1 Le modèle probabiliste

En tant que théorie mathématique, la théorie des probabilités n'a pas à être justifiée : une fois ses axiomes posés, elle se développe de façon autonome par rapport à la réalité concrète.

Il en va autrement lorsque l'on cherche à appliquer le calcul des probabilités : on ne peut alors éluder la question de la nature de la probabilité et de la validité du modèle probabiliste. Après trois paragraphes consacrés à un exposé simple[1] de la théorie on se proposera de donner quelques éléments de réflexion sur le concept de probabilité.

1.1 ESPACE PROBABILISABLE

On expose ici la formalisation d'une expérience où intervient le « hasard ».

1.1.1 Expérience aléatoire et événements

Une ***expérience*** est qualifiée d'***aléatoire*** si l'on ne peut prévoir par avance son résultat et si, répétée dans des conditions identiques, elle peut (on aurait pu s'il s'agit d'une expérience par nature unique) donner lieu à des résultats différents.

On représente le résultat de cette expérience comme un élément ω de l'ensemble Ω de tous les résultats possibles : Ω est appelé ***l'ensemble fondamental*** ou encore l'univers des possibles.

Ainsi à l'expérience aléatoire qui consiste à lancer deux dés, on peut associer l'ensemble $\Omega = \{(1.1), (1.2), (1.3)\ldots\}$ à 36 éléments.

Il convient de noter ici que l'ensemble Ω ne se déduit pas de manière unique de l'expérience mais dépend de l'usage qui doit être fait des résultats : ainsi, si l'on convient une fois pour toutes qu'on ne retiendra de l'expérience des deux dés que la somme des points affichés, on peut très bien se contenter d'un ensemble $\Omega' = \{2, 3, 4 \ldots 12\}$.

---

1 Un exposé complet des fondements théoriques, comprenant en particulier le théorème de prolongement, dépasserait le cadre de ce livre. On se reportera à l'ouvrage de J. Neveu (1964).

Un ***événement*** est une assertion ou proposition logique relative au résultat de l'expérience (ex. : la somme des points est supérieure à 10). On dira qu'un événement est réalisé ou non suivant que la proposition est vraie ou fausse une fois l'expérience accomplie.

A la réalisation d'un événement on peut donc associer tous les résultats de l'épreuve correspondante ; ainsi la somme supérieure ou égale à 10 est l'ensemble de résultats suivants :

$$\{(4.6)\ ;\ (5.6)\ ;\ (6.6)\ ;\ (6.4)\ ;\ (6.5)\}$$

c'est-à-dire une partie de Ω. Désormais nous identifierons un événement à la partie de Ω pour laquelle cet événement est réalisé.

On appelle ***événement élémentaire*** une partie de Ω réduite à un seul élément.

1.1.2 Algèbre des événements

Réciproquement toute partie de Ω peut-elle être considérée comme un événement, ou du moins est-il utile qu'il en soit ainsi ? Afin de répondre à cette question nous allons supposer pour l'instant que l'ensemble des événements constitue une classe $\mathscr{C}$ de parties de Ω dont nous allons définir les propriétés en nous référant à des besoins usuels ; nous en profiterons pour introduire le vocabulaire probabiliste.

A tout événement A, on associe son contraire noté $\overline{A}$ tel que si A est réalisé alors $\overline{A}$ ne l'est pas, et réciproquement. $\overline{A}$ est donc représenté dans Ω par la partie complémentaire de A.

Il sera donc naturel d'exiger de $\mathscr{C}$ la propriété suivante : si $A \in \mathscr{C}$ alors $\overline{A} \in \mathscr{C}$.

Étant donné deux événements A, B on est conduit à s'intéresser à leur union A ou B ($A \cup B$) et à leur intersection (A et B ou $A \cap B$). Il faudra donc que si $A, B \in \mathscr{C}$, $A \cup B$ et $A \cap B \in \mathscr{C}$, et ceci d'une manière générale pour un nombre quelconque d'événements.

On définit également l'événement certain représenté par Ω tout entier et l'événement logiquement impossible (tel que avoir une somme de points égale à 13) représenté par l'ensemble vide $\varnothing$.

Nous pouvons maintenant définir la classe $\mathscr{C}$ par les trois axiomes :

- $\forall A \in \mathscr{C}, \overline{A} \in \mathscr{C}$;
- pour tout ensemble fini ou dénombrable $A_1, A_2, \ldots, A_n$ d'éléments de $\mathscr{C}$, $\bigcup_i A_i \in \mathscr{C}$;
- $\Omega \in \mathscr{C}$.

On peut montrer à titre d'exercice que ces axiomes impliquent que $\varnothing \in \mathscr{C}$ et que $\bigcap_i A_i \in \mathscr{C}$.

Les propriétés précédentes définissent ce que l'on appelle une σ-algèbre de Boole ou une tribu. $\mathscr{P}(\Omega)$ est une σ-algèbre particulière, la plus grosse, mais il n'est pas toujours utile ni souhaitable de l'utiliser.

On peut donc donner maintenant la définition d'un espace probabilisable :

DÉFINITION

On appelle espace probabilisable le couple $(\Omega\ ;\ \mathscr{C})$ où $\mathscr{C}$ constitue une tribu de parties de Ω.

Donnons encore quelques définitions utiles :

DÉFINITIONS

Événements incompatibles. *Deux événements A et B sont dits incompatibles si la réalisation de l'un exclut celle de l'autre, autrement dit si les parties A et B de Ω sont disjointes $A \cap B = \varnothing$.*

Système complet d'événements. $A_1, A_2, \ldots, A_n$ *forment un système complet d'événements si les parties* $A_1, \ldots, A_n$ *de* Ω *constituent une partition de* Ω :

$$\begin{cases} \forall i \neq j & A_i \cap A_j = \varnothing \\ \cup A_i = \Omega \end{cases}$$

1.2 ESPACE PROBABILISÉ

1.2.1 L'axiomatique de Kolmogorov

A chaque événement on associe un nombre positif compris entre 0 et 1, sa probabilité. Afin d'éviter toute discussion de nature philosophique sur le hasard, la théorie moderne des probabilités repose sur l'axiomatique suivante :

DÉFINITIONS

On appelle ***probabilité sur*** $(\Omega, \mathscr{C})$ (***ou loi de probabilité***) *une application P de* $\mathscr{C}$ *dans* $[0, 1]$ *telle que :*

- $P(\Omega) = 1$;
- *pour tout ensemble dénombrable d'événements incompatibles* $A_1, A_2, \ldots, A_n$, *on a* $P(\cup A_i) = \Sigma P(A_i)$.

On appelle ***espace probabilisé*** *le triplet* $(\Omega, \mathscr{C}, P)$.

Une loi de probabilité n'est donc rien d'autre qu'une mesure positive de masse totale 1 et la théorie des probabilités s'inscrit dans le cadre de la théorie de la mesure.

1.2.2 Propriétés élémentaires

Des axiomes on déduit immédiatement les propriétés suivantes :

Propriété 1 : $P(\varnothing) = 0$.
Propriété 2 : $P(\overline{A}) = 1 - P(A)$.
Propriété 3 : $P(A) \leq P(B) \quad$ *si* $A \subset B$.
Propriété 4 : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
Propriété 5 : $P(\cup A_i) \leq \sum_i P(A_i)$.
Propriété 6 : *Si* $A_i \downarrow \varnothing$, *alors* $\lim P(A_i) = 0$ (*continuité monotone séquentielle*).
Propriété 7 : *Théorème des probabilités totales : Soit* B_i *un système complet d'événements alors* $\forall A : P(A) = \sum_i P(A \cap B_i)$.

FORMULE DE POINCARÉ

Cette formule permet de calculer la probabilité de la réunion d'un nombre quelconque d'évenements ; elle se démontre par récurrence :

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{1\le i<j\le 1}^{n} P(A_i \cap A_j)$$
$$+ \sum_{1\le i<j<k\le 1}^{n} P(A_i \cap A_j \cap A_k) + \cdots + (-1)^{n-1} P(A_1 \cap A_2 \ldots \cap A_n)$$

Remarque : $P(A) = 0$ n'implique pas nécessairement $A = \varnothing$. Un événement de probabilité nulle n'est pas nécessairement impossible : soit $\Omega = [0, 1]$ muni de la loi de probabilité uniforme (c'est-à-dire de la mesure de Lebesgue) alors $P(\omega) = 0 \quad \forall \omega$.

De même $P(A) = 1$ n'implique pas que A soit l'événement certain : on parlera d'événement presque certain et dans le cas précédent d'événement presque impossible.

Les événements de probabilité nulle sont en réalité très communs, comme on le verra dans l'étude des variables aléatoires continues possédant une densité : tous les évènements $\{X = x\}$ sont de probabilité nulle mais aucun n'est impossible. La variable X prend une valeur précise une fois l'expérience réalisée. Cela est comparable au fait qu'un intervalle de longueur donnée l est formé d'une infinité de points de longueur nulle.

1.3 LOIS DE PROBABILITÉS CONDITIONNELLES, INDÉPENDANCE

Les concepts suivants sont purement probabilistes.

1.3.1 Introduction et définitions

Supposons que l'on s'intéresse à la réalisation d'un événement A, tout en sachant qu'un événement B est réalisé (fig. 1.1). Si A et B sont incompatibles la question est tranchée : A ne se réalisera pas, mais si $A \cap B \neq \varnothing$, il est possible que A se réalise ; cependant, l'univers des possibles n'est plus Ω tout entier, mais est restreint à B ; en fait, seule nous intéresse la réalisation de A à l'intérieur de B, c'est-à-dire $A \cap B$ par rapport à B.

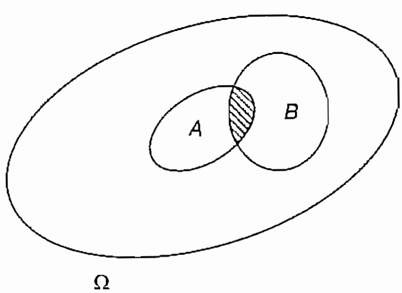

FIGURE 1.1

Ceci justifie la définition suivante :

DÉFINITION

Soit B un événement de probabilité non nulle. On appelle ***probabilité conditionnelle de A sachant B*** *(ou encore de A si B) le rapport noté P(A/B) :*

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

Il faut s'assurer que le nom de probabilité est justifié. Vérifions les axiomes :

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

$$P\left(\bigcup_i A_i/B\right) = \frac{P\left[\left(\bigcup_i A_i\right) \cap B\right]}{P(B)} = \frac{P\left[\bigcup_i (A_i \cap B)\right]}{P(B)}$$

$$\sum_i \frac{P(A_i \cap B)}{P(B)} = \sum_i P(A_i/B) \qquad \text{c.q.f.d}$$

On peut donc munir $(\Omega, \mathscr{C})$ d'une nouvelle loi de probabilité, la loi de probabilité conditionnelle à B fixé et ceci pour tout B de probabilité non-nulle.

Il sera nécessaire d'étendre la notion de loi de probabilité conditionnelle lorsque B est de probabilité nulle (rappelons que la tribu $\mathscr{C}$ contient de tels événements) : cela sera fait au chapitre 3 dans certains cas particuliers.

Exemple : En fiabilité (ou en assurance sur la vie), on considère la fonction de survie $R(t)$ définie comme la probabilité qu'un individu vive au-delà d'une date t : $R(t) = P(X > t)$. Cette fonction définit une loi de probabilité sur $\mathbb{R}^+$ et :

$$P(t_1 \leq X < t_2) = R(t_1) - R(t_2)$$

La probabilité conditionnelle de défaillance (ou de décès) entre t_1 et t_2 sachant que l'individu a déjà fonctionné (ou vécu) jusqu'à t_1 est :

$$P(t_1 \leq X < t_2/X > t_1) = \frac{R(t_1) - R(t_2)}{R(t_1)}$$

Pour la loi de survie exponentielle $P(X > t) = \exp(-ct)$ on constate que cette probabilité conditionnelle vaut :

$$1 - \exp(-c\,(t_2 - t_1)) = P(X < t_2 - t_1)$$

il n'y a pas de vieillissement : la probabilité de fonctionner pendant $t_2 - t_1$ à partir de t_1 est la même qu'au démarrage. Ce modèle est couramment utilisé en électronique.

1.3.2 Indépendance

1.3.2.1 Indépendance de deux événements

Définition

A est indépendant de B si $P(A/B) = P(A)$.

Autrement dit, la connaissance de B ne change pas les « chances » de réalisation de A.

Propriété

A indépendant de $B \Longrightarrow B$ indépendant de A.

On parlera désormais d'événements indépendants sans autre précision.

En effet, si $P(A/B) = P(A)$, alors :

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

et :

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

On a démontré au passage l'importante formule :

$$\boxed{P(A \cap B) = P(A)P(B)}$$

si et seulement si A et B sont indépendants.

N.B. : La notion d'indépendance n'est pas une notion purement ensembliste comme l'incompatibilité : deux événements peuvent être indépendants pour une loi de probabilité P et pas pour une autre P'. On s'en convaincra en vérifiant qu'en général si A et B sont indépendants, ils ne le sont plus conditionnellement à un troisième événement C.

1.3.2.2 Indépendance deux à deux et indépendance mutuelle

Soient $A_1, A_2, \ldots, A_n$ des événements ; ils sont dits mutuellement indépendants si pour toute partie I de l'ensemble des indices allant de 1 à n on a :

$$P\left[\bigcap_I A_i\right] = \prod_I P(A_i)$$

Cette condition est beaucoup plus forte que l'indépendance deux à deux, qui ne lui est pas équivalente mais en est une simple conséquence.

Remarque : Dans les applications il est assez fréquent que l'on n'ait pas à démontrer l'indépendance de deux événements car celle-ci est une propriété de l'expérience aléatoire. Ainsi lorsqu'on procède à un tirage avec remise de n individus dans une population finie les événements relatifs aux différents tirages sont indépendants entre eux par construction.

1.3.3 Formules de Bayes

Elles ont pour but d'exprimer $P(A/B)$ en fonction de $P(B/A)$.

Première formule de Bayes :

$$\boxed{P(B/A) = \frac{P(A/B)P(B)}{P(A)}}$$

Il suffit d'éliminer $P(A \cap B)$ entre $P(A/B) = \dfrac{P(A \cap B)}{P(B)}$ et $P(B/A) = \dfrac{P(A \cap B)}{P(A)}$.

Soit B_i un système complet d'événements. On peut écrire : $P(A \cap B_i) = P(A/B_i)P(B_i)$.

Le théorème des probabilités totales devient donc :

$$\boxed{P(A) = \sum_i P(A/B_i)P(B_i)}$$

On en déduit alors la ***deuxième formule de Bayes*** :

$$\boxed{P(B_i/A) = \frac{P(A/B_i)P(B_i)}{\sum_k P(A/B_k)P(B_k)}}$$

Exemple : Dans une usine trois machines M_1, M_2, M_3 fabriquent des boulons de même type. M_1 sort en moyenne 0.3 % de boulons défectueux, M_2 0.8 % et M_3 1 %. On mélange 1 000 boulons dans une caisse, 500 provenant de M_1, 350 de M_2 et 150 de M_3. On tire un boulon au hasard dans la caisse ; il est défectueux. Quelle est la probabilité qu'il ait été fabriqué par M_1 (ou M_2 ou M_3) ?

Lorsque l'on tire un boulon au hasard les probabilités dites *a priori* qu'il provienne de M_1, M_2 ou M_3 sont évidemment $P(M_1) = 0.50$, $P(M_2) = 0.35$, $P(M_3) = 0.15$.

Lorsque l'on sait qu'il est défectueux, événement noté D, il faut alors calculer les probabilités conditionnelles :

$$P(M_1/D),\ P(M_2/D),\ P(M_3/D)$$

Comme on connaît $P(D/M_1) = 0.003$, $P(D/M_2) = 0.008$ et $P(D/M_3) = 0.01$ la deuxième formule de Bayes permet d'écrire :

$$\begin{aligned} P(M_1/D) &= \frac{P(D/M_1)P(M_1)}{P(D/M_1)P(M_1) + P(D/M_2)P(M_2) + P(D/M_3)P(M_3)} \\ &= \frac{0.003 \times 0.5}{0.003 \times 0.5 + 0.008 \times 0.35 + 0.01 \times 0.15} \\ &\simeq 0.26 \end{aligned}$$

On trouverait de même $P(M_2/D) \simeq 0.48 \qquad P(M_3/D) \simeq 0.26$.

Ce sont les probabilités *a posteriori*, sachant que le boulon est défectueux. On voit donc que la prise en compte d'une information (le boulon est défectueux) modifie les valeurs des probabilités de M_1, M_2 et M_3.

Le théorème de Bayes, simple conséquence des axiomes et de la définition de la probabilité conditionnelle, tient une place à part dans le calcul des probabilités en raison de son importance pratique considérable et des controverses auxquelles son application a donné lieu : il est à la base de toute une branche de la statistique appelée ***statistique bayésienne.***

Parmi les applications courantes citons : en diagnostic médical la révision des probabilités de telle ou telle affection après obtention des résultats d'examens de laboratoire, en matière financière la détermination du risque de faillite des entreprises après observations de certains ratios.

Le théorème de Bayes est souvent appelée théorème sur la « probabilité des causes » ce qui se conçoit aisément sur l'exemple précédent. Son application générale a donné lieu à de violentes critiques de la part des logiciens pour qui causalité et aléatoire sont antinomiques : il n'y a qu'une cause possible parmi des causes mutuellement exclusives et leur donner des probabilités n'aurait aucun sens.

Certains auteurs interprètent le fait que les formules de Bayes ont été publiées à titre posthume (en 1763) par la crainte du sacrilège : Thomas Bayes était en effet un ecclésiastique et l'application de sa formule à la recherche des causes ultimes d'un événement aurait pu conduire à probabiliser l'existence de Dieu. . .

1.4 RÉFLEXIONS SUR LE CONCEPT DE PROBABILITÉ

La théorie mathématique des probabilités ne dit pas quelle loi de probabilité mettre sur un ensemble Ω parmi toutes les lois possibles (et elles sont nombreuses. . .). Ce problème concerne ceux qui veulent appliquer le calcul des probabilités, et renvoie à la nature « physique », si l'on peut dire, du concept de probabilité qui formalise et quantifie le sentiment d'incertitude vis-à-vis d'un événement.

1.4.1 La conception objectiviste

Pour les tenants de ce point de vue, la probabilité d'un événement peut être déterminée de manière unique.

1.4.1.1 La vision classique

C'est celle qui est héritée des jeux de hasard. Ω est en général fini et des raisons de symétrie conduisent à donner à chaque événement élémentaire la même probabilité : ainsi le lancer d'un dé parfait conduit à un ensemble Ω à 6 éléments équiprobables.

Le calcul des probabilités n'est donc plus qu'une affaire de dénombrement, d'où la célèbre formule :

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

L'analyse combinatoire fournit alors les réponses aux cas classiques.

Cette approche ne s'étend pas aux cas où Ω n'est plus dénombrable (voir plus loin) et repose sur une conception idéalisée de l'expérience aléatoire : les symétries parfaites n'existent pas ; ainsi le dé parfait n'est qu'une vue de l'esprit et ses 6 faces ne sont pas en réalité équiprobables en raison de la non homogénéité de la matière et surtout des gravures des numéros sur les faces.

1.4.1.2 *Un paradoxe célèbre*

Les limites de la vision classique apparaissent, nous semble-t-il, assez bien dans le célèbre paradoxe de Bertrand.

Considérons un triangle équilatéral et son cercle circonscrit. On tire une corde au hasard. Quelle est la probabilité que sa longueur soit supérieure à celle du côté du triangle ?

Reproduisons ici les commentaires de Renyi (1966) :

- **Première solution.** Comme la longueur de la corde est déterminée par la position de son milieu, le choix de la corde peut consister à marquer un point au hasard à l'intérieur du cercle. La probabilité pour que la corde soit plus longue que le côté du triangle équilatéral inscrit est alors évidemment égale à la probabilité pour que le milieu de la corde soit intérieur au cercle inscrit qui est de rayon moitié (*cf.* fig. 1.2).

Si l'on admet que la répartition de ce point est uniforme dans le cercle, on trouve pour la probabilité demandée :

$$\frac{\pi(r/2)^2}{\pi r^2} = \frac{1}{4}$$

- **Deuxième solution.** La longueur de la corde est déterminée par la distance de son milieu au centre du cercle. Par raison de symétrie nous pouvons considérer que le milieu de la corde est pris sur un rayon donné du cercle et supposer que la répartition de ce point sur le rayon est uniforme. La corde sera plus longue que le côté du triangle équilatéral inscrit si son milieu est à une distance du centre inférieur à $r/2$; la probabilité cherchée est alors $1/2$ (*cf.* fig. 1.3).

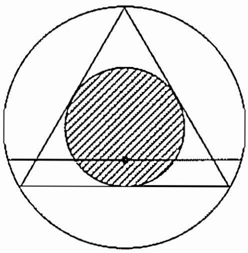

FIGURE 1.2

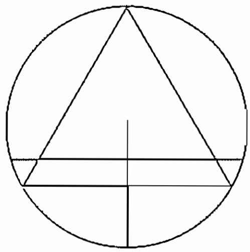

FIGURE 1.3

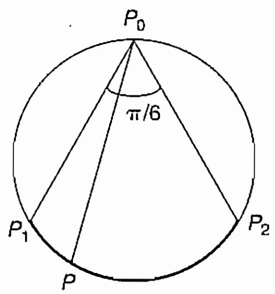

FIGURE 1.4

• **Troisième solution.** Par raison de symétrie nous pouvons supposer qu'on a fixé une des extrémités de la corde, soit P_0. L'autre sera choisie au hasard sur la circonférence. Si l'on admet que la probabilité pour que l'autre extrémité P tombe sur un arc donné de la circonférence est proportionnelle à la longueur de cet arc, la corde P_0P est plus grande que le côté du triangle équilatéral inscrit quand P se trouve sur l'arc P_1P_2 donc la longueur est le $1/3$ de celle de la circonférence (*cf.* fig. 1.4) ; la probabilité est alors $1/3$.

Il est clair que ces trois hypothèses de répartition, sont également réalisables. L'exemple parut paradoxal en son temps uniquement parce qu'on ne comprenait pas que des conditions expérimentales différentes pour le choix au hasard de la corde, dans les trois procédés décrits, conduisaient à des mesures-probabilités différentes sur la même algèbre d'événements.

1.4.1.3 La vision fréquentiste

Elle repose sur la loi des grands nombres (voir chapitre 2). Une seule expérience ne suffisant pas pour évaluer la probabilité d'un événement on va répéter un très grand nombre de fois l'expérience. Ainsi du lancer d'un dé : la probabilité d'observer la face 6 est la limite du rapport :

$$\frac{\text{Nombre de 6 obtenus}}{\text{Nombre d'essais}} = f$$

lorsque le nombre d'essais augmente indéfiniment. En effet la loi des grands nombres assure que f converge vers la probabilité p de l'événement.

Du point de vue pratique il est clair que la vision fréquentiste ne permet pas de trouver la probabilité d'un événement puisqu'un tel processus nécessitant une infinité d'observations est physiquement irréalisable : cela permet tout au plus de donner une définition de la probabilité comme limite d'une fréquence. Remarquons que dans la conception fréquentiste il est impossible de donner une valeur et même un sens à la probabilité d'un événement non répétable du genre « neigera-t-il le 25 octobre 2990 » ; ce qui limite le champ d'application du calcul des probabilités.

Cependant la critique la plus radicale du point de vue fréquentiste est la suivante : la définition de la probabilité repose sur la loi des grands nombres, or celle-ci est un théorème de probabilités qui suppose donc défini le concept de probabilité : il y a donc un cercle vicieux.

1.4.2 La conception subjectiviste

Le point de vue classique étant trop limité, le fréquentisme logiquement intenable, la probabilité d'un événement sujette à révision en fonction d'informations nouvelles (théorème de Bayes), l'existence même de probabilités objectives a été niée par beaucoup. C'est ainsi que le magistral Traité de Probabilités de de Finetti (1974) commence par l'affirmation en lettres capitales « *La Probabilité n'existe pas* » et continue par :

« *L'abandon de croyances superstitieuses sur l'existence du phlogistique, de l'éther, de l'espace et du temps absolu. . . ou des fées, a été une étape essentielle dans la pensée scientifique. La probabilité, considérée comme quelque chose ayant une existence objective est également une conception erronée et dangereuse, une tentative d'extérioriser ou de matérialiser nos véritables conceptions probabilistes !* »

1.4.2.1 Mesure d'incertitude

La probabilité objective d'un événement n'existe pas et n'est donc pas une grandeur mesurable analogue à la masse d'un corps, c'est simplement une ***mesure d'incertitude***, pouvant varier avec les circonstances et l'observateur, donc ***subjective***, la seule exigence étant qu'elle satisfasse aux axiomes du calcul des probabilités.

Les tenants de l'école subjectiviste proposent alors des méthodes permettant de passer d'une probabilité qualitative c'est-à-dire d'un simple pré-ordre sur les événements, à une mesure de probabilité.

Puisque la répétition n'est plus nécessaire on peut probabiliser des événements non répétables et étendre le domaine d'application du calcul des probabilités en particulier pour tout ce qui concerne les décisions économiques.

1.4.2.2 Le bayésianisme

Un pas de plus va être franchi par l'école bayésienne (ou plus exactement néo-bayésienne vu les deux siècles de décalage entre Bayes et ceux qui s'en réclament actuellement) qui va probabiliser tout ce qui est incertain et même des phénomènes non aléatoires.

Pour illustrer la théorie bayésienne modifions quelque peu l'exemple précédent de la fabrication des boulons : supposons qu'il n'y ait plus qu'une machine et que l'on cherche à estimer le pourcentage p de boulons défectueux produit en moyenne par la machine : si l'on admet qu'il n'y a que trois valeurs possibles p_1, p_2, p_3 respectivement égales à 0.3 %, 0.8 %, 1 % de probabilités *a priori* π_1, π_2, π_3 respectivement, la solution est inchangée et la valeur la plus probable *a posteriori* est 0.008 (si l'on tire un seul bouton défectueux). Supposons qu'on tire maintenant n boulons et que le nombre de boulons défectueux soit k, la probabilité que le pourcentage de défectueux produit par la machine soit p_2 est alors :

$$\frac{C_n^k p_2^k (1-p_2)^{n-k}\pi_2}{\sum_{i=1}^{3} C_n^k p_i^k (1-p_i)^{n-k}\pi_i}$$

On peut encore généraliser et supposer que p prenne toutes les valeurs possibles dans l'intervalle [0, 1]. Si l'on connaît la loi de probabilité de p sur [0, 1] et qu'elle admet une densité $f(p)$ par rapport à la mesure de Lebesgue, la formule de Bayes s'écrit :

$$P(p/k) = \frac{C_n^k p^k (1-p)^{n-k} f(p)}{\int_0^1 C_n^k p^k (1-p)^{n-k} f(p)\,\mathrm{d}p}$$

(voir chapitre 3).

A condition de connaître une distribution de probabilité *a priori* sur les valeurs de p, on peut donc en déduire les valeurs de p *a posteriori* les plus probables, donc estimer p.

On aura remarqué que p n'est pas aléatoire mais un paramètre fixe de valeur inconnue et que l'on a modélisé notre incertitude sur ses valeurs, par une mesure de probabilité. Mais comment choisir cette mesure *a priori ?* on retombe sur la difficulté signalée plus haut et, si cette probabilité est subjective, quel statut scientifique donner à une grandeur qui peut varier d'un observateur à l'autre ? Telles sont les critiques usuelles faites par les objectivistes. De plus on a montré qu'un ordre de probabilités donné n'induisait pas nécessairement une mesure de probabilité unique P sur Ω, compatible avec la relation d'ordre. P n'existe pas forcément ou encore, si P existe, P n'est pas toujours unique.

Nous arrêterons là ces quelques remarques et sans prendre parti dans une querelle qui dure encore, rappelons que le modèle probabiliste a prouvé son efficacité dans de nombreuses applications mais que comme tout modèle ce n'est qu'une représentation simplificatrice de la réalité et que ses hypothèses doivent être mises à l'épreuve des faits.

Nous renvoyons le lecteur intéressé par la philosophie des probabilités aux travaux de de Finetti (1974), Matalon (1967), Matheron (1978) et Savage (1954), cités en références.