Zeitschrift: Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker = Bulletin / Association des Actuaires Suisses = Bulletin / Association of Swiss Actuaries

Herausgeber: Vereinigung Schweizerischer Versicherungsmathematiker

Band: 55 (1955)

Artikel: La notion de probabilité et ses applications

Autor: Muller, Maurice

DOI: https://doi.org/10.5169/seals-550953

La notion de probabilité et ses applications

Par *Maurice Muller*, Zurich

Les problèmes relatifs à la nature et à la signification du calcul des probabilités ont fait, depuis le dix-huitième siècle surtout, l'objet de discussions intéressantes. Mais les solutions que l'on a essayé d'apporter à ces problèmes ont toujours été fonction du climat philosophique de l'époque. Au siècle du cartésianisme, le dix-septième, Fermat et même Pascal fondaient le calcul des probabilités sur une intuition claire des rapports entre les cas possibles. Au dix-huitième siècle, l'empirisme des philosophes anglais tend à se substituer au rationalisme cartésien et on commence à juger le calcul des probabilités d'après ses rapports avec l'expérience. Aux intuitions intellectuellement claires des cartésiens, on substitue l'idée en tant qu'abstraite de l'expérience sensible; on cherche à trouver dans l'expérience la justification ultime des mathématiques elles-mêmes, et l'on se demande, comme d'Alembert, si le calcul des probabilités est applicable à la réalité. En ce qui concerne le dix-neuvième siècle, on doit constater d'abord l'influence (et le prestige) des constructions mécanistes et déterministes de Lagrange et de Laplace – où, une fois posés les principes dont les applications garantissent le bien-fondé, la science se développe au moyen d'un processus rigoureusement analytique – ensuite et surtout, depuis le milieu du siècle, le succès de la philosophie kantienne auprès d'un grand nombre de savants. A l'empirisme des philosophes anglais et français, Kant avait opposé une doctrine fondée sur l'élaboration des données de l'expérience au moyen de fonctions tout intellectuelles, élaboration commandée par des jugements synthétiques *a priori* dont le fondement repose sur l'activité de l'entendement. Placé dans une telle perspective, le calcul des probabilités ne posait pas de question philosophique brûlante, dans la mesure toutefois où l'on n'opposait pas, au déterminisme issu de Laplace, un probabilisme qui aurait constitué l'essence même de la connaissance scientifique.

Sans doute, cette perspective historique est fortement simplifiée. Elle nous paraît cependant exacte dans la mesure où même les dissidents situent leur pensée par opposition aux principes implicitement ou ouvertement admis par leurs contemporains. Ceci est vrai, par exemple, de Cournot; et c'est aussi bien contre Kant que pour Leibniz qu'à la fin du dix-neuvième et au début du vingtième siècles un grand nombre de logiciens et de mathématiciens ont pris position.

C'est d'ailleurs avec cette réaction antikantienne que nous entrons dans le vif de notre exposé. Que prétendait, dans ses grandes lignes, la philosophie nouvelle? D'abord qu'il n'y a pas d'activité synthétique *a priori* au sens kantien. Le développement des sciences est essentiellement tautologique: les mathématiques partent d'axiomes, de principes, de conventions, librement posés ou consentis, et le mouvement qui conduit des principes aux conséquences n'est pas créateur de vérités nouvelles, n'est pas synthétique, mais analytique, les mathématiques étant près de se confondre avec la logique pour constituer le langage de la science. Ensuite que le fondement de la connaissance du monde réel, du monde tel qu'il nous est donné dans l'expérience, est essentiellement empirique: c'est au moyen de principes synthétiques issus de l'expérience, venant en quelque sorte alimenter la déduction, que le contact de ce langage avec la réalité est assuré.

Bien entendu, autour de cette position centrale ou par rapport à elle se sont orientées d'autres philosophies et d'autres attitudes scientifiques. Qu'il nous suffise de rappeler la protestation des analystes français tels que Baire, Borel et Lebesgue contre les excès du cantorisme, ou l'empirisme brouwerien, tout différent de l'empirisme logique du Cercle de Vienne auquel nous avons pensé plus spécialement. En nous plaçant dans une perspective proche de celle de l'empirisme logique, nous pourrons peut-être dégager quelques aspects de la notion de probabilité. Nous n'ignorons pas cependant que dans ces matières la façon même de poser un problème commande, dans une certaine mesure, les perspectives de la solution.

Nous allons essayer de distinguer la *structure logique* d'un calcul (nous ne disons pas du calcul des probabilités) de l'*usage* de ce calcul dans les problèmes pratiques, sans d'ailleurs nous inquiéter ici de savoir si les notions dont nous ferons usage ne font pas appel à quelque hypothèse implicite, ou ne sont pas surabondantes.

A cet effet, nous allons procéder de la manière suivante: nous partons d'un domaine que nous désignons par S et que nous admettons décomposable en sous-domaines A, B, C, etc., en nombre limité ou illimité, ayant ou n'ayant pas de parties communes. Il n'est pas du tout nécessaire de savoir ce qu'est un domaine; la notion de domaine a ici le rôle d'une notion indéfinissable propre à fixer notre langage. Nous admettons que notre domaine S est mesurable, c'est-à-dire qu'à ce domaine nous pourrons, au moyen d'un procédé de mesure dont il ne nous est pas utile d'indiquer la nature, faire correspondre un nombre positif bien défini. Nous admettons encore que la même règle est applicable d'une manière bien déterminée et univoque aux sous-domaines et surtout que, quelle que soit la manière dont un domaine est décomposé en sous-domaines, sans parties communes, le total des mesures des sous-domaines est égal à la mesure s du domaine S. Si deux sous-domaines ont une partie commune, il faut que cette partie commune ait la même mesure, qu'elle soit considérée comme sous-domaine de l'un ou de l'autre sous-domaine dont elle est partie.

D'une manière toute symbolique, nous pourrons représenter (comme en algèbre de la logique) le domaine S par une certaine surface limitée, et chaque sous-domaine par une partie de cette surface. Nous désignons par A et B deux sous-domaines (qui peuvent avoir ou ne pas avoir de parties communes et qui ne sont pas nécessairement chacun d'un seul tenant), par AB leur partie commune (AB peut être égal à 0) et par $A+B$ leur somme au sens de l'algèbre de la logique (pour laquelle $A+A=A$ et $AA=A$). Ecrivons mesure de A:

$$\text{mesure de} \quad A = m(A) = a;$$

on obtiendra évidemment les relations suivantes:

$$m(A+B) = a + b - m(AB)$$

et

$$\frac{m(A+B)}{m(S)} = \frac{a}{s} + \frac{b}{s} - \frac{m(AB)}{s}, \tag{I}$$

$$\frac{m(AB)}{m(S)} = \frac{m(AB)}{a} \cdot \frac{a}{s} = \frac{m(AB)}{b} \cdot \frac{b}{s}. \tag{II}$$

L'emploi, pour l'addition et la multiplication logiques, des mêmes signes que pour les opérations de même nom en arithmétique ne crée pas de confusion, puisque les lettres majuscules sont réservées ici aux termes logiques.

La relation I (ou la relation qui l'a précédée) fait en réalité partie de la notion de mesure telle que nous l'avons décrite. Le terme soustrait au second membre est introduit afin que la mesure du produit AB ne soit pas comptée deux fois. Quant à la relation II, elle est une simple tautologie.

Envisageons maintenant trois domaines, A_1, A_2, *disjoints* par hypothèse, et B, tels que B soit partie de $A_1 + A_2$ (ce qui entraîne la première des relations suivantes III); nous pouvons écrire:

$$m(A_1B) + m(A_2B) = m(B) = b$$

$$\frac{m(A_kB)}{m(B)} = \frac{\dfrac{a_k}{s} \cdot \dfrac{m(A_kB)}{a_k}}{\dfrac{a_1}{s} \cdot \dfrac{m(A_1B)}{a_1} + \dfrac{a_2}{s} \cdot \dfrac{m(A_2B)}{a_2}} \qquad \text{(III)}$$

avec, ici, $k = 1$ ou 2,

la seconde relation étant valable à condition que la première soit satisfaite. Ces relations III peuvent être facilement généralisées aux cas où l'on est en présence de n domaines A_k disjoints, B étant partie de la somme des domaines A_k. Si les sous-domaines A_k n'étaient pas disjoints, les relations III devraient être modifiées.

Nous avons admis que l'algèbre de la logique (plus particulièrement le calcul des classes au sens de cette algèbre), et que l'arithmétique sont applicables aux domaines considérés, les domaines pouvant être définis à l'aide de propositions plus ou moins compliquées, mais en prenant toute précaution indispensable lorsque les classes sont infinies. La mesure des domaines – que nous avons laissée indéterminée – pourra être définie, s'il s'agit de domaines indéfiniment décomposables, au moyen de procédés empruntés à la théorie des ensembles; elle pourra consister, si la décomposition du domaine S s'arrête de toute façon à un nombre fini de sous-domaines (ou, si l'on veut, d'éléments) en un simple dénombrement de sous-domaines et l'on pourra, dans certains cas, passer à la limite. Si le nombre des sous-domaines devient infini, on exigera que la somme s soit finie ou que, de toute façon, les rapports a/s aient un sens. Seules seront d'ailleurs acceptées les décompositions pour lesquelles la relation I est vérifiée; en d'autres termes, les notions de sous-domaine et de mesure retentissent l'une sur l'autre, et la relation I peut être regardée comme un principe du calcul.

Les propositions qui définissent les domaines et sous-domaines pourront être très différentes de nature. Sans porter de jugement préalable sur la signification du calcul des probabilités, nous allons adopter provisoirement le langage de la théorie élémentaire des probabilités et montrer à l'aide d'un exemple simple la signification que peuvent prendre les relations que nous avons établies.

Nous partons des propositions suivantes:

On jette un dé dont les six faces sont numérotées de un à six; l'une ou l'autre des faces sort $= S$.

Le dé montre une face paire $= A$.

Le dé montre une face dont le numéro est divisible par trois $= B$.

Le domaine S comprend les six cas possibles correspondant aux six faces du dé; le domaine A comprend les trois cas possibles 2, 4 et 6 et le domaine B les cas possibles 3 et 6. Comme convention de mesure des domaines, nous choisissons le dénombrement des cas qui y sont contenus. Visiblement, nous obtenons pour

$$\frac{a}{s}, \text{ la probabilité de sortir un nombre pair,}$$

$$\frac{b}{s}, \text{ la probabilité de sortir 3 ou 6,}$$

$$\frac{m(AB)}{s}, \text{ la probabilité de sortir 6,}$$

et la relation I correspond au théorème des probabilités totales (ici la probabilité de sortie d'un nombre pair ou divisible par 3); d'autre part,

$$\frac{m(AB)}{a} \text{ représente la probabilité pour la face 6 d'être sortie, sachant qu'un nombre pair est sorti,}$$

et la relation II représente le théorème des probabilités composées.

On constaterait de la même manière, à l'aide d'un schéma basé sur l'extraction de boules d'une série d'urnes données, que les relations III correspondent à la formule de Bayes. Il faut naturellement construire le domaine S et définir sa mesure d'une manière adéquate.

Toutefois, en adoptant un procédé de mesure étranger au calcul classique des probabilités, nous pourrions obtenir d'autres significations pour les relations I, II et III. Par exemple, si au lieu de mesurer les domaines suivant le nombre des cas élémentaires dits également possibles qu'ils contiennent, nous les mesurions en attribuant à chacun des cas élémentaires un nombre selon une règle arbitraire, nous obtiendrions en général une mesure cohérente, mais les relations I, II et III, tout en restant valables, ne seraient plus des relations entre des probabilités au sens classique; elles pourraient avoir, par exemple, le sens de relations entre des espérances mathématiques. Ceci montre bien que les relations I, II et III ont un caractère tout à fait général. Ce sont des relations formelles, sans contenu bien déterminé, applicables chaque fois que nous pourrons définir des domaines satisfaisant aux conditions que nous avons posées, en particulier au calcul élémentaire des probabilités; mais ni ces relations, ni la notion de domaine ne sont liées aux modèles classiques de ce calcul.

Pour passer au calcul des probabilités, il est indispensable de donner un sens aux signes utilisés dans le calcul formel. C'est à la recherche de la nature de ce sens que nous devons nous appliquer. Sans prétendre épuiser le sujet, nous le ferons à l'aide de quelques remarques. Nous voudrions cependant, avant de poursuivre, mentionner que notre schéma de domaines s'adapte facilement aux problèmes classiques du calcul des probabilités. On en déduit en effet facilement, sous une forme générale, les notions d'espérance mathématique, de moment et d'écart.

Nous nous arrêtons brièvement à la notion d'espérance mathématique.

Supposons que nous ayons divisé le domaine en sous-domaines *disjoints*, mais recouvrant ensemble le domaine S, de deux manières différentes:

$$A_1, A_2, \ldots, A_k \quad \text{et} \quad B_1, B_2, \ldots, B_e.$$

La partie commune à deux sous-domaines A_m et B_n sera

$$A_m B_n = C_{mn}$$

et l'on voit facilement que

$$\sum_n C_{mn} = A_m \quad \text{et que} \quad \sum_m C_{mn} = B_n.$$

Attribuons un nombre x_m à chaque sous-domaine A_m et un nombre y_n à chaque domaine B_n et désignons pour simplifier par p_m le rapport a_m/s, par q_n le rapport b_n/s et par p_{mn} le rapport c_{mn}/s.

Les espérances mathématiques sont définies par:

$$e(x) = \sum p_i x_i$$

ou

$$e(y) = \sum q_j y_j .$$

Le théorème des probabilités composées (II) donne

$$p_{mn} = \frac{m(A_m B_n)}{s} = p_m \frac{m(A_m B_n)}{a_m} = q_n \frac{m(A_m B_n)}{b_n} .$$

On voit aisément que l'on a $\sum (A_m B_n) = S$ et $\sum p_{mn} = 1$ ainsi que, par exemple, $\sum_n p_{mn} = p_m$. On aura donc

$$\sum_n p_{mn} x_m = p_m x_m$$

et on en déduit que

$$e(x + y) = \sum p_{mn}(x_m + y_n) = e(x) + e(y) .$$

De plus, si les probabilités sont indépendantes, c'est-à-dire si l'on a

$$\frac{m(A_m B_n)}{a_m} = q_n$$

on aura

$$p_{mn} = p_m q_n$$

et

$$e(xy) = \sum p_m q_n x_m y_n = e(x)\, e(y) .$$

Nous n'avons utilisé ici le terme de probabilité que pour la commodité du langage, ces relations ayant un caractère absolument général. Dans les applications au calcul des probabilités, il y aura lieu de morceler convenablement le domaine S de manière à obtenir des sous-domaines répondant aux problèmes posés [1]).

Si l'on joue avec deux dés, on sait que les espérances mathématiques s'attachant séparément aux deux dés satisfont au théorème sur la somme des espérances mathématiques. Mais le domaine S sera construit en tenant compte d'une manière convenable des combinaisons possibles entre les faces des dés dans les décompositions en sous-domaines auxquels seront affectées des valeurs x_m, y_n, ... etc., adéquates. On procédera d'une façon analogue pour les propositions sur les épreuves répétées, en constituant le domaine relatif à n épreuves. On parviendra ainsi au théorème de Jacques Bernoulli, en passant par la notion d'écart moyen quadratique et par l'inégalité de Tchebichef qu'on démontre aisément.

[1]) En général, il conviendra de faire correspondre des domaines adéquats à des valeurs x_i (ou y_i) distinctes.

La notion de domaine s'appliquera aussi, convenablement utilisée, aux problèmes des jeux, quelquefois relativement compliqués dans le cadre du calcul élémentaire des probabilités, où le jeu s'arrête non pas après une série bien déterminée de coups, mais lorsqu'un des joueurs a gagné. Il est clair que la décomposition des domaines et la mesure des sous-domaines dépendra de la nature du problème traité, et des informations, supposées suffisantes, en notre possession. On pourra aussi imaginer des chaînes de domaines, ou encore faire usage de la notion de domaine en théorie de la corrélation.

Cette notion de domaine s'étendra également au calcul des probabilités appliqué à des ensembles mesurables. Nous avons admis que la somme des mesures des sous-domaines (sans partie commune ou, dans la théorie des ensembles, sans point commun) doit être égale à la mesure du domaine total. C'est un postulat énoncé par Borel pour les ensembles mesurables. Conformément aux idées de Borel sur ces ensembles, il faut encore que la construction des sous-domaines réponde à une suite dénombrable d'opérations. Une décomposition en une infinité non dénombrable de sous-domaines nous exposerait d'ailleurs à toutes les difficultés qui s'attachent à l'étude d'ensembles ayant la puissance du continu, ce qui ne signifie naturellement pas que nous écartions la notion de probabilité continue telle qu'on la rencontre dans le calcul des probabilités, ni les cas où l'on est amené à faire dépendre la probabilité d'une fonction de point.

Nous pouvons maintenant procéder aux remarques annoncées.

Première remarque

Tout d'abord, on pourrait prétendre qu'en traitant notre problème d'une façon trop générale, nous avons nous-mêmes barré la route qui nous mènerait à la notion de probabilité. En restreignant notre notion logique ou mathématique de domaine, au risque de quelque définition complémentaire, nous devrions, semble-t-il, obtenir une intuition claire de la notion de probabilité. Nous voudrions montrer rapidement qu'on se ferait quelque illusion sur la valeur d'une telle manière de faire. Pour cela, nous admettrons que le domaine total S n'est décomposable qu'en un nombre fini de sous-domaines élémentaires, et nous nous bornerons aux schémas classiques du calcul des probabilités. Or, on a souvent constaté, à propos de ces schémas, qu'il s'agisse du jeu de pile ou face, ou des jeux de dés, ou des modèles d'urnes, que le calcul des probabilités en tant que calcul se réduit à une simple analyse combinatoire. Si nous disons qu'au jeu de pile ou face la probabilité de sortir pile seul en deux coups est de $^1/_4$, nous envisageons quatre combinaisons

$$PP,\ PF,\ FP,\ FF$$

et constatons que PP représente une combinaison sur quatre; du point de vue strictement mathématique nous n'avons rien de plus. Il est en effet clair que si nous avons le droit d'attribuer le *nom* de probabilité à un rapport entre des nombres de combinaisons, nous pourrions aussi attribuer à ce rapport, sans inconvénient, un autre nom, pour la simple commodité du langage. C'est en passant aux applications que des rapports de ce genre prennent un sens, et ce sens est ici physique. En effet, pour appliquer le calcul élémentaire des probabilités, nous admettons implicitement que la pièce de monnaie utilisée est physiquement telle que les deux éventualités P ou F sont de même poids ou, en langage de probabilités, sont également possibles. Il en résulte que les termes «également possibles» ou «également probables» ont aussi une signification essentiellement physique et non logique ou mathématique. Le fait que nous ayons besoin d'une information physique sur les urnes, la symétrie de la pièce ou du dé pour décider de l'application du calcul élémentaire des probabilités, en est une confirmation.

Deuxième remarque

Notre seconde remarque concerne un premier essai de comparaison entre les notions de probabilité et de fréquence. Si l'on suppose que la notion de probabilité est abstraite de l'expérience, on doit se demander jusqu'à quel point on peut en faire une fréquence idéale. En fait, les statisticiens affirment en général que, tout au moins en statistique, la fréquence (relative) joue le même rôle que la probabilité. Pour la comparaison qui nous occupe, nous ferons un instant appel à un mathématicien du dix-huitième siècle, d'Alembert, largement influencé par les idées empiristes des Anglais. Nous ne modifions pas notre perspective initiale, puisque le point de départ que nous avons adopté peut être regardé comme une combinaison de philosophie empiriste et de logique leibnizienne.

D'Alembert ne croyait pas beaucoup à l'application illimitée des mathématiques aux sciences de la nature, et le calcul des probabilités lui semblait une construction mathématique dont les applications aux phénomènes naturels et aux jeux de hasard sont des plus douteuses; il proposait de lui substituer un art des conjectures plus général et moins mathématique. Je pense qu'il aurait écarté, comme inutilisable, le postulat de Bertrand selon lequel les dés n'ont ni conscience ni mémoire. D'Alembert se sert en particulier de l'exemple suivant: deux joueurs, A et B, jouent à pile ou face. Si pile sort au premier coup, le jeu s'arrête et A a gagné; si face sort, on joue encore une fois et une seule; si pile sort au second coup, A a gagné, mais de toute façon le jeu s'arrête. D'Alembert, profitant du fait que le jeu s'arrête si pile sort au premier coup, ne considère que les trois combinaisons pile, face-pile et face-face et pense qu'il pourrait être opportun de fixer à $^2/_3$ la chance de gain du joueur A, alors que le calcul élémentaire des probabilités la fixe à $^3/_4$. Je pense que si l'on prenait d'Alembert au mot, on construirait un calcul des probabilités contradictoire. En effet, il suffit d'une simple modification à l'énoncé de son problème, pour aboutir au nombre prévu par le calcul des probabilités. Si on énonce le problème posé de la manière suivante: «un joueur A a le droit de lancer deux fois une pièce de monnaie: il sera gagnant contre B si pile sort au moins une fois», il est clair que si pile sort au premier coup, A aura de toute façon gagné et le résultat du second coup lui deviendra indifférent. Mais alors l'énoncé du problème entraîne clairement la considération des quatre combinaisons pile-pile, pile-face, face-pile et face-face.

D'Alembert ne prétendait d'ailleurs pas fonder un calcul des probabilités sur un tel exemple. Il admettait volontiers que le résultat $^3/_4$ est «mathématiquement parlant» (c'est l'expression qu'il emploie) plus juste que le résultat $^2/_3$. Mais il pensait qu'il ne l'est peut-être pas physiquement et que l'art des conjectures ne peut pas être soumis à un calcul également net dans ses principes et dans ses résultats.

Pour décider de la valeur du calcul classique des probabilités, un empiriste procéderait à des expériences. D'Alembert semblait admettre que des expériences ne conduiraient qu'à des résultats incertains ou, d'une manière plus elliptique, que le hasard ne peut pas être soumis au calcul. Un pur empiriste, plus favorable que d'Alembert au calcul des probabilités, penserait au contraire que les fréquences observées dans des conditions convenables garantiraient la légitimité du calcul. Mais il faut encore que ce calcul, en tant que mathématique appliquée, ne soit pas contradictoire dans son ensemble. En d'autres termes, il faut que les règles de mesure adoptées soient cohérentes pour tous les problèmes de même nature physique. On peut sans doute ajouter que la nature physique d'un problème commande le procédé de mesure qui lui sera appliqué, mais l'on devra convenir que la cohérence du calcul classique des probabilités en tant que mathématique appliquée est due en grande partie à la notion d'éventualités physiques posées *a priori* comme également possibles, ou comme possibles dans des rapports supposés *a priori* connus (même s'ils restent indéterminés dans le calcul).

Troisième remarque

Poursuivons, au moyen d'une troisième remarque, notre comparaison entre fréquence et probabilité. Nous supposons que nous sommes en possession d'une pièce de monnaie parfaitement symétrique, éprouvée au moyen d'un grand nombre d'épreuves, par exemple 10 000. Nous admettons qu'après nous être assurés de cette symétrie physique dans la mesure la plus grande possible, nous lancions la pièce un nombre impair de fois, ni trop grand, ni trop petit, par exemple 99, et que nous obtenions au cours de cette série d'épreuves 52 sorties de pile et 47 sorties de face. Tout en leur exposant les données desquelles nous sommes partis, ici la symétrie de la pièce expérimentalement constatée et le résultat global de la série de 99 épreuves, nous nous adresserons à deux personnes, un probabiliste comme Bertrand et un pur statisticien. Nous leur demanderons quelle est, selon eux, la probabilité de sortie de pile pour la dixième épreuve dans la série de 99 coup déjà effectués. Nous supposons que le premier réponde que la probabilité est de $^1/_2$, puisque la pièce n'a ni conscience ni mémoire, et le second, dérivant la probabilité de la fréquence, que cette probabilité est de $^{52}/_{99}$. Sans vouloir accorder trop d'importance au fait qu'ayant limité la série à un nombre impair d'épreuves effectivement réalisées, la seconde réponse ($^{52}/_{99}$) ne pouvant jamais être la même que la première ($^1/_2$), nous voudrions insister sur la différence de point de vue et mettre en relief le caractère idéal et abstrait de la notion de probabilité.

Tout d'abord la probabilité peut appartenir à l'essence d'un seul événement: la fréquence, jamais. Ceci est d'observation courante, presque triviale, mais ce n'est pas une distinction limitée au langage; on peut parler de la probabilité de décès d'une personne, alors que les tables de mortalité ne livrent en réalité que des fréquences. Dans le jeu de pile ou face, on peut observer la fréquence de pile dans une série donnée, mais on parle de la probabilité de sortir pile au premier coup. Et même un empiriste pourra convenir qu'un tel processus d'abstraction a quelque chose de légitime. C'est d'ailleurs par un décret porté par le théoricien qu'une signification de probabilité ou de fréquence est accordée à certains rapports tout formels de la théorie mathématique, notamment aux rapports entre des mesures de domaines et de sous-domaines.

Ensuite, les deux réponses sont des réponses à une question interprétée de deux manières différentes. Le probabiliste se place à la fois «dans l'abstrait» et, pour employer un terme d'inspiration bergsonienne, dans l'attitude prospective du «se faisant»; le statisticien se place dans l'attitude rétrospective du déjà fait, et se demande combien il a de chances, s'il choisit l'un des coups *déjà effectués,* de rencontrer pile. La nature de ces réponses suffit cependant à montrer qu'une nuance sépare la notion de probabilité de celle de fréquence.

De nos trois premières remarques, nous pouvons tirer une conclusion unique et plus générale, conforme à nos prémisses: la notion de probabilité n'appartient pas au calcul lui-même, mais à la signification que l'on donne au calcul, que celle-ci ait un caractère purement empirique comme dans la fréquence statistique ou qu'elle ait un caractère plus abstrait, ou même intuitif comme une idée claire au sens de Descartes. Ceci correspond d'ailleurs à une opinion fort répandue selon laquelle la structure logique d'une théorie mathématique doit être distinguée des significations que la théorie peut recevoir. Nous avons en outre constaté que la notion de probabilité ne résultait pas immédiatement de celle de fréquence.

Nous passons maintenant à une quatrième remarque relative à l'application du calcul des probabilités dans la théorie cinétique des gaz ou des ensembles de particules.

Quatrième remarque

Admettons que nous disposions de trois particules a_1, a_2, a_3 et de quatre cellules e_1, e_2, e_3, e_4 et que les trois particules soient distribuées de toutes les manières possibles entre les quatre cellules. Nous obtiendrons 120 distributions possibles (ou 120 complexions, pour utiliser le langage des physiciens), à condition de considérer les trois particules comme parfaitement individualisables; en d'autres termes, si nous plaçons les particules a_1 et a_2 dans la cellule e_i, ou dans deux cellules différentes, nous admettrons que l'échange des deux particules entre elles conduit de toute façon à une complexion nouvelle. Ces 120 possibilités représentent la manière la plus générale de distribuer les trois particules bien individualisées selon les quatre cellules. Le domaine S pourra alors être morcelé en 120 cases (une par complexion) et une probabilité relative à un sous-domaine sera mesurée au nombre de cases qu'il contient. Tel n'est cependant pas le point de vue de la mécanique statistique classique qui admet que l'échange entre particules situées dans la même cellule ne donne pas de complexion nouvelle et que l'échange entre particules de cellules différentes donne lieu à des complexions nouvelles. Ceci a pour effet de réduire à 64 le nombre de complexions. Si l'on supprime cette restriction à l'indiscernabilité des particules, c'est-à-dire si l'on admet qu'aucun échange entre particules ne modifie le nombre de complexions possibles, il ne reste que 20 complexions. C'est le point de vue de la statistique de Bose. Sans doute on pourrait, pour chaque statistique, créer un domaine S adéquat selon le caractère (discernable ou indiscernable) des particules, celles-ci étant alors distribuées selon un autre critère dans la statistique de Bose que dans la statistique classique. Mais si l'on part du domaine initial à 120 cases, on remarquera que pour passer soit à la statistique classique, soit à la statistique de Bose, on doit poser certaines restrictions à la formation des sous-domaines d'égale probabilité. Ceci est encore plus apparent avec la statistique de Fermi, qui attribue une probabilité nulle à toutes les cases pour lesquelles une cellule contiendrait plus d'une particule.

Dans ces statistiques, c'est l'expérience qui tranchera entre les différentes manières de morceler et de mesurer le domaine initial S et qui décidera des complexions qui devront être regardées comme d'égale probabilité. Et il est clair qu'on pourrait imaginer d'autres manières de mesurer S, qui toutes satisferont aux règles essentielles du calcul des probabilités si la relation fondamentale des probabilités totales est satisfaite. Aucune considération *a priori* ne permettra de décider à elle seule entre les différentes manières possibles de mesurer le domaine S. Que ce soit directement ou par le détour de la mécanique ondulatoire, l'expérience décidera en dernière analyse de la validité des hypothèses faites, c'est-à-dire fixera la notion d'également probable; en particulier, si l'on pouvait vérifier sans équivoque qu'un ensemble de particules d'un certain genre (par exemple de photons pour la statistique de Bose et d'électrons pour la statistique de Fermi) obéit à l'une ou à l'autre des statistiques, on vérifierait du même coup les hypothèses qui ont servi de base à la mesure du domaine S, en particulier l'hypothèse qui détermine les complexions d'égale probabilité. Mais l'on voit immédiatement que cette vérification ne passerait pas nécessairement par un détour qui consisterait à observer la fréquence d'événements successifs comme au jeu de pile ou face. Ici encore la notion de fréquence statistique ne nous paraît pas recouvrir entièrement la notion de probabilité. Ici, la probabilité posée *a priori* trouverait sa justification aussi bien dans des considérations de symétrie et d'indiscernabilité (et, pour la statistique de Fermi, dans le principe d'exclusion), que dans les vérifications expérimentales propres à l'ensemble de la théorie des quanta.

Cinquième remarque

Nous voudrions encore montrer très brièvement que le formalisme de la mécanique ondulatoire autorise plusieurs interprétations de l'indéterminisme propre aux phénomènes physiques, en nous bornant toutefois à quelques indications très générales.

On sait que l'état d'un corpuscule peut être représenté par une fonction ψ, solution de l'équation des ondes

$$H\psi = \frac{h}{2\pi i} \cdot \frac{\partial \psi}{\partial t}$$

(H est l'opérateur hamiltonien),

et qu'un principe d'interprétation physique de la mécanique ondulatoire fait correspondre à toute grandeur observable des opérateurs A d'une certaine nature (ils sont linéaires et hermitiens), dont les valeurs propres α_i, relatives à l'équation

$$A\varphi = \alpha\varphi,$$

sont les valeurs possibles de la grandeur physique mesurable. Les fonctions φ_i, solutions de cette équation, sont *orthogonales* et on les suppose *normées* ainsi que la fonction ψ. En d'autres termes, on doit avoir [1]):

$$\int \psi\psi^* d\tau = 1; \quad \int \varphi_n\varphi_m^* d\tau = 0; \quad \int \varphi_n\varphi_n^* d\tau = 1.$$

On montre alors qu'on peut développer ψ en fonction des φ_i et l'on obtient

$$\psi = \sum c_n\varphi_n$$

avec, en formant le produit $\psi\psi^*$ (où ψ^* est l'imaginaire conjugué de ψ),

$$\sum c_n c_n^* = 1.$$

On peut distinguer trois manières de voir:

1° Le produit $\varepsilon\psi\psi^*$ (où ε est la charge électrique) représente une densité de charge et celle-ci est disséminée dans tout l'espace: dans ce cas, l'introduction de la notion de probabilité est évitée. On sait que cette première interprétation a dû être abandonnée, mais elle semble bien montrer que les schémas théoriques n'ont pas nécessairement une signification probabiliste.

2° Le produit $\psi\psi^*$ est une densité de probabilité et fournit une probabilité de présence (d'une particule dans un domaine élémentaire). Si ψ est décomposé selon l'hamiltonien H, on aura

$$\psi = \sum c_n\psi_n,$$

les ψ_n étant les solutions de l'équation de Schrödinger, $H\psi = \alpha\psi$, qui fournit les états stables de l'atome. Si l'on ne connaît pas ψ, l'on devra effectuer des mesures et les fréquences des réponses $\psi_1, \ldots, \psi_n$, etc., permettront de fixer les amplitudes c_n. C'est une interprétation statistique, la probabilité de présence étant alors fondée sur un calcul statistique.

3° Si la fonction ψ est supposée connue ou est choisie *a priori* (on peut la faire correspondre à l'un ou l'autre des états stables), alors les coefficients c_n dans le développement

$$\psi = \sum c_n\varphi_n$$

[1]) $d\tau$ est l'élément d'espace.

seront parfaitement déterminés. Nous aurons une interprétation probabiliste, les $|c_n|^2$ étant des probabilités ne dépendant que du choix de la fonction ψ.

Il est difficile de choisir entre la seconde et la troisième interprétation, et même de les distinguer entièrement; elles dépendent de toute évidence de la *manière* dont la question est posée. Les fondateurs de la mécanique matricielle – Sommerfeld l'avait rappelé dans un volume complémentaire à son traité *Atombau und Spektrallinien* – partaient d'une conception positiviste proche du point de vue statistique. Mais le succès de la mécanique ondulatoire et l'indéterminisme qui lui est inhérent feraient plutôt pencher pour la troisième interprétation, la notion de probabilité étant alors partie *a priori* des principes d'interprétation physique de la mécanique ondulatoire.

Au point de vue général auquel nous nous sommes placés, on peut faire correspondre à la fonction ψ un domaine S, et aux opérateurs A différentes manières de décomposer S en sous-domaines.

* * *

L'analyse de la notion de probabilité à laquelle nous venons de nous livrer en nous servant d'exemples pratiques n'aboutit pas nécessairement à une conclusion nette. Nous voudrions simplement essayer d'en tirer une ou deux brèves leçons.

Nous avons tout d'abord observé que la notion de domaine que nous avons adoptée n'est pas nécessairement attachée à une interprétation probabiliste ou statistique. Il faut attribuer aux rapports entre des mesures de domaines et sous-domaines la signification de probabilités ou de fréquences relatives pour passer à un calcul des probabilités ou à un calcul statistique.

Nous avons ensuite distingué, dans la notion de probabilité, une idée abstraite qui n'intervient pas de la même manière dans la notion plus directement expérimentale de fréquence relative. L'idée de probabilité ne coïncide pas exactement avec celle de fréquence statistique et ne semble pas pouvoir y être ramenée entièrement.

Nous pouvons d'autre part observer qu'une différence, semble-t-il capitale, sépare l'usage de la notion de probabilité dans la statistique usuelle de son application dans les théories physiques. Un statisticien est dans la pratique de son art véritablement à la recherche de fréquences dont il tend à faire dériver des probabilités; pour lui, une loi statistique ne pourra être vérifiée qu'en observant des fréquences effectives. Au contraire, la théorie physique ne reçoit pas ses vérifications de l'observation de fréquences – ou tout au moins pas exclusivement – mais d'expériences qui peuvent être uniques, d'observations et de mesures comme celles effectuées sur les raies du spectre de l'hydrogène ou de l'hélium. C'est à la fois la cohérence de la théorie physique et sa vérification par des expériences qui autorisent l'usage de la notion de probabilité en physique.

Nous avons laissé en suspens la notion de *limite* que celle de fréquence semble supposer et à laquelle on a voulu ramener la notion de probabilité. En effet, cette notion de limite de fréquence semble dériver d'une conception idéale et abstraite de la fréquence qui, de toute façon, procède d'une rupture entre la fréquence physique et sa limite idéale. Ou bien on tend à confondre la notion de limite de fréquence avec la notion de probabilité – et alors on a changé le nom sans changer la signification – ou bien on tend à rattacher la probabilité à une vue expérimentale des choses – et alors on s'expose à des difficultés insurmontables.

On sait par exemple que le jeu de pile ou face peut être assimilé à la construction d'un nombre dans un système numérique binaire et que l'ensemble de tous les résultats possibles, obtenus en supposant la suite des épreuves *idéalement* réalisée jusqu'à l'infini (ce qui semble déjà contestable), s'étale entre 0 et 1 et a la puissance du continu. Si l'on admet que l'on peut choisir l'une de ces limites, on risque de se heurter aux difficultés de l'axiome du choix de Zermelo. Sans doute, on doit penser que les limites s'accumuleront en fait dans la région $^1/_2$, mais pour l'affirmer on invoque directement ou indirectement les résultats généraux du calcul *classique* des probabilités.

Pour retrouver les notions de probabilité et de fréquence en partant de nos domaines, il faut, ainsi que nous l'avons vu, prêter une signification aux éléments de notre calcul, mais il est difficile, sans faire appel à l'intuition, d'établir une frontière nette entre les deux notions. Afin de les distinguer, nous allons employer, en le simplifiant, un langage emprunté à Husserl. Husserl, revenant à des positions cartésiennes, a repoussé une philosophie kantienne selon laquelle la pensée opère une synthèse unifiante du «divers éparpillé [1])» livré par la sensibilité. Pour

[1]) Cette expression est de Bréhier.

Husserl, ce qui caractérise la conscience, c'est son intentionalité, c'est-à-dire une manière de contenir autre chose qu'elle, de prêter un sens aux structures qui s'offrent à elle et dont la nature varie avec celle de l'objet, formel ou non, qu'elle se donne. Or, si nous voulons distinguer plus profondément que nous ne l'avons fait la notion de probabilité de celle de fréquence, un langage emprunté à Husserl pourrait convenir. Alors que la notion de fréquence aurait un sens essentiellement dérivé de l'expérience, la notion de probabilité répondrait à l'acte par lequel nous prêtons ou donnons un sens aux rapports formels de nos calculs. Et, comme chez Descartes, ce sens pourrait être impliqué dans une intuition. Sans doute, on peut bien dire que la probabilité est une fréquence idéale ou abstraite, mais l'on devra alors convenir que l'adjonction des termes «idéale» ou «abstraite» (comme l'usage du terme de limite) change considérablement la signification, essentiellement expérimentale, du terme de fréquence.

Nos dernières remarques sur la notion de probabilité viseront ses rapports avec d'autres notions, celle de vérité et celle de temps ou, d'une façon toute générale, celle de contingence. On sait que pour Aristote, dès que l'on n'a plus affaire au nécessaire éternel ou au réel actuel et donné, mais au futur contingent, le principe de contradiction cesse d'avoir une fonction déterminante. C'est d'une manière *indéterminée* qu'il est vrai ou qu'il est faux que «demain il y aura bataille navale». Or, en dégageant les propriétés formelles du calcul des probabilités, nous sommes restés dans un monde entièrement intemporel, où le temps ne peut intervenir qu'en qualité de paramètre qui ne trouve de signification que dans les applications. Si, en passant à l'*interprétation pratique* du calcul des probabilités, nous ordonnons selon une dimension temporelle les réponses aux questions relatives à la probabilité d'un *événement futur*, alors le caractère d'indétermination dénoncé par Aristote apparaît clairement. Comme d'autre part les notions de vérité et de fausseté *relatives* n'apparaissent pas comme telles dans notre calcul, en tant que fondé sur la logique des classes, nous pouvons admettre qu'un événement futur sera certain (et la proposition affirmative correspondante absolument vraie) si la probabilité qui lui correspond est égale à *un*, qu'il sera impossible si sa probabilité est égale à *zéro* (et la proposition affirmative correspondante sera fausse) et qu'il sera incertain dans les cas intermédiaires (et la proposition correspondante ni absolument vraie, ni absolument fausse). On peut dès lors penser que des logiques à plusieurs valeurs conviendraient au calcul des probabilités, comme elles semblent convenir aux théories physiques contemporaines.

Mais la notion de probabilité physique entretiendrait une relation encore plus étroite avec celle de temps. L'évolution d'un système physique dans le temps, selon le second principe de la thermodynamique, se ferait toujours dans le sens des états les moins probables vers les états les plus probables. Par rapport à un système qui n'évolue pas, le temps est, si l'on peut dire, comme s'il n'existait pas; pour que le temps soit, il faut que le monde évolue, qu'il passe d'états moins probables à des états plus probables. De là à faire dépendre la notion de temps de celle de probabilité, il n'y a qu'un pas, il est vrai difficile à franchir, puisque la notion de temps paraît plus immédiate et plus qualitative que la notion de probabilité.

Sur la technique actuarielle

Supposons que des observateurs aient choisi au hasard, dans une population humaine déterminée, une collection de cent mille individus, de sexe masculin par exemple, âgés de 20 ans, et qu'ils aient noté combien d'individus de la collection sont encore en vie à 21 ans, à 22 ans et ainsi de suite, jusqu'à extinction. La collection initiale constitue un domaine S, et les collections successives d'individus encore en vie aux âges de 21 ans, 22 ans, etc., des sous-domaines de S, mesurés par le nombre d'individus qu'ils contiennent. D'autres sous-domaines seront constitués par les différences entre un sous-domaine et les sous-domaines qu'il contient. Le rapport entre la mesure d'un sous-domaine à celle d'un sous-domaine qui le contient (ou au domaine S) aura la signification d'une fréquence statistique. Si l'on admet que les observateurs ont exclusivement porté leur attention sur des nombres d'individus, sans distinguer les individus les uns des autres, c'est-à-dire sans noter que tel individu est mort à tel âge, ils pourront, leur statistique étant construite, choisir un individu particulier de la collection initiale et se demander combien il y a de chances pour que *cet* individu ait été encore en vie à l'âge de 21 ans, ou 22 ans, etc. Nous sommes donc dans la situation décrite au cours de notre troisième remarque: nos statisticiens répondront, correctement puisque l'individu choisi appartient à la collection initiale, que les chances sont égales aux fréquences relatives qu'ils ont calculées. Mais si l'individu choisi n'appartient pas à la collection initiale – tout en appartenant à la même population – on postulera que les fréquences relatives déduites de la collection donnée peuvent être approximativement assimilées à des probabilités dont les valeurs exactes sont inconnues et sont, en tant que telles, applicables à l'individu choisi. Ceci suffit à l'actuaire qui, en construisant une table de mortalité, admet implicitement que les fréquences oscilleront à l'intérieur de limites relativement étroites et qui ajuste ses tables ou modifie ses fréquences relatives en fonction du but qu'il désire atteindre.

En fait, il n'est pas nécessaire d'admettre l'existence de probabilités abstraites auxquelles la mortalité, au sein d'une population donnée, serait soumise. L'existence de probabilités abstraites supposerait que la mortalité générale évolue selon des lois qu'on pourrait à la rigueur formuler analytiquement et dont ces probabilités pourraient aisément être déduites, ou que la mortalité est véritablement assimilable aux modèles classiques du calcul des probabilités, modèles dont la notion de domaine, applicable aussi bien aux probabilités qu'aux fréquences, nous a libérés. Des mathématiciens, Volterra notamment, ont sans doute créé une biologie analytique relative à l'évolution des populations animales et végétales en concurrence les unes avec les autres, mais c'est seulement dans des cas très simples ou très théoriques qu'ils ont pu obtenir des résultats précis ou pratiquement utilisables. Si l'on observe que l'homme est un destructeur des équilibres biologiques naturels, que les populations humaines sont en concurrence les unes avec les autres (au moins actuellement), que la croissance d'une population ou son accès à un certain degré de maturité politique ou économique peuvent avoir des répercussions sur l'évolution des populations voisines, on doutera qu'une loi concernant l'évolution d'une population déterminée (et en conséquence l'évolution de sa mortalité) puisse être jamais énoncée. Ce n'est pas sans raison que des économistes s'inquiètent de l'introduction, moralement si légitime, de l'hygiène dans des pays à forte natalité. En outre, l'épuisement des sources de matières premières d'un pays et la perte d'un marché pour ses produits industriels pourront modifier radicalement la structure de la population qui lui appartient.

L'inexistence de lois concernant l'évolution des populations semble donc bien entraîner l'inexistence de «probabilités» relatives à la mortalité au sens classique. Ceci n'empêche d'ailleurs nullement de construire des modèles empruntés au calcul des probabilités, ou d'user de fréquences observées ou ajustées en technique actuarielle.

On a quelquefois tendance à restreindre l'application du terme «actuariel» aux techniques de l'assurance sur la vie, l'invalidité ou la maladie. En fait, beaucoup d'autres formes d'assurance relèvent des mêmes principes généraux et reposent sur la notion de fréquence relative, ou sur des rapports analogues (mais qui peuvent devenir supérieurs à un) tels que celui du total des paiements à la suite de sinistres au total des sommes assurées. Mais, dans ces formes d'assurance, une notion nous paraît s'imposer: celle de pari. Car, ici, le passé n'est pas, comme dans l'assurance sur la vie (ou tout au moins dans la même mesure), garant de l'avenir. Les opérations financières qui se traitaient aux Lloyd's ont tenu, au moins à leur début, du pari. Et le pari est presque toujours fondé sur cet art des conjectures que souhaitait d'Alembert – où les probabilités deviennent qualitatives avec des degrés de plus ou de moins – pour lequel la notion de probabilité totale ne serait pas valable sans restriction. D'ailleurs, dans l'assurance sur la vie elle-même, l'acceptation ou le refus d'un risque font l'objet d'une appréciation qualitative, d'un pari orienté par les renseignements contenus dans la proposition d'assurance ou dans le rapport médical. Cette notion de pari nous semble aussi s'introduire dans les jeux pour lesquels les informations sont incomplètes. Si nous sommes mal renseignés sur la symétrie d'une pièce et si nous jouons à pile ou face avec un partenaire, nous pouvons parier que la pièce est symétrique ou qu'elle ne l'est pas dans un sens que nous choisissons, et évaluer nos chances en conséquence. Le défaut d'information suffisante entraîne donc l'effacement de la notion de probabilité quantitative, et ceci pourrait aussi être vrai pour l'excès d'information. Il y a également une part de pari dans tous les problèmes où la probabilité est fonction de notre ignorance. Ainsi, lorsque nous disons que la probabilité pour que la cent-millième décimale du nombre e soit un 7 est de $^1/_{10}$, nous parions que l'événement peut être considéré comme soumis au hasard.

Pour terminer, nous voudrions relever que nous nous sommes abstenus, au cours de notre exposé, de prendre position sur des recherches bien connues sur le fondement du calcul des probabilités: celles de von Mises, dont la notion de limite de fréquence a été fort contestée, celles de Wald qui a continué l'œuvre de von Mises, celles de Kolmogoroff, qui fonde sa théorie sur les ensembles, celles de Reichenbach, sans oublier les recherches de Borel sur la mesure des ensembles, capitales pour l'objet qui nous a occupés, celles de Paul Lévy, de Fréchet, de de Finetti et de bien d'autres auteurs. La simple énumération de ces noms montre que notre aperçu sommaire est loin d'épuiser un sujet très vaste.