega, \mathcal{P}(\Omega))$ s'appelle un **espace probabilisable**.

Cependant, même si Ω est fini, le cardinal de $\mathcal{P}(\Omega)$ est $2^{card\,\Omega}$, qui peut être un nombre très grand. Dans ce cas, il peut être souhaitable de ne considérer qu'une famille restreinte $\mathcal{A}$ de parties de Ω, $\mathcal{A} \subset \mathcal{P}(\Omega)$. Pour que le résultat des opérations ensemblistes (union, intersection, complémentaire) soit encore un événement, il est nécessaire que cette famille d'événements retenue soit fermée, ou stable, vis-à-vis de ces opérations, c'est-à-dire qu'il soit bien un élément de la famille (par exemple, si on retient la famille des nombres impairs, elle n'est pas stable pour l'addition puisque le résultat est un nombre pair). De plus, les événements « certain », Ω, et « impossible », $\emptyset$, doivent également appartenir à cet ensemble. Ainsi, on associera à une épreuve aléatoire un ensemble non vide de parties de Ω, noté $\mathcal{A}$, qui vérifiera :

– $\boxed{C_1}$ pour tout $A \in \mathcal{A}$ alors $\overline{A} \in \mathcal{A}$;

– $\boxed{C_2}$ pour tout $A \in \mathcal{A}$ et tout $B \in \mathcal{A}$ alors $A \cup B \in \mathcal{A}$.

Il y a fermeture pour le complémentaire et l'union. Cet ensemble $\mathcal{A}$ s'appelle une **algèbre** de parties de Ω. Bien entendu, grâce aux lois de Morgan, on a une définition équivalente en remplaçant la condition C_2 par :

– $\boxed{C_2'}$ pour tout $A \in \mathcal{A}$ et tout $B \in \mathcal{A}$ alors $A \cap B \in \mathcal{A}$.

Exemple 1.9

L'algèbre la plus élémentaire est réduite à $\mathcal{A} = \{\emptyset, \Omega\}$.

Exemple 1.10

À partir d'un événement quelconque A, on peut constituer l'algèbre :

$$\mathcal{A} = \{\emptyset, A, \overline{A}, \Omega\}$$

Exemple 1.11

On peut générer une algèbre à partir d'une partition. À partir de la partition de $\Omega = \{a, b, c, d\}$ en trois ensembles $\{a, b\}, \{c\}, \{d\}$, on construit l'algèbre :

$$\mathcal{A} = \{\emptyset, \{a, b\}, \{c\}, \{d\}, \{a, b, c, \}, \{a, b, d\}, \{c, d\}, \Omega\}$$

avec card $\mathcal{A} = 2^3$.

Exemple 1.12

L'algèbre la plus complète est bien entendu $\mathcal{P}(\Omega)$.

Théorème 1.1

Propriétés d'une algèbre

$\boxed{P_1}$ La famille étant non vide, on en conclut que :

$$\emptyset \in \mathcal{A}, \Omega \in \mathcal{A}$$

$\boxed{P_2}$ Si $A_j \in \mathcal{A}$ pour $1 \leqslant j \leqslant n$, on démontre par récurrence que :

$$\bigcup_{j=1}^{n} A_j \in \mathcal{A}$$

$\boxed{P_3}$ Si $A_j \in \mathcal{A}$ pour $1 \leqslant j \leqslant n$, on démontre également par passage au complémentaire que :

$$\bigcap_{j=1}^{n} A_j \in \mathcal{A}$$

Cependant, certaines expériences peuvent se dérouler indéfiniment (au moins théoriquement), comme par exemple lancer un dé jusqu'à obtenir le chiffre 6. Si A_n représente l'événement « obtenir le chiffre 6 au n-ème lancer », l'événement « obtenir le chiffre 6 » s'écrira $A = \bigcup_{n=1}^{\infty} A_n$. On a donc besoin de renforcer la propriété P_2 de fermeture pour l'union finie par une condition de fermeture pour l'union dénombrable, soit :

– $\boxed{C_3}$ Si $A_n \in \mathcal{A}$ pour tout $n \in \mathbb{N}$ alors :

$$\bigcup_{n=0}^{\infty} A_n \in \mathcal{A}$$

condition qui exprime que toute union dénombrable d'événements est encore un événement. L'ensemble $\mathcal{A}$ auquel on impose les conditions C_1 et C_3 s'appelle alors une σ – **algèbre** ou **tribu** d'événements.

REMARQUE On aurait pu remplacer dans la condition C_3 l'union par l'intersection (par passage au complémentaire l'union se transforme en intersection). Bien entendu toute algèbre finie est une σ – **algèbre**.

Cependant, dans les cas simples où Ω est un ensemble fini, ou éventuellement infini dénombrable, on peut retenir comme tribu des événements $\mathcal{P}(\Omega)$ tout entier. Ce n'est que dans les autres cas que l'on est amené à considérer des ensembles $\mathcal{A}$ plus réduits que $\mathcal{P}(\Omega)$ qui est alors trop vaste. Le couple formé de l'ensemble fondamental Ω et de la tribu d'événements associée $\mathcal{A}$ s'appelle un **espace probabilisable**. Cette terminologie signifie que l'on va pouvoir associer une **probabilité**, notée P, à ce modèle $(\Omega, \mathcal{A})$.

1.3 Probabilité

Une fois défini l'ensemble des événements auxquels on s'intéresse, on va essayer de traduire par un nombre leurs « possibilités » de réalisation. Cela revient à affecter une mesure de « croyance » à chaque événement, c'est-à-dire un degré de certitude que l'on a que l'événement se produise ou non. Afin de correspondre à la notion intuitive, une probabilité sera un nombre associé à un événement, compris entre 0 et 1, pour pouvoir se convertir en pourcentage de « chances » ; l'événement certain Ω se voit attribuer la probabilité 1 et l'événement impossible $\emptyset$ la probabilité 0. Nous verrons dans l'exemple 6.6 que cette définition axiomatique est adaptée à la réalité, puisque la fréquence observée d'un événement a pour limite sa probabilité ainsi définie. D'autre part, si deux événements sont incompatibles, c'est-à-dire ne peuvent pas se réaliser simultanément, la probabilité de réalisation de l'un des deux sera la somme de leurs probabilités respectives (par exemple pour un jet de dé $P(1 \text{ ou } 2) = P(1) + P(2) = 1/6 + 1/6 = 1/3$). D'où la définition suivante :

DÉFINITION

On appelle probabilité P sur $(\Omega, \mathcal{A})$ une application $P : \mathcal{A} \to [0, 1]$ telle que :
(i) $P(\Omega) = 1$;

(ii) pour toute suite A_n d'événements incompatibles, soit $A_n \in \mathcal{A}$ avec $A_m \bigcap A_n = \emptyset$ pour $m \neq n$:

$$P\left(\bigcup_{n=0}^{\infty} A_n\right) = \sum_{n=0}^{\infty} P(A_n)$$

propriété dite de σ – additivité.

REMARQUE Pour toute suite d'événements quelconques, c'est-à-dire non disjoints, on a l'inégalité de Boole :

$$P\left(\bigcup_{n=0}^{\infty} A_n\right) \leqslant \sum_{n=0}^{\infty} P(A_n)$$

Une probabilité est donc une application qui à un événement va associer un nombre. Le triplet $(\Omega, \mathcal{A}, P)$ s'appelle un **espace probabilisé**. Comme conséquences de la définition on déduit les propriétés suivantes.

Théorème 1.2

$\boxed{P_1}$ L'événement impossible est de probabilité nulle :

$$P(\emptyset) = 0$$

$\boxed{P_2}$ La probabilité de l'événement complémentaire d'un événement quelconque A s'obtient par :

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

$\boxed{P_3}$ Si un événement en implique un autre, sa probabilité est plus petite :

$$A \subset B \Rightarrow P(A) \leqslant P(B)$$

$\boxed{P_4}$ La probabilité de l'union de deux événements s'obtient par la formule de Poincaré :

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

a) Cas où Ω est fini

Si $\Omega = \{\omega_1, \ldots, \omega_n\}$, la donnée de n nombres $p_i, 1 \leqslant i \leqslant n$, associés à chacun des événements élémentaires par $p_i = P(\{\omega_i\})$, tels que $0 \leqslant p_i \leqslant 1$ et $\sum_{i=1}^{n} p_i = 1$, suffit à déterminer une probabilité P sur $(\Omega, \mathcal{P}(\Omega))$ par :

$$P(A) = \sum \{p_i / \omega_i \in A\}$$

c'est-à-dire que la probabilité d'un événement quelconque A de $\mathcal{P}(\Omega)$ est définie comme la somme des probabilités de tous les événements élémentaires qui y sont inclus.

b) Cas particulier : équiprobabilité

Il s'agit d'un cas assez fréquent où tous les événements élémentaires ont la même probabilité, ce qui correspond à la loi uniforme discrète définie par :

$$p_i = \frac{1}{n}, \quad 1 \leqslant i \leqslant n$$

Cette particularité est souvent sous-entendue ou précisée par l'affirmation que les résultats de l'expérience sont obtenus **au hasard**. On obtient alors :

$$P(A) = \sum_{\omega_i \in A} \frac{1}{n} = \frac{1}{n} \times card\ A = \frac{card\ A}{card\ \Omega}$$

puisque $n = card\ \Omega$. Ce résultat s'énonce souvent sous la forme très dangereuse de la règle énoncée par Laplace au XVIII^{e} siècle :

$$P(A) = \frac{\text{nbre de cas favorables}}{\text{nbre de cas possibles}}$$

un cas favorable étant un événement élémentaire qui réalise A. On est alors ramené à un simple problème de dénombrement. Mais il faut bien faire attention que cette règle ne s'applique que dans le cas d'**équiprobabilité** des événements élémentaires. Si en jouant au loto il y a deux événements possibles, gagner ou non le gros lot, il n'y a malheureusement pas une chance sur deux pour l'événement favorable qui est de le gagner ! Cette règle ne peut donc en aucun cas servir de définition pour une probabilité.

ATTENTION !

C'est ici qu'il est important de bien préciser l'ensemble fondamental Ω. Si par exemple on lance deux pièces de monnaie identiques, on pourrait être tenté de retenir comme ensemble fondamental $\Omega = \{PP, PF, FF\}$, mais dans ce cas il n'y aurait pas équiprobabilité des événements élémentaires, car PP ou FF ne peut être obtenu que d'une seule façon alors que PF peut être obtenu de deux façons distinctes, le résultat pile pouvant être réalisé sur une pièce ou l'autre. Ces événements ont donc respectivement comme probabilité 1/4, 1/4 et 1/2. Il vaut donc mieux faire comme si les pièces étaient distinctes et retenir $\Omega = \{PP, PF, FP, FF\}$ où les événements élémentaires sont équiprobables.

REMARQUES

1. Dans le cas où Ω est fini, on peut toujours construire un espace probabilisé tel que chaque événement élémentaire, au sens de singleton, ait une probabilité nulle. Par exemple pour $\Omega = \{a, b, c\}$ on choisit $\mathcal{A} = \{\emptyset, \{a\}, \{b, c\}, \Omega\}$ et la probabilité P définie par $P(\{a\}) = 0, P(\{b, c\}) = 1$, ce qui correspond au cas de deux événements indiscernables, avec un seul événement élémentaire.

2. De même, on peut construire un espace probabilisé tel que chaque événement élémentaire, toujours au sens de singleton, ait une probabilité égale à 1. Dans l'ensemble précédent, la probabilité P est alors définie par $P(\{a\}) = 1$ et $P(\{b, c\}) = 0$.

3. Même dans le cas où les événements A et B ne sont pas disjoints, on peut avoir l'égalité $P(A \cup B) = P(A) + P(B)$; il suffit pour cela que $P(A \cap B) = 0$. Par exemple, définissons sur $\Omega = \{a, b, c\}$ la probabilité P par $P(\{a\}) = P(\{b\}) = 1/2$ et $P(\{c\}) = 0$. Avec $A = \{a, c\}$ et $B = \{b, c\}$ on obtient $P(A) = P(B) = 1/2$ et $P(A \cup B) = P(\Omega) = 1 = 1/2 + 1/2 = P(A) + P(B)$ et pourtant $A \cap B = \{c\} \neq \emptyset$.

Terminons ce paragraphe par deux exemples, paradoxaux à première vue.

Exemple 1.13

Dans un tournoi exhibition de tennis, Djokovic doit affronter Nadal et Federer au cours de trois sets successifs où ses deux adversaires alterneront. Il remportera ce tournoi s'il gagne deux sets consécutifs. Il considère que la probabilité p de battre Federer est supérieure à celle q de battre Nadal : $p > q$. Quel adversaire va-t-il choisir d'affronter en premier ?

Pour une succession Fed-Nadal-Fed, sa probabilité de gain est :

$$p_1 = pq + (1 - p)qp.$$

Pour une succession Nadal-Fed-Nadal, cette probabilité est :

$$p_2 = qp + (1 - q)pq = p_1 + pq(p - q) > p_1.$$

Il doit donc choisir d'affronter le joueur le plus fort en premier, ce paradoxe s'expliquant par le fait qu'il a plus de chances de le battre une fois sur deux parties que sur une seule.

Exemple 1.14

Déterminons la répartition la plus probable de six atouts entre deux partenaires de bridge. Il y a $\binom{26}{13}$ répartitions possibles des 26 cartes des deux joueurs. Les dénombrements des répartitions de six atouts, et donc des vingt non-atouts, entre les deux joueurs sont indiquées dans le tableau ciaprès :

(1, 5)	(2, 4)	(3, 3)	(0, 6)
$2\binom{6}{1}\binom{20}{12}$	$2\binom{6}{2}\binom{20}{11}$	$2\binom{6}{3}\binom{20}{11}$	$2\binom{20}{13}$

ce qui correspond aux probabilités :

$$P\{(1,5)\} = \frac{117}{805}, P\{(2,4)\} = \frac{390}{805}, P\{(3,3)\} = \frac{286}{805}, P\{(0,6)\} = \frac{12}{805}$$

C'est donc la répartition (2, 4) qui est la plus probable.

On demande maintenant de répondre à la même question dans le cas où les deux partenaires ont fourni aux deux premiers tours d'atout. Les seules répartitions initiales possibles sont alors (2, 4) et (3, 3), donnant respectivement les répartitions (0, 2) et (1, 1) après deux tours d'atout. Il reste $\binom{22}{11}$ répartitions équiprobables entre les deux joueurs, les répartitions restantes des deux atouts, et donc des vingt non-atouts, ayant comme probabilités :

$$P\{(0,2)\} = 2\frac{\binom{20}{11}}{\binom{22}{11}} = \frac{10}{21}, \qquad P\{(1,1)\} = \frac{\binom{2}{1}\binom{20}{10}}{\binom{22}{11}} = \frac{11}{21}$$

La répartition la plus probable *a posteriori*, c'est-à-dire après les deux tours d'atouts, est donc issue de la moins probable des deux *a priori*.

2 Probabilités conditionnelles

On considère l'espace probabilisé $(\Omega, \mathcal{A}, P)$ et un événement particulier B de $\mathcal{A}$ tel que $P(B) > 0$. La connaissance de la réalisation de B modifie la probabilité de réalisation d'un événement élémentaire, puisque l'ensemble des résultats possibles est devenu B et non plus Ω.

Exemple 1.15

À l'issue d'un jet de dé, on sait que le résultat est supérieur à trois et on s'intéresse à l'événement A = « obtenir une face paire ». Initialement on avait $P(A) = 3/6 = 1/2$; maintenant Ω est devenu $\Omega|B = \{4, 5, 6\}$ et $P(A|B) = 2/3 > 1/2$.

Cette nouvelle probabilité, notée $P(.|B)$, est définie sur la tribu conditionnelle :

$$\mathcal{A}|B = \{A \cap B / A \in \mathcal{A}\}$$

par :

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

Seuls les événements ayant une partie commune avec B peuvent se réaliser et la figure 1.1 visualise cette situation où l'ensemble fondamental est devenu B et donc seule la part de A incluse dans B est prise en compte dans le calcul de la probabilité conditionnelle.

Vérifions que cette application π de $\mathcal{A}$ dans $\mathbb{R}$, définie par $\pi(A) = P(A \cap B)/P(B)$ est bien une probabilité. On a bien entendu $\pi(A) \geqslant 0$ et comme $A \cap B \subset B$ on a également $\pi(A) \leqslant 1$. D'autre part :

$$\pi(\Omega) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

donc π vérifie la première condition de la définition.

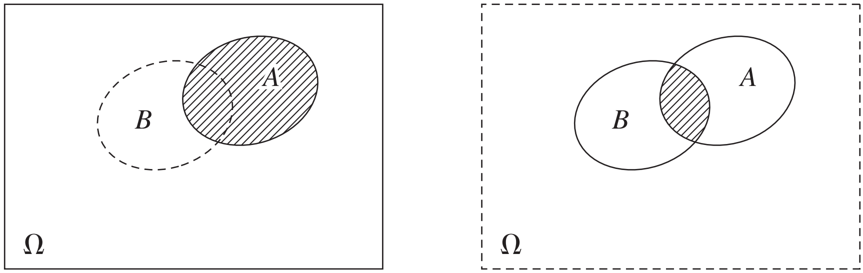

Figure 1.1

Enfin, si $A_n \in \mathcal{A}$ avec $A_m \cap A_n = \emptyset$ pour $m \neq n$, alors :

$$\pi\left(\bigcup_{n=0}^{\infty} A_n\right) = \frac{1}{P(B)} P\left\{\bigcup_{n=0}^{\infty} (A_n \cap B)\right\} = \sum_{n=0}^{\infty} \frac{P(A_n \cap B)}{P(B)} = \sum_{n=0}^{\infty} \pi(A_n)$$

donc la condition 2 est aussi vérifiée.

Exemple 1.16

On tire sans remise deux cartes successivement d'un jeu de 52 et on cherche la probabilité de tirer un as au deuxième coup sachant que l'on en a obtenu un au premier. Avec des notations évidentes :

$$P(A_2 | A_1) = \frac{P(A_1 A_2)}{P(A_1)} = \frac{(4 \times 3)/(52 \times 51)}{4/52} = \frac{3}{51} = \frac{1}{17}$$

alors que la probabilité non conditionnelle est :

$$P(A_2) = P(A_1 A_2) + P(\overline{A}_1 A_2) = \frac{4 \times 3}{52 \times 51} + \frac{48 \times 4}{52 \times 51} = \frac{4}{52} = P(A_1) = \frac{1}{13}$$

donc valeur plus élevée (avoir tiré un as au premier coup diminue la probabilité d'en tirer un au second).

Exemple 1.17

On lance trois fois une pièce de monnaie et on considère les événements A = « obtenir au moins deux face » et B = « obtenir face au premier coup ». L'ensemble fondamental retenu est $\Omega = \{P, F\}^3$, ensemble des triplets ordonnés, bien que l'ordre des résultats n'intervienne pas, mais pour qu'il y ait équiprobabilité des événements élémentaires, soit $P(\{\omega\}) = 1/8$ puisque card $\Omega = 2^3 = 8$. Ces événements s'écrivent $A = \{FFP, FPF, PFF, FFF\}$ et $B = \{FFF, FFP, FPF, FPP\}$, avec card A = card $B = 4$ donc $P(A) = P(B) = 4/8 = 1/2$. Calculons maintenant la probabilité conditionnelle $P(A|B)$. On a $A \cap B = \{FFP, FPF, FFF\}$ donc card $A \cap B = 3$ et $P(A \cap B) = 3/8$. Ainsi :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3}{4} > \frac{1}{2}$$

la probabilité conditionnelle a ici augmenté.