THEORIE DES PROBABILITES

INTRODUCTION

La statistique a une origine très ancienne, se réduisant initialement à une collecte d'observations, notamment le dénombrement des hommes (recensement). Cependant le terme statistique est apparu assez récemment, vers le milieu du 17ème siècle ; il vient du latin statisticus, relatif à l'état (status) et est employé 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 pour étendre les résultats et dégager des lois (l'inférence). Il s'agit alors dans ce contexte 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.

Le début de la méthodologie statistique peut se situer au XVIIe 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 azzahr (jeu de dés, alea en latin) au XIIe siècle d'où est venue cette expression jeu de hasard au XVIe siècle. 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. 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 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 le cadre de ce cours composé de 6 chapitres.

- Le chapitre 1 intitulé « Notion de probabilité » aura pour objectif de montrer que le modèle probabiliste est choisi en fonction du but que l'on poursuit, qui se résume essentiellement à la construction du modèle d'échantillonnage, base de la modélisation statistique.

- Le chapitre 2 intitulé « Variable aléatoire » aura pour objectif de retenir que dans les cas usuels la loi d'une variable aléatoire est déterminée par les probabilités des points dans le cas discret, par la densité dans le cas continu, et que cette loi est souvent caractérisée par les deux premiers moments qui sont l'espérance, caractéristique de valeur centrale, et la variance, caractéristique de dispersion autour de cette valeur centrale.
- Le chapitre 3 intitulé « Lois usuelles » aura pour objectif de présenter les principales lois de probabilité pouvant être retenues dans la modélisation statistique.
- Le chapitre 4 intitulé « Couple et vecteur aléatoires » aura pour objectif de généraliser les notions de variable aléatoire, d'espérance et de variance au cas multidimensionnel ; définir les lois conditionnelles, la fonction de régression et la convolution de deux lois.
- Le chapitre 5 intitulé « Loi empirique » aura pour objectif d'introduire la notion d'échantillon, définir les lois de Student et de Fisher-Snedecor associées à un échantillon gaussien et présenter deux tests d'adéquation à une loi donnée.
- Le chapitre 6 intitulé « Comportement asymptotique » aura pour objectif de présenter les deux théorèmes fondamentaux de la statistique asymptotique, la loi des grands nombres et le théorème central limite, associés aux deux modes principaux de convergence, la convergence en loi et la convergence en probabilité.

Chapitre 1 NOTION DE PROBABILITE

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

1. MODELE 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* $\Omega = \{noire, blanche, rouge\}$.

Chaque élément $\omega \in \Omega$ représente donc un événement élémentaire, et toute partie $A \subset \Omega$ (ou $A \in P(\Omega)$) 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 $P(\Omega)$, 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 = {*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 Ω = {*trèfle, carreau, cœur, pique*} *ou* Ω = {*rouge*, *noir*}.

Cet ensemble 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 : $\Omega = [0, +\infty[= R+$ 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 $\Omega = R+$; cet ensemble contient des 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. ALGEBRE ET TRIBU D'EVENEMENTS

Un événement étant un élément de $P(\Omega)$ 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 $(\Omega, P(\Omega))$ s'appelle un espace probabilisable. Cependant, même si Ω est fini, le cardinal de $P(\Omega)$ 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 A de parties de Ω, $A \subset P(\Omega)$. 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 A, qui vérifiera :

– C1 pour tout $A \in A$ alors $\bar{A} \in A$;

– C2 pour tout $A \in A$ et tout $B \in A$ alors $A \cup B \in A$.

Il y a fermeture pour le complémentaire et l'union. Cet ensemble 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 C2 par :

– C'2 pour tout $A \in A$ et tout $B \in A$ alors $A \cap B \in A$.

Exemple 1.9

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

Exemple 1.10

À partir d'un événement quelconque A, on peut constituer l'algèbre : A = {Ø, A, $\bar{A}$, Ω}

Exemple 1.11

On peut générer une algèbre à partir d'une partition. À partir de la partition

de Ω = {a, b, c, d} en trois ensembles {a, b} , {c} , {d} on construit l'algèbre :A = {Ø, {a, b} , {c} , {d} , {a, b, c,} , {a, b, d} , {c, d}, Ω} avec card A = 2^3.

Exemple 1.12

L'algèbre la plus complète est bien entendu P(Ω)

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 $A = \cup_{n=1}^{\infty} A_n$. 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 :

– *C*3 : si $A_n \in$ A pour tout $n \in$ 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.

NB : 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 P(Ω) tout entier. Ce n'est que dans les autres cas que l'on est amené à considérer des ensembles $\mathcal{A}$ plus réduits que P(Ω) 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 (Ω, $\mathcal{A}$).

1.3. PROBABILITE

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 Ø 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 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 $A_m \cap A_n = \emptyset$ 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 :

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

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

Propriétés

$\boxed{P_1}$ 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 :

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

.3.1. Cas où Ω est fini.

Si Ω = {ω1, . . ., ω*n*}, la donnée de *n* nombres *pi*, 1 ≤ *i* ≤ *n*, associés à chacun des événements élémentaires par *pi* = *P* ({ω*i* }), tels que 0 ≤ *pi* ≤ 1 et $\sum_1^n p_i = 1$ suffit à déterminer une probabilité *P* sur (**Ω**, P (**Ω**)) par :

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

C'est-à-dire que la probabilité d'un événement quelconque *A* de P (**Ω**) est définie comme la somme des probabilités de tous les événements élémentaires qui y sont inclus.

.3.2. **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 Ω*. Ce résultat s'énonce souvent sous la forme très dangereuse de la règle énoncée par Laplace au XVIIIe siècle :

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

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

a. 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.

b. 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 $P(\{b, c\}) = 0$.

c. 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 $P(A \cup B) = P(\Omega) = 1 = 1/2 + ½ = P(A) + P(B)$ et pourtant $A \cap B = \{c\} \neq \emptyset$.

PROBABILITES 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.

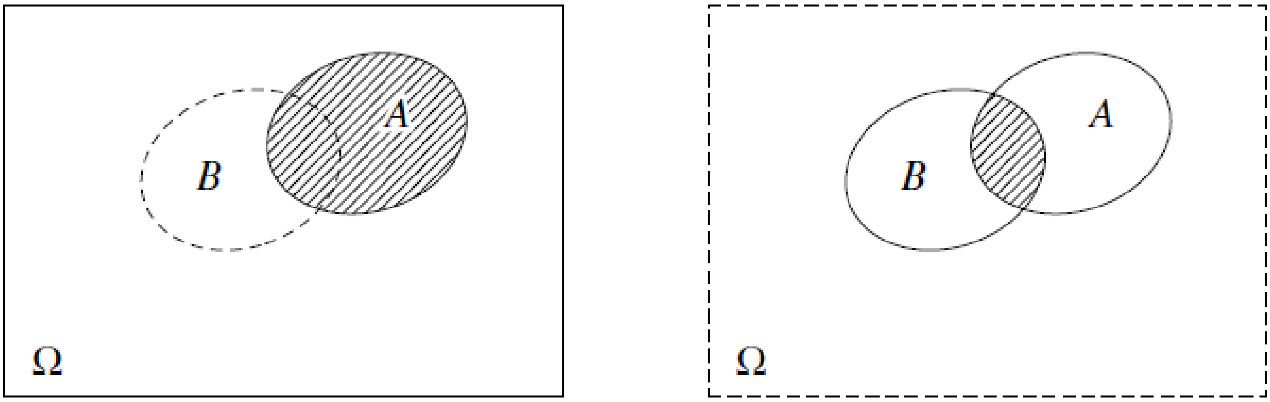

Figure 1.1

Vérifions que cette application π de A dans R, définie par π(A) = P(A ∩ B)/P(B) est bien une probabilité. On a bien entendu π(A) ≥ 0 et comme A ∩ B ⊂ B on a également π(A) ≤ 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.

Enfin, si $A_n \in A$ avec $A_m \cap A_n = \emptyset$ pour m ≠ 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 Ω = $\{P, F\}^3$, 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({ω}) = 1/8 puisque card _ = 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 ∩ B = {FFP, FPF, FFF} donc card (A ∩ B) = 3 et P(A ∩ 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.

La formule de définition de la probabilité conditionnelle peut aussi s'écrire, si P(A) > 0 :

$$P(A \cap B) = P(B)P(A|B) = P(A)P(B|A)$$

et s'appelle parfois formule des probabilités composées ; elle se généralise par récurrence :

$$\begin{aligned} P(A_1 \cap A_2 \cap \ldots \cap A_n) = & P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \\ & \ldots P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \\ = & P(A_1)\prod_{k=2}^{n} P\left(A_k|\bigcap_{i=1}^{k-1} A_i\right) \end{aligned}$$

Exemple 1.18

Dans une urne qui contient deux boules rouges et trois noires, quatre personnes tirent successivement une boule sans la remettre ; la première qui tire une rouge gagne. Calculons la probabilité de gain de chaque personne A, B, C et D :

$$P\,(A) = P\,(R_1) = \frac{2}{5}$$
$$P\,(B) = P(N_1)P(R_2|N_1) = \frac{3}{5} \times \frac{2}{4} = \frac{3}{10}$$
$$P\,(C) = P(N_1)P(N_2|N_1)P(R_3|N_1N_2) = \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} = \frac{1}{5}$$
$$P\,(D) = P(N_1)P(N_2|N_1)P(N_3|N_1N_2)P(R_4|N_1N_2N_3) = \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} \times \frac{2}{2} = \frac{1}{10}$$

THEOREME DE BAYES

Nous allons voir à partir d'un exemple dans quelles conditions on peut être amené à utiliser la formule, ou théorème, de Bayes.

Exemple 1.19

Considérons une expérience aléatoire qui se déroule en deux étapes : on tire au sort entre deux urnes U1 et U2, avec des probabilités respectives de 1/5 et 4/5 puis on tire une boule dans l'urne choisie. Leurs compositions *respectives sont 6 blanches, 4 noires et 3 blanches, 7 noires. La probabilité a priori de U1 est donc 1/5. Sa probabilité a posteriori, sachant qu'on a obtenu une boule blanche, va être plus élevée car la probabilité de tirer une blanche dans U1 est plus forte que dans U2.*

On a :

$$P(U_1|B) = \frac{P(U_1 \cap B)}{P(B)}$$

avec :

$$\begin{aligned} P(U_1 \cap B) &= P(U_1)P(B|U_1) = \frac{1}{5} \times \frac{6}{10} = \frac{3}{25} \\ P(B) &= P(B \cap U_1) + P(B \cap U_2) \\ &= P(U_1)P(B|U_1) + P(U_2)P(B|U_2) \\ &= \frac{1}{5} \times \frac{6}{10} + \frac{4}{5} \times \frac{3}{10} = \frac{9}{25} \end{aligned}$$

ainsi :

$$P(U_1|B) = \frac{3/25}{9/25} = \frac{1}{3} > P(U_1) = \frac{1}{5}$$

Cet exemple peut se traiter par application de la formule de Bayes que nous allons établir en considérant un système complet d'événements, c'est-à-dire une partition de Ω en événements {A1, . . . , An} de probabilités strictement positives,

P(Ai) > 0 pour 1 ≤ i ≤ n, et incompatibles deux à deux, i.e. avec Ai ∩ Aj = ∅ pour i ≠ j et $\sum_{i=1}^{n} P(Ai) = 1$. *On suppose que les probabilités des événements inclus dans chacun des Ai sont connues et on va donc décomposer un événement quelconque B sur ce système :*

$$B = B \cap \Omega = B \cap \left(\bigcup_{i=1}^{n} A_i \right) = \bigcup_{i=1}^{n} (A_i \cap B)$$

On aboutit ainsi à la formule de la probabilité totale :

$$P(B) = \sum_{i=1}^{n} P(A_i \cap B) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$$

Ceci va nous permettre de calculer les probabilités a posteriori P(Ai |B), après réalisation d'un événement B, à partir des probabilités a priori P(Ai), 1 ≤ i ≤ n :

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)}$$

Résultat appelé formule de Bayes ou parfois théorème de Bayes.

Exemple 1.20

On tire au sort entre trois urnes dont les compositions sont indiquées dans le tableau ci-après. Sachant qu'on a obtenu une boule rouge, on se pose la question de savoir quelle est la probabilité qu'elle provienne de l'urne U2 .

	Rouge	*Bleue*	*Verte*
U_1	3	4	1
U_2	1	2	3
U_3	4	3	2

$$P(R)=\sum_{i=1}^{3}P(R\cap U_i)=\sum_{i=1}^{3}P(U_i)P(R|U_i)=\frac{1}{3}\left(\frac{3}{8}+\frac{1}{6}+\frac{4}{9}\right)$$

La probabilité a posteriori de l'urne U2 étant donc :

$$P(U_2|R)=\frac{P(U_2)P(R|U_2)}{P(R)}=\frac{1/6}{3/8+1/6+4/9}$$
$$=\frac{12}{71}=\frac{36}{213}<\frac{71}{213}=P(U_2)$$

INDEPENDANCE EN PROBABILITE

Définition

Deux événements A et B sont dits indépendants, relativement à la probabilité P, si :

$$P(A \cap B) = P(A)P(B)$$

La probabilité de réalisation simultanée de deux événements indépendants est égale au produit des probabilités que chacun de ces événements se produise séparément. En conséquence, si P (B) > 0 :

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

la réalisation d'un événement ne modifie pas la probabilité de réalisation de l'autre.

Exemple 1.21

On jette un dé rouge et un dé vert et on considère les événements A = « le dé vert marque 6 », et B = « le dé rouge marque 5 ». Il nous faut démontrer que ces deux événements sont indépendants (bien entendu ce résultat est évident, il n'y a pas d'influence d'un dé sur l'autre !). L'ensemble fondamental retenu est Ω = E $\times$ E, avec

E = {1, 2, 3, 4, 5, 6}, sur lequel il y a équiprobabilité : P({ω}) = $1/6^2$. Comme A = {6} × E, B = E $\times$ {5} et A $\cap$ B = {(6, 5)} on obtient :

$$P\ (A)=\frac{card\ A}{card\ \Omega}=\frac{6}{6^2}=\frac{1}{6}$$
$$P\ (B)=\frac{card\ B}{card\ \Omega}=\frac{6}{6^2}=\frac{1}{6}$$
$$P\ (A\cap B)=\frac{card\ \ (A\cap B)}{card\ \Omega}=\frac{1}{6^2}=P(A)P(B)$$

et les événements A et B sont donc bien indépendants.

• Indépendance mutuelle

Si l'on considère n événements Ai, avec n > 2, il y a lieu de distinguer l'indépendance deux à deux qui impose :

$$P(A_i \cap A_j) = P(A_i)P(A_j), \quad 1 \leqslant i \neq j \leqslant n$$

de l'indépendance mutuelle, condition plus forte qui s'écrit :

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \ldots P(A_{i_k}), \quad 2 \leqslant k \leqslant n,$$

pour tous les sous-ensembles {i1, . . . , ik} de {1, 2, . . . , n}, ce qui correspond à :

$$\sum_{k=2}^{n} \binom{n}{k} = 2^n - \binom{n}{0} - \binom{n}{1} = 2^n - 1 - n$$

conditions qui doivent être réalisées.

_ ***Exemple 1.22***

On lance deux pièces de monnaie distinctes et on s'intéresse aux événements A = « obtenir pile sur la pièce 1 », B = « obtenir pile sur la pièce 2 » et C = « obtenir pile sur une seule pièce ». À chaque pièce est attaché le même ensemble fondamental E = {P, F} et à l'expérience l'ensemble Ω = E × E. On peut écrire A = {P} × E = {PP, PF} donc P(A) = 2/4,

B = E × {P} = {PP, FP} donc P(B) = 2/4 et C = {PF, FP} donc P(C) = 2/4.

On considère maintenant les événements deux à deux : A ∩ B = {PP} donc P(A ∩ B) = ¼ = P(A)P(B), A ∩ = C {PF} donc P(A ∩ C) = 1/4 = P(A)P(C) et B ∩ C = {FP} donc P(B ∩ C) = ¼ = P(B)P(C) , ces événements sont indépendants deux à deux. Mais A ∩ B ∩ C = Ø donc P(A ∩ B ∩ C) = 0 ≠ P(A)P(B)P(C) et ces événements ne sont pas indépendants dans leur ensemble.