

À PROPOS DU TERME «PROBABILITÉ»

H. BRENY (*)

«*Must* a name mean something?» Alice asked doubtfully.

«Of course it must » Humpty Dumpty said with a short laugh; «*my* name means the shape I am – and a good handsome shape it is, too. With a name like yours, you might be *any* shape – almost.»

1. POLYSÉMIE

Lorsqu'on parcourt l'un quelconque des grands ouvrages qu'a suscités ces derniers temps la philosophie des sciences, l'œil voit défiler, page après page, ces termes généraux: «théorie», «hypothèse», «modèle» (ou «langage», selon le cas) qu'aucun auteur ne se risquerait à inclure dans son Index rerum. Il semble bien que, après ceux-là, l'un des plus fréquemment employés est «probabilité». Il est bien clair qu'une telle fréquence d'occurrence n'est guère un indice d'univocité. Et, en effet, «probabilité» est, parmi les termes scientifiques, l'un de ceux qui ont le plus de sens différents. Carnap [1] en avait distingué deux; mais il s'était arrêté avant le but – encore que ce fût sur le bon chemin. La première tâche de l'épistémologie, à ce sujet, est donc d'énumérer ces sens; il lui faudra ensuite étudier s'ils sont, ou non, en relation les uns avec les autres; et, ensuite seulement, étudier en détail les divers rôles qu'ils assignent à la probabilité dans le processus de la connaissances humaine.

Le premier, le plus simple de ces sens, celui aussi qui soulève le moins de discussions, fait de «probabilité» l'ingrédient essentiel d'une certaine structure mathématique; sens purement formel et abstrait («application bornée et totalement additive d'une tribu d'ensembles dans $[0;1]$ ») qui ne retiendra pas, ici, l'attention. Il faut pourtant remarquer qu'une autre structure mathématique a, pour un temps,

(*) professeur ordinaire, Université de l'État à Liège (Institut de mathématique)

tenu la vedette, celle qui, due à von Mises [18], prit par après le nom de «théorie des collectifs»; elle met en jeu, comme modèle prétendument universel d'un phénomène fortuit, une suite prenant ses valeurs dans un ensemble quelconque et jouissant des deux propriétés que voici: d'une part, la fréquence d'apparition d'une de ces valeurs (quelle qu'elle soit) parmi les n premiers termes tend vers une limite bien définie (qui est, précisément, sa probabilité); d'autre part, la succession de ces valeurs ne montre pas la moindre trace de régularité. Non seulement cette structure n'est que médiocrement satisfaisante pour le mathématicien (car «ne présenter aucune trace de régularité» est une propriété mal définie, voire contradictoire, et on n'arrive à une définition mathématiquement satisfaisante qu'en faisant fi de propriétés qui paraissent pourtant s'imposer à l'intuition) mais encore (mais surtout!) n'est-elle susceptible d'aucune application, si ce n'est aux jeux de hasard indéfiniment répétés (et encore!). C'est pourtant elle seule que citent, à ce propos, les philosophes (singulièrement, Carnap tout au long de ses œuvres et Popper avant 1957; Popper a même contribué à l'étude *mathématique* de cette structure). Ce fait est fort curieux, car la «théorie classique modernisée», comme on l'a appelée, c'est à dire celle qui fait tout uniment emploi de la structure de tribu probabilisée, est d'usage commun, parmi les mathématiciens tournés vers les applications, depuis 1920 (voyez p. ex. [19]). Or, la théorie des collectifs a des particularités malheureuses qui colorent de façon fort discutable l'analyse épistémologique qui en fait usage; de ce point de vue, une partie des textes les plus classiques de Carnap serait à réécrire.

Un autre sens est relatif au monde concret, au monde de la nature sensible. On y distingue des situations, dites «aléatoires», donnant lieu à des événements, dits «fortuits», dont on suppose qu'ils ont, chacun, une tendance plus ou moins forte à se produire; on suppose en outre que ces tendances sont mesurables, et on appelle «probabilités» leurs mesures numériques. [C'est ce sens-là, encore qu'avec des nuances, que Carnap note «probabilité₂».]

Tout au contraire, les deux autres sens se rapportent non pas au monde physique mais à la connaissance humaine elle-même (ou, pour les plus positivistes d'entre nous, aux énoncés qui l'expriment); ils sont donc dits «épistémiques». «Ils sont dits», au pluriel: car ici une nouvelle distinction s'impose absolument. Pour les uns en effet – et

c'est assurément le sens épistémique le plus obvie, c'est sans doute aussi le sens du mot pour l'homme de la rue – «probabilité» se rapporte aux opinions auxquelles on adhère avec une conviction moins grande que la certitude proprement dite, et donc sans exclure tout à fait les opinions contraires; comme quand la mère au foyer se rassure quant au retard du père parce qu'il a *probablement* manqué son train, quand un historien déclare *probables* les motifs qu'il invoque pour le surprenant abandon de l'assaut sur Malte en juin 1942, ou quand un prévisionniste de la météo, prudent plus qu'en moyenne, prophétise qu'il fera *probablement* beau durant les trois jours à venir. Mais pour d'autres, la probabilité n'est nullement un élément subjectif, évalué par chacun selon ses dispositions du moment, mais un élément purement logique, qui – s'il est correctement calculé – doit s'imposer à tout sujet disposant d'un même corpus d'information. L'exemple le plus clair de ce sens purement logique se trouve dans les efforts de ceux qui cherchent à déterminer la *probabilité* d'une théorie donnée, compte tenu de toutes les informations disponibles. En ce sens, la théorie des probabilités est, pour beaucoup, un élément indispensable de la logique de l'induction.

Il n'existe aucune unanimité sur les qualificatifs à employer pour distinguer ces divers sens. On pourrait peut-être utiliser un système de dénnotations qui, à la fois, soit affectivement neutre – chose importante dans un domaine où, curieusement, les passions s'éveillent à la moindre occasion – et soulage la mémoire. On nommerait la structure formelle des probabilités d'après son analogué le plus élémentaire, les structures dites «ultra-simples» chères aux manuels scolaires: *p. numerus combinationum* (PNC). En hommage à de Moivre, on rappellerait que la mesure des propensions à l'occurrence se réalise le plus simplement dans les jeux de hasard: *p. mensura sortis* (PMS). Pour ce qui est du sens épistémique subjectif, faisant revivre un terme qui servit beaucoup voici quelque six siècles, on parlerait de *p. latitudo fidei* (PLF). Enfin, pour le sens épistémique logique, se souvenant que Carnap fait de cette probabilité-là une inférence partielle, on parlerait de *p. vis illationis* (PVI). Quoi qu'il en soit de ces notations, les distinctions qu'elles rappellent sont d'importance capitale. La confusion qui a régné longtemps dans ce domaine est bien mise en lumière par l'anecdote que voici. Jeffreys, un tenant des plus décidés de la probabilité épistémique [2], reproche à un contradicteur

(L. Ellis) d'avoir écrit: «plus j'y pense, moins je puis douter qu'un événement plus probable arrive plus souvent»; dans ce cas, dit Jeffreys, que l'on vous présente un dé qui a toute apparence d'être symétrique; vous attribuerez à chaque point de la même probabilité; donc, à long terme, chaque point arrivera – selon vous – avec la même fréquence, même si, à votre insu, le dé est pipé. Il faut se dominer pour ne pas accuser Jeffreys de mauvaise foi: les deux opposants prennent manifestement «probabilité» en des sens divers, mais Jeffreys est totalement incapable de se rendre compte que d'autres sens sont possibles que celui qu'il admet.

2. DE MENSURA SORTIS

Pour ce qui est de la théorie physique des phénomènes fortuits, l'épistémologie n'a quasiment rien à dire de spécifique: elle s'analyse comme toute autre théorie physique, ni plus ni moins. Pourtant, trois remarques valent d'être faites.

Il est certain que le statut épistémologique de cette théorie s'est grandement clarifié lorsque Popper [3] a proposé de manière explicite la notion de «propension», donnant ainsi droit de cité aux dispositions dans le domaine physique. L'historien des sciences sait que Popper était loin d'être le premier à interpréter ainsi la notion de probabilité MS. Déjà au 17^{ème} siècle, Huyghens et, avec lui, J. Bernoulli parlent, à propos d'événements fortuits, de «facilitas» ou «proclivitas» (ainsi, J. Bernoulli [4], définissant ce qui serait pour nous le «modèle formel» du jeu de dé, écrit «sex esse jactus diversos, quorum utrumvis æque in procilivi esse»). Plus tard, alors même que la doctrine subjectiviste de Laplace tenait le haut du pavé, Laurent [5] écrivait «le calcul des probabilités a pour but la mesure des chances qui président à l'arrivée des événements soumis au hasard ... leur chance est susceptible de plus ou de moins, et il y a lieu de chercher une mesure précise de cette chance». La citation la plus ancienne, de ce point de vue, est sans doute aussi la plus surprenante, et il est bien regrettable qu'elle soit tombée dans l'oubli; elle est de Thomas d'Aquin: «de futuris contingentibus secundum se consideratis non potest esse scientia; sed cum in suis causis considerantur potest de eis scientia esse, prout aliquæ scientiæ cognoscunt esse inclinationes quasdam ad tales effectus» [6].

«Inclinatio», «proclivitas», «propensity», ces termes expriment tous une même idée ; mais le mérite de Popper n'en est nullement diminué : en donnant à cette interprétation l'appui de son autorité (encore accru par le fait qu'il s'agissait bel et bien d'une conversion ; voir [7]), il l'a fait revivre et l'a solidement implantée. (La valeur épistémologique de cette notion est bien mise en évidence par le passage déjà cité de Ellis [2b] : sa discussion dans [1b] est peu claire et ambiguë, parce qu'elle ne tient pas compte de l'aspect tendanciel de la vue du monde qu'Ellis exprime à cet endroit : «... dans une longue série d'essais, ... une *tendance* à se produire avec une fréquence proportionnelle à sa probabilité» ; l'anglais «tendency» est rendu par l'allemand «Neigung», mais cette indication bien explicite est négligée dans l'analyse.)

Il est clair, aussi, que la théorie PMS doit être tenue pour vide par ceux qui tiennent le monde physique pour une machinerie rigide déterminée, où le fortuit n'a aucune place, mais que, à l'inverse, ceux qui admettent l'existence du hasard n'ont pas, de ce fait même, à contester la validité des interprétations épistémiques ; pourtant, il en est souvent ainsi, et Popper lui-même en est un exemple [8].

Nombreux sont les ouvrages élémentaires sur la théorie des probabilités dont les auteurs croient devoir avertir leurs lecteurs qu'ils vont avoir affaire à une «nouvelle logique», à un «mode de raisonner bien différent de la déduction mathématique», un mode dont les conclusions sont partiellement incertaines. Le tort fait aux débutants par l'énoncé de cette erreur est difficilement appréciable ; sans doute est-il atténué par le fait que ces auteurs raisonnent comme tout un chacun, et usent de la même mathématique que tout le monde. Il serait assurément plus indiqué d'attirer l'attention des débutants sur le fait que la théorie PMS est une «théorie physique» à l'instar de bien d'autres, et que la question si telle situation concrète est de fait l'application de telle partie de la théorie se tranche, comme partout, par la comparaison de valeurs théoriques calculées et de nombres observés – en l'occurrence, de probabilités et de fréquences. Un événement peut se produire ou non, c'est évident ; et il en est ainsi que sa probabilité soit grande, moyenne, ou petite ; son occurrence (ou sa non-occurrence) ne prouve donc rien, c'est vrai ; mais conclure de là que la théorie PMS est en dehors des normes de la Logica major est un sophisme pur et simple.

3. DE LATITUDINE FIDEI

3.1. *La pesée des opinions*

Sous sa forme latine, «probabilité» se trouve en abondance dans certains écrits philosophiques du Moyen Âge, en relation avec l'*opinion* (contredistinguée de la science proprement dite); une opinion peut être dite probable à cause des «autorités» qui la mentionnent et l'approuvent; ou à cause des arguments, dialectiques plutôt que démonstratifs, que l'on peut faire valoir en sa faveur; ou enfin, de façon plutôt péjorative, précisément parce qu'elle n'est pas démontrable mais «seulement probable». Au 17^{ème} siècle, le «probabilisme» (et ses diverses nuances) défendu par les casuistes a provoqué les *Lettres provinciales* de Pascal; en 1870, Newman, dans *A grammar of assent*, use en abondance de «probable opinion»; de nos jours encore, ce sens-là du terme est parfaitement vivant – c'est même peut-être le seul sens vivant dans «l'usage de la tribu»: une proposition quelconque reçoit, de la part d'un individu donné, une certaine approbation, une certaine acceptation, qui toutefois n'est pas celle que l'on accorde aux propositions connues avec certitude. Il est bien entendu que chacun se forme sa propre probabilité: même à égalité d'information, *ceteris paribus* en quelque sorte, les probabilités d'opinion peuvent différer de personne à personne. Ainsi, lorsque au soir du 4 juin 1944 Eisenhower réunit ses commandants inter-armes pour éclairer son propre jugement avant la décision qu'il devait prendre, il y avait là présentes des personnes également bien au fait des plans de l'opération projetée, également conscientes des avantages ou inconvénients d'un succès, d'un échec, ou d'une remise, et qui venaient d'entendre ensemble le même rapport du service météorologique; la question qu'il leur posa, quels qu'en fussent les termes, revenait à leur demander quelle probabilité avait, pour chacun d'eux, la proposition «le débarquement réussira»; or, Montgomery était pour, Leigh-Mallory était nettement contre, et Tedder hésitait: c'est un exemple typique de probabilité d'opinion, chaque personne ayant la sienne, même à égalité d'information préalable.

Cet exemple montre d'ailleurs clairement que la probabilité d'une opinion peut être plus ou moins forte, allant de la quasi-certaine au doute nettement caractérisé. Il est donc normal, la science moderne

ayant pris le tour quantitatif que l'on sait, que l'on ait cherché à quantifier la probabilité des opinions. On a un exemple pratique d'une telle quantification dans la méthode parfois utilisée pour noter les examens par questionnaires à choix multiples: on ne se borne pas à contraindre la «victime» à choisir une (et une seule) des réponses qui lui sont proposées, afin de la récompenser si elle choisit la bonne réponse, de la pénaliser sinon; on l'oblige à indiquer en outre quelle probabilité elle attribue à cette réponse, sur une échelle numérique (p. ex.: [4], je suis absolument certain de ma réponse; [3], je crois fort bien que ma réponse est correcte; [2], j'ai des doutes sérieux; [1], je ne suis vraiment pas certain du tout). Les adversaires de cette façon de faire objectent que rien ne garantit que cette échelle a la même signification pour tous (tel étudiant, bon connaisseur de l'histoire de France, très cultivé, mais méticuleux, notera [3] sa réponse «Blanche de Castille était la mère de St Louis» alors que tel autre, aventureux et tête en l'air, mais confiant en son étoile, notera [3], lui aussi, sa réponse «Louis X mourut en 1270 au siège de St Jean d'Acre»). Cette objection met en effet le doigt sur une difficulté tout à fait réelle.

Étant admis que, pour chaque personne, le degré de croyance qu'elle accorde à telle ou telle proposition est quantifiable, il convient tout d'abord de fixer, par des conventions ad hoc, son échelle de mesure. On trouve tout naturel d'accorder la note 1 à la certitude:

« $PLF_X(h) = 1$ » veut dire «X regarde h comme certaine».

On pourrait trouver tout aussi naturel d'accorder la note 0 au doute intégral et la note -1 à la certitude de la contradictoire; mais on a choisi une autre convention:

« $PLF_X(h) = 0$ » veut dire «X regarde **Nh** comme certaine».

En outre, bien entendu,

« $PLF_X(h_1) > PLF_X(h_2)$ »

veut dire que, pour X, la proposition h_1 entraîne davantage l'assentiment que la proposition h_2 .

Reste, ces points fixes étant conventionnellement choisis, à assurer l'objectivité de l'échelle de mesure dans son ensemble. Ramsey et de

Finetti ont mis à l'honneur, sur ce point, la méthode des paris [9]. Soit à déterminer quelle probabilité la personne X attribue, hic et nunc, à la proposition h . On lui propose une série de paris : elle-même miserait a F contre un adversaire qui en miserait A ; si, vérification faite, h était reconnue vraie, X emporterait tout l'enjeu, $a + A$; sinon, X perdrait sa mise, a , au profit de l'adversaire (il se peut que la vérification en cause soit fort difficile, voire impossible ; peu importe : il s'agit non de paris véritables mais de paris seulement envisagés) ; soit alors α tel que

«le pari est accepté (par X) si $a < \alpha$,
et refusé si $a > \alpha$ » ;

c'est qu'alors, pour X,

$$PLF(h) = \frac{\alpha}{\alpha + A}.$$

Il convient d'introduire une hypothèse quant au comportement de X : X n'accepte jamais un système de paris où il est sûr de perdre

[un tel système serait, p.ex., celui-ci :

sur la vérité de h : a contre A } avec $a > B$
sur la vérité de $\neg h$: b contre B } et $b > A$;

en effet, quelle que soit l'issue de la vérification, X perd soit $a - B$ soit $b - A$].

Un tel X est dit *cohérent*. On peut alors démontrer ceci :

si $\mathcal{H}$ est un ensemble de propositions, clos pour les opérations de la logique propositionnelle (**N**, **&**, **V**) et si X attribue à chaque élément de $\mathcal{H}$ une probabilité (définissant ainsi une application, PLF_X , de $\mathcal{H}$ dans $[0;1]$), alors

«X est cohérent» équivaut à « PLF_X est une probabilité PNC».

Ainsi, le raccord est fait avec une structure mathématique bien connue, et les théoriciens ont tout lieu d'être satisfaits.

3,2 Les magiciens

3,21. Qu'arrive-t-il lorsqu'un déterministe use de la théorie des probabilités ? Il adopte, bien entendu, un sens épistémique ; pour la plupart des auteurs, il s'agit du sens subjectif, PLF. En voici, parmi bien d'autres, trois exemples :

- *P. Mansion* [10], qui enseigne à Gand vers 1910 :

«tous les événements qui se produisent dans la nature sont soumis à des lois fixes et immuables. ... une immense quantité d'événements sont produits par des causes trop compliquées ... L'arrivée de ces événements est vulgairement attribuée au *hasard*, bien qu'il n'y ait d'autre hasard que l'ignorance où nous sommes des vraies causes des événements ... Ainsi, la probabilité morale est la raison que l'on a de croire ou de ne pas croire à l'arrivée d'un événement futur, ou à son existence s'il est passé ... Les géomètres ont cherché à représenter la probabilité morale par une expression numérique ... à laquelle ils ont donné le nom de probabilité mathématique.»

- *P. Lévy* [11], le plus brillant probabiliste d'Europe occidentale dans les années 1920 :

«on dit qu'un cas est *plus probable* qu'un autre si l'on a des raisons de s'attendre au premier de ces cas plutôt qu'au second ... Ainsi, à la roulette, la position d'arrêt de la boule pourrait être prévue si l'on connaissait exactement sa vitesse ... mais la diminution de cette vitesse n'est pas perceptible pour nous ... cela suffit pour que les cas possibles nous apparaissent équivalents».

- *A. Jacquard* [12], généticien contemporain :

«que recouvre ce mot, *probabilité* ? Il correspond à notre degré de confiance dans la réalisation d'un événement. ... Une épreuve est probabilisable si, compte tenu des opinions ou des informations que nous avons sur les conditions de son déroulement, nous sommes en mesure d'attacher à chacun de ses résultats possibles un nombre qui traduit le niveau de confiance dans sa réalisation».

De toute évidence, il s'agit là, chaque fois, d'une «pesée des opinions», d'une probabilité PLF. Chaque fois, l'auteur admet qu'elle se mathématise en une structure PNC, sans pour autant invoquer

explicitement les résultats formels de Ramsey et de Finetti.

3,22. Or, il est bien clair qu'une opinion à propos de tel ou tel aspect de la réalité concrète n'influence pas, comme telle, cette réalité même; seuls les magiciens pensent la contraire. Et, de fait, on a vu ci-dessus comment Jeffreys, en subjectiviste passionné mais néanmoins logique, refuse d'admettre qu'un événement est «plus fréquent» pour la seule raison qu'il est ressenti comme «plus probable». Mais très rares sont les probabilistes subjectivistes qui restent ainsi dans les limites de la saine raison: la tentation de mettre en évidence les résultats pratiques de la théorie des probabilités est quasiment irrésistible, et que seraient des résultats «pratiques» qui n'iraient pas au-delà des opinions? Pour d'aucuns, le passage de l'un à l'autre confine à l'illogisme pur et simple; ainsi, Jacquard, à la fin de son livre, a totalement changé d'avis (op.cit, p. 102): «cette probabilité est une caractéristique de l'épreuve, indépendante de l'observateur; elle est 'objective' mais elle est inconnue» et (p. 123) «cette croyance en la possibilité d'un accord prouve que nous attribuons à la probabilité en question une existence objective indépendante de l'observateur».

[Il est digne de note que ce dernier argument serait catégoriquement rejeté par de Finetti: il ne s'agit nullement, dirait-il, d'objectivité, mais d'un simple accord inter-subjectif: si une pièce de monnaie paraît symétrique et régulière, il est normal – encore que non logiquement contraignant – que la plupart des observateurs considèrent les deux résultats possibles comme «échangeables» et leur attribuent, en conséquence, à tous deux la même probabilité, $\frac{1}{2}$; ce n'en est pas moins une probabilité purement subjective.]

Pour d'autres, on est bien contraint de parler de *magie*, au sens plein du mot. Ainsi, Mansion (op.cit. p. 137) écrit: «Les probabilités de la vie humaine ont donné lieu aux spéculations par lesquelles on se procure, pour le présent ou pour l'avenir, des avantages pécuniaires» ou encore (p. 278), après avoir calculé, pour le jeu de baccarat, les probabilités (subjectives, puisque ce sont les seules!) des résultats possibles, «les bénéfices du banquier sont très considérables ... au moins équivalents à ceux d'un usurier qui prêterait les mises à 1 % par minute». De même, Lévy, après avoir fort clairement redit les limites qu'impose la logique, dans ce domaine (op. cit. p. 34):

«ces deux cas nous apparaissent comme aussi probables l'un que l'autre, parce que nous ne savons faire entre eux aucune distinction ; mais cela n'est qu'un effet de notre ignorance, et nous ne devons attribuer aucune valeur pratique aux conséquences que nous pourrions être tentés d'en déduire»

s'efforce pourtant de montrer «qu'un tel scepticisme serait dénué de fondement et que le *hasard* obéit effectivement aux lois que nous avons découvertes» : le hasard obéit donc aux lois de notre pensée, et même de notre ignorance ; «je néglige une probabilité (PLF [H.B.]) de l'ordre du millième ; il n'est pas déraisonnable d'agir ainsi, à condition de ne le faire qu'une fois ; si je le fais souvent, il arrivera quelques fois que j'aie tort» (p. 24) ; dans le même sens (pp. 31-33), «la loi des grands nombres précise la signification de l'énoncé : on a une chance sur q de se tromper ... cela veut dire qu'on se trompera en effet environ une fois sur q ... le hasard a ses lois ; il y obéit avec un certain caprice, mais à ce caprice on peut assigner des limites» ; ou encore (p. 43) «nous cherchons, par une étude aussi objective que possible, à montrer que notre idée subjective s'adapte bien à la réalité (quelle réalité ? une réalité PMS ? [H.B.]) ... 'L'on ne saurait concevoir' : cela est encore subjectif ... mais si cette conclusion s'impose à nous après une étude objective des conditions d'une expérience, c'est sans doute qu'il n'y a réellement [sic !] aucune raison de favoriser l'un des cas au détriment de l'autre».

3,23. Tout cela est déplorable. Que la théorie PLF soit logiquement sans défaut, on peut aisément l'accepter ; qu'elle ait un objet réel et précis (le «poids» des opinions) est sans doute vrai. Mais en partant d'opinions subjectives après nié l'existence de propriétés objectives correspondantes («le hasard n'existe pas», il n'y a donc pas non plus de propensions que l'on pourrait mesurer, ou même invoquer) on ne peut aboutir qu'à des conclusions subjectives. J'estime, *subjectivement*, «pile» et «face» également probables : alors, en vertu de la loi des grands nombres, et à peine d'incohérence, je dois aussi estimer, *subjectivement*, que, sur une longue série de coups à venir, de gros écarts de fréquences sont «peu probables». Il n'en résulte nullement que, dans la réalité des faits, «pile» et «face» se produiront également souvent. S'ils le font, et si c'est là plus qu'une simple coïncidence

(mais comment le savoir ?), c'est que mon estimation était «justifiée» : le seul sens possible de cette affirmation, c'est qu'il existe, dans la réalité des choses, une caractéristique objective que mon estimation cerne de façon adéquate. Le subjectiviste «conscient et organisé» récuse cette caractéristique ; c'est son droit ; mais si, alors, il ne s'en tient pas à des conclusions aussi purement subjectives que ses prémisses, il n'est rien de mieux qu'un magicien.

3,24. Les plus remarquables de ces magiciens sont les théoriciens de la mécanique statistique et ceux de la physique quantique.

Les premiers se représentent (p.ex.) un gaz comme un ensemble de particules obéissant strictement aux lois rigides de la mécanique newtonnienne (équations de Lagrange, théorème de conservation de Liouville) mais étudient les propriétés de leurs mouvements à l'aide de notions probabilistes. Dans un environnement aussi rigide et déterministe, ils ne peuvent prendre «probabilité» qu'en un sens subjectif. Ainsi, pour eux, la pression qu'exerce un gaz est la valeur moyenne de l'impulsion communiquée à la paroi lors d'un choc. Une «valeur moyenne» est un paramètre descriptif d'une loi de probabilité, et une probabilité est un degré d'ignorance ; ainsi, la pression d'un gaz n'est pas, comme on aurait pu le croire, une grandeur physique, mais un paramètre descriptif d'un état d'esprit !

Parmi les seconds, certains se préoccupent essentiellement de formalisme, et emploient «probabilité» sans paraître trop y penser (encore que l'ensemble des sous-espaces d'un espace de Hilbert ne soit pas une tribu, alors que l'ensemble des événements observables, au sens vulgaire du mot, en soit nécessairement une). Mais un grand nombre, qui suivent servilement Bohr, présentent les applications de ce formalisme en faisant de la probabilité une simple mesure de l'ignorance (soit que, à l'instar de Jauch et Piron, il fassent de cette interprétation subjective le pilier de la théorie, soit que, comme la plupart, il ne reculent pas devant une solide dose d'inconséquence). Les enseignants procèdent quasi tous de cette façon, et les admonestations de ceux que chagrine cette magie [13] restent sans effet : les étudiants ainsi formés s'en vont à leur tour répétant les mêmes aberrations, de génération en génération.

3,25. Il est pour le moins curieux qu'aucun de ces magiciens ne

semble s'être avisé d'une porte de sortie qui, à tout le moins, leur éviterait le ridicule de leur position actuelle. Ils maintiennent que le hasard n'existe pas, que l'univers tout entier est une machinerie rigide ment déterminée: soit, chacun a le droit d'avoir tort, pourvu qu'il reste logique. Pour eux, donc, la théorie PMS est vide de toute application, elle est (non pas fausse mais) sans objet. Ils pourraient pourtant invoquer le principe que voici:

lorsque les lois d'un phénomène sont suffisamment compliquées, tout se passe comme s'il n'y avait pas de lois

(principe qu'ils pourraient motiver, heuristiquement, par des exemples: la génération sur ordinateurs, par des calculs d'une complication suffisante, de longues séries de nombres «pseudoaléatoires»; ou les institutions communautaires de la Belgique); et dès lors ils pourraient dire: voici une certaine classe de phénomènes; elle a ses lois, mais nous les ignorons (sauf peut-être pour le noyau essentiel); cette ignorance même est le signe de leur extrême complication; dès lors, la théorie des probabilités s'applique à ces phénomènes (bien qu'elle soit, en droit, dénuée d'objet) à titre d'approximation *anémique* d'une situation *polyplaconomique*. Cette façon de faire aurait, à tout le moins, le mérite de sauvegarder l'emploi des probabilistes (surtout si le régime politique en venait à imposer le déterminisme, ce qui n'est pas tellement «improbable»!).

3,3. *Décideurs cohérents*

3,31. L'usage des paris dans l'objectivation des degrés de croyance, quelque jolie performance qu'il constitue, n'est pourtant pas exempt d'inconvénients. En premier lieu, il n'est pas du tout certain que l'on réagit en présence de paris supposés de la même façon qu'en présence de paris réels: ce défaut est irrémédiable, il faut bien qu'on s'en accommode. En second lieu, la probabilité LF déduite d'un système de paris ne dépend que du rapport des enjeux, puisque

$$PLF = a/(a + A);$$

or, cela est contraire au bon sens le plus élémentaire: tel qui accepterait sans barguigner de risquer 100 F pour gagner 1000 F

n'accepterait peut-être pas de risquer 100 000 F (dans les mêmes conditions) pour gagner même 2 000 000 F, parce que la perte de 100 F lui serait indifférente alors que la perte de 100 000 F lui serait insupportable (déjà au 18ème siècle, D. Bernoulli avait émis l'idée que l'utilité d'une somme d'argent est proportionnelle à son logarithme). Ainsi, l'idée apparemment bénigne d'avoir recours à la méthode des paris a conduit à modifier complètement le statut de la théorie PLF, pour en faire une théorie du «comportement du décideur cohérent». En bref, elle suppose qu'une personne donnée, X, doit choisir UNE des diverses actions possibles: $A_1, A_2, \dots, A_r$, sachant que «l'état du monde» peut être soit E_1 , ou E_2 , ..., ou E_s , et que

l'action A_i
lorsque l'état du monde est E_j
a pour conséquence $C_{i,j}$.

Mais l'état du monde est inconnu, et X doit se fixer sa propre «matrice des probabilités subjectives»

$$p_{i,j} = PLF_X(C_{i,j}) \quad [\sum_j p_{i,j} = 1] ;$$

en outre, X doit se fixer sa propre échelle d'utilités, U_X , et donc la «matrice des utilités subjectives»

$$v_{i,j} = U_X(C_{i,j}) .$$

Diverses considérations de cohérence, analogues en esprit à la proscription de tout système de paris où la perte est certaine (et qui font intervenir des «conditions de convexité» sur lesquelles l'accord n'est pas unanime) conduisent à la conclusion que voici:

à peine d'être *incohérent*

X doit choisir celle des actions A_i qui rend maximum la valeur moyenne (*subjective*) de l'utilité (*subjective*) de la conséquence qui en résulte:

A_i est choisie
si et seulement si

$$\forall k \neq j; \quad \sum_{j=1}^S p_{i,j} v_{i,j} > \sum_{j=1}^S p_{k,j} v_{k,j} .$$

3,32. De toute évidence, on est ainsi sorti du domaine de la connaissance pour entrer dans le domaine de la pratique, et l'épistémologie n'a plus grand'chose à dire à ce sujet. Il faut pourtant remarquer que (même en acceptant sans restrictions les diverses hypothèses de la théorie) on n'obtient ainsi qu'une règle de comportement cohérent, et non une règle de comportement efficace. Et aussi, plus généralement, qu'il semble bien impossible de définir ce qu'est un comportement efficace en régime d'incertitude (c'est l'origine des probabilités $p_{i,j}$) non répétitive (la répétition ayant le double avantage de permettre et l'apprentissage et l'évaluation globale du comportement); dans un tel cas, en effet, une règle de comportement pourrait fort bien être efficace et conduire pourtant à l'échec

(ce fut sans doute cette idée qui amena le Sénat romain à garder sa confiance à Paul Emile après la bataille de Cannes)

tout comme un succès peut fort bien résulter d'une règle inefficace

(c'est sans doute ce que pensait Churchill lorsque – fort injustement – il limogea Auchinleck après l'éclatant succès de la bataille défensive d'El Alamein [1.7.42] et le demi-succès de la bataille offensive qui la suivit [26.7.42]; voyez [13]).

Mais l'esprit magicien souffle ici aussi, à un point tel que cette théorie est couramment enseignée dans les Hautes Écoles de commerce et d'administration. Il est vrai que, dans ce cas, une interprétation moins désastreuse est possible (indépendamment du fait que dans l'administration la cohérence compte bien plus que l'efficacité): un bon décideur est celui qui, *souvent*, attribue une grande probabilité (subjective) à l'état de fait effectivement réalisé. Mais cette interprétation n'a un sens qu'en régime suffisamment répétitif pour que «souvent» puisse avoir un sens précis; et on peut sérieusement douter qu'alors les $p_{i,j}$ soient encore d'authentiques probabilités subjectives.

4. DE VI ILLATIONIS

4,1. *Logique du probable*

4,11. En dépit de variations dans l'expression, les logiciens – Aris-

tote, Leibniz, Frege, Hilbert, Gentzen, Tarski, etc. – font de la *logique*, pour l'essentiel, l'étude des procédures conservatrices de vérité. Il est immédiatement évident que les plus communes de ces procédures – p.ex. le syllogisme, ou la «dédution naturelle» de Gentzen – sont «épiphaniques» et nullement «créatrices»; elles n'augmentent pas le «contenu de vérité» des systèmes de propositions qu'elles forment, elles le rendent seulement plus manifeste. N'existe-t-il donc pas de procédures à la fois conservatrices de vérité et amplificatrices de contenu? de procédures qui permettent à la connaissance une progression authentique et néanmoins parfaitement sûre? Il n'était pas besoin des analyses de Hume pour établir que la réponse est NON, sans le moindre doute possible; mais après Hume nul ne peut même plus se permettre de feindre d'en douter.

Eppur, si muove: la science, c'est incontestable, avance. Pour d'autres, la coexistence de ces deux conclusions, également irréfragables, est le signe que la logique déductive n'est pas le seul outil de la connaissance scientifique, ni même de sa seule partie rationnelle, et qu'en conséquence une autre logique est nécessaire, moins ambitieuse que la logique déductive classique (mais plus utile, sans doute): LA LOGIQUE DU PROBABLE. Alors que la logique classique sacrifie la nouveauté à la sécurité, la logique du probable, au prix d'une partie *soigneusement calculée* de la sécurité, cherche à faire à la nouveauté la place que, sans conteste, le mouvement de la science montre qui lui revient.

4,12. La figure essentielle de la logique classique est le principe de détachement connu sous le nom de *modus ponens*:

$$\begin{array}{ccc}
 . & . & . \\
 . & . & . \\
 d & \Rightarrow & h \\
 . & . & . \\
 d & & \\
 \hline
 h
 \end{array}$$

Il conviendrait donc d'adjoindre au connecteur propositionnel $\Rightarrow$ un

connecteur plus faible – en d'autres termes, de compléter l'opération logique «implication» par une opération plus ample, une *implication partielle*, dont la signification intuitive serait celle-ci: la donnée de la proposition d ne permet pas de poser h comme proposition vraie, mais seulement comme proposition «probable», à un degré plus ou moins fort:

si d, alors (h au degré p)

ou, en signes, $d \Rightarrow_p h$.

Mais l'habitude s'est établie de noter cela sous la forme

$$P(h, d) = p$$

(lu: «la probabilité de h, si d, vaut p) que nous écrirons, au bénéfice d'une clarté plus explicite, $PVI(h, d) = p$ (1)

[l'ordre d'énonciation «h si d» expliquant pourquoi h est le premier argument de P, et d le second, alors que l'on écrit d avant h dans $d \Rightarrow h$].

Il est donc bien clair que

(1) est une **relation purement logique**
entre **deux propositions**.

Les cas extrêmes de (1):

$$PVI(h, d) = 1 \qquad PVI(h, d) = 0$$

sont respectivement synonymes de

$$d \Rightarrow h \qquad d \Rightarrow Nh.$$

4,13. Une logique qui mettrait en œuvre, en toute rigueur, une telle notion d'implication partielle serait une doctrine d'une valeur inestimable; les services qu'elle rendrait, dans tous les domaines, seraient extraordinaires; ainsi, à titre d'exemples:

- dans le domaine de la justice pénale, notant h la proposition «l'accusé est coupable» et d la conjonction des énoncés du code pénal et de ceux du plumeitif d'audience, la logique du probable permettrait de calculer, sans erreurs ni variations, $PVI(h, d)$; il suffirait alors que la loi ait prescrit, une fois pour toutes, la conduite à tenir (acquitte-

ment ou condamnation plus ou moins sévère) en fonction du nombre ainsi obtenu

- dans le domaine militaire, tout chef d'armée s'appêtant à livrer bataille confierait à la section ad hoc de son état-major le soin de calculer $PVI(h, d)$, h étant la proposition «la bataille sera gagnée» et d la conjonction des informations sur sa propre armée (fournie par son troisième bureau et ses commandements subordonnés) et des informations sur l'adversaire (fournies par son deuxième bureau); l'autorité militaire suprême préciserait une fois pour toutes, dans le règlement général sur la conduite des armées, ce qu'il convient de faire en fonction de la probabilité ainsi obtenue (ou encore, si divers plans sont en concurrence, de faire le calcul pour chacun d'eux, lui-même n'ayant plus qu'à mettre en œuvre le plan qui offre la plus grande probabilité de succès).

- dans le domaine scientifique, l'évaluation des hypothèses et des théories deviendrait enfin claire et efficace: h étant l'hypothèse en question (ou la conjonction des propositions qui expriment la théorie) et d la conjonction des protocoles d'expériences disponibles, on calculerait $PVI(h, d)$ (dont on observerait la variation à mesure que croît le contenu de d).

4,14. Peut-être n'est-il pas surprenant que ces mirifiques programmes n'aient jamais existé: ce serait vraiment trop beau! Mais il y a là, pourtant, un authentique paradoxe: l'idée même d'une probabilité VI paraît bien s'imposer, les principes de sa mise en œuvre ne sont guère discutables, l'intérêt (pratique ou scientifique) des résultats espérés est énorme, et le poids de la littérature sur ce sujet est plus que respectable. Pour quelle raison n'en résulte-t-il en fait *rien*? On ne peut guère le voir qu'en examinant d'un peu près la technique de cette théorie. On peut cependant faire d'emblée quelques remarques.

On explique – ou on excuse – la stérilité pratique de la théorie par l'extrême complication des données à prendre en compte. Il est vrai que l'étendue des faits dans un grand procès d'assises, ou la préparation d'une bataille, ou l'évaluation d'une théorie complexe – la relativité, p.ex. – est énorme à couper le souffle. Mais pourquoi personne n'a-t-il jamais essayé de traiter un cas simple, comme un procès facile de correctionnelle, ou bien, dans le cadre d'une théorie générale qui ne serait pas mise en cause, le choix entre quelques

hypothèses précises relatives à une question controversée mais bien délimitée (en somme, une théorie PVI de l'*experimentum crucis*)?

On dit aussi qu'il ne suffit pas de prendre comme donnée une proposition (supposée vraie, évidemment) car l'information dont on dispose est, bien souvent, elle-même seulement probable; ainsi, le deuxième bureau d'un chef d'armée pourrait estimer à $\frac{3}{4}$ la probabilité que le gros des blindés adverses est opposé à sa propre aile gauche. Mais cette objection ne tient pas: une telle information «probable» résulte elle-même de données de fait exprimées par des propositions vraies (l'on a entendu, pendant deux nuits de suite, des bruits de chenillés de telle intensité, et les photos aériennes font apparaître tels et tels objets à tels et tels endroits).

Enfin, on fait remarquer que la logique mène à tout, à condition, très littéralement, d'en sortir: elle a besoin d'un principe de détachement et, dans le cas de la logique du probable, un tel principe n'existe pas. Cette dernière affirmation est exacte; en voici un exemple. On note, respectivement, V_u et $A_{x,y}$ les propositions «M.X vivra encore au moins u ans» et «M.X est âgé d'entre x et y ans»; on a, d'après les tables de mortalité,

$$P(V_{25}, A_{30,35}) = 0,841 \quad P(V_{25}, A_{30,40}) = 0,763;$$

en outre, M. X est âgé de 32 ans, donc $A_{30,35}$ est vraie; donc, par détachement,

$$P(V_{25}) = 0,841; \tag{2a}$$

mais $A_{30,40}$ est vraie aussi, donc

$$P(V_{25}) = 0,763 \tag{2b}$$

Les énoncés (2a) et (2b) sont contradictoires! On a tenté de contrer cette objection en imposant un «principe de l'information totale» qui, dans le cas actuel, interdirait d'utiliser $A_{30,40}$ si on sait que la proposition plus précise $A_{30,35}$ est vraie. Cette réponse serait peut-être valable s'il s'agissait de règles de «comportement raisonnable», mais il s'agit ici de logique, et de rien d'autre. La vraie réponse est beaucoup plus triste: les énoncés (2a) et (2b) [que nous devrions écrire $PVI(V_{25}) = 0,841$ $PVI(V_{25}) = 0,763$] sont dénués de sens, puisque PVI est une relation entre DEUX propositions: la logique PVI n'a pas de «porte de sortie»! Aussi n'était-il pas question,

dans les deux premiers exemples, «d'en sortir» par la seule logique : il existait, là, de règles *externes* de comportement ; et, dans le troisième exemple, la logique se suffisait à elle-même, il n'était pas question d'en sortir le moins du monde.

4.2. Le formalisme PVI

4.21. La construction la mieux achevée d'une logique du probable [15] suppose donnés

un ensemble dénombrable d'*individus*, $\mathcal{E}$,

un ensemble fini de *caractères* : $\mathcal{C} = \{C_1, \dots, C_m\}$

pour chaque caractère, un ensemble dénombrable de modalités, appelées *attributs* (l'ensemble $\mathcal{A}_i = \{A_{i,1}, A_{i,2}, \dots, A_{i,n} \dots\}$ pour le caractère C_i).

Il est entendu que chaque $\mathcal{A}_i$ est une partition de $\mathcal{E}$ (autrement dit, tout individu est «porteur» d'un et un seul attribut de chaque caractère).

On appelle *modèle* toute fonction de $\mathcal{E}$ dans l'ensemble-produit des $\mathcal{A}_i$ (tout modèle est donc un «état possible du monde») :

$$z \left| \begin{array}{l} \mathcal{E} \rightarrow \prod_{i=1}^m \mathcal{A}_i \\ x \rightarrow \parallel A_{1,i}, \dots, A_{m,i_m} \parallel = \parallel z_1(x) \dots z_m(x) \parallel \end{array} \right.$$

et on note $\mathcal{Z}$ l'ensemble des modèles.

Toute *proposition* est un ensemble de modèles : une *proposition atomique* est du type

$$P(x; i, j) : \text{«l'individu } x \text{ porte l'attribut } A_{i,j}\text{»}$$

c'est à dire

$$P(x; i, j) = \{z \in \mathcal{Z} \mid z_i(x) = A_{i,j}\}$$

et les propositions *admissibles* correspondent aux éléments de la tribu engendrée par les propositions atomiques (en gros : les assemblages dénombrables de propositions atomiques et de connecteurs logiques **N**, **&**, **V**). L'ensemble des propositions admissibles est noté $\mathcal{P}$.

4.22. On démontre alors ceci :

si une fonction K , définie sur $\mathcal{P} \times \mathcal{P}$ et à valeurs réelles

$$K \left| \begin{array}{l} \mathcal{P} \times \mathcal{P} \rightarrow \mathbb{R} \\ (h, d) \rightarrow K(h, d) \end{array} \right.$$

jouit des propriétés suivantes :

$$\forall h, d; K(h, d) \geq 0 \quad (A1)$$

$$\forall d; K(d, d) = 1 \quad (A2)$$

$$\forall h, d; K(h, d) + K(\bar{h}, d) = 1 \quad (A3)$$

$$\forall h', h'', d; K(h' \& h'', d) = K(h', d) \cdot K(h'', d \& h') \quad (A4)$$

alors, il existe une application de $\mathcal{P}$ dans $[0; 1]$ – disons Φ – qui est une probabilité NC, et K est la probabilité conditionnelle déduite de Φ par contraction à son second argument; dans le cas élémentaire, $K(h, d) = \Phi(h \& d)/\Phi(d)$.

4,23. Ce résultat est remarquable. *Malheureusement*, il admet une réciproque: toute mesure probabiliste définie sur $\mathcal{P}$ définit, par conditionnement, une fonction définie sur $\mathcal{P} \times \mathcal{P}$ et jouissant des propriétés (A1), ..., (A4). En d'autres termes,

la difficulté n'est pas, étant donnés les ensembles $\mathcal{E}, \mathcal{C}, \mathcal{A}_i$ qui décrivent «le monde», de définir sur $\mathcal{P} \times \mathcal{P}$ une fonction PVI (notée ci-dessus K)

mais bien plutôt d'en choisir UNE parmi l'immensité de celles qui sont possibles.

Tout bien réfléchi, on se retrouve exactement au point où l'on objectait aux aristotéliens dégénérés de la fin du Moyen-Âge que la ratiocination logique, laissée à elle-même, ne fournit aucun résultat réellement utilisable; comme on l'a dit, c'était trop beau pour être vrai.

4,24. Bien entendu, les logiciens du probable ne s'avouent pas vaincus pour autant; il cherchent à préciser les conditions imposées à K afin de rétrécir les possibilités de choix de Φ .

Ils exigent, par exemple, que $K(h, d) = 1$ ne puisse avoir lieu *que* si $d \Rightarrow h$ (condition «de régularité»); cela est, en effet, intuitivement acceptable.

Ils exigent aussi que $K(h, d)$ ne dépende en aucune manière des individus ni des caractères qui ne sont mentionnés ni dans d ni dans h (condition «d'invariance»); cela aussi est intuitivement acceptable (mais ne le serait plus si «caractères» y était remplacé par «attributs»).

La condition «de symétrie» s'exprime ainsi: si g est une permutation de $\mathcal{E}$ qui ne touche qu'un nombre fini d'individus, h est une proposition, et h_g se déduit de h en y remplaçant toute occurrence d'un individu, disons x , par $g(x)$, alors

$$\Phi(h) = \Phi(h_g) .$$

À première vue, cette condition paraît exagérément restrictive car elle équivaut au postulat «tous les individus sont échangeables entre eux». Mais on fait remarquer qu'elle porte sur la fonction Φ , qui est un élément, en quelque sorte, entièrement a priori, et que, a priori, les individus ne sont, en effet, distingués en aucune manière: toute distinction entre individus s'exprime par des propriétés du modèle. Il en est que cette argumentation convainc, en effet.

Enfin, on impose aussi à Φ d'être telle que l'apprentissage par l'expérience soit possible, de la façon que voici:

si les individus x et y sont distincts, alors

$$PVI(P(x; i, j), d \ \& \ P(y; i, j)) \geq PVI(P(x; i, j), d) .$$

Il est vrai que l'expérience semble nous montrer que le monde est ainsi fait que l'apprentissage y est possible; mais est-ce bien là le type d'argument valable pour un logicien?

[on ne sera sans doute pas étonné d'apprendre que Wittgenstein a proposé l'emploi d'une fonction Φ qui ne satisfait pas à cette exigence].

4,25. Hélas, même si on accepte ces restrictions a priori, l'indétermination de Φ reste énorme. On démontre en effet que, sous ces conditions, on a (i étant fixé)

si d est la conjonction de

s propositions du type $P(x_k; i, A_{i, j_k})$

parmi lesquelles s_1 ont $j = 1, \dots, s_n$ ont $j = n, \dots$

et de propositions où n'intervient pas le caractère C_i

alors, si y est distinct des x_k (eux-mêmes 2 à 2 distincts),

$$K(P(y; i, A_{i,j}), d) = \frac{s_j + \lambda \cdot \gamma_j}{s + \lambda}$$

λ étant un réel non autrement déterminé et γ_j un coefficient lié d'une certaine façon à l'étendue de l'attribut $A_{i,j}$ («étendue» ayant un sens passablement vague ; si p.ex. il s'agit du caractère «couleur», avec les attributs

rouge, jaune, vert, bleu clair, bleu moyen, bleu foncé, chacun des trois derniers a une étendue moindre que chacun des trois premiers).

L'indétermination reste donc fort grande (un paramètre réel pour chaque caractère). D'autres auteurs ont même proposé, à partir d'hypothèses un peu différentes, une indétermination encore plus grande (2 paramètres réels pour chaque caractère ; voir [16]).

4,26. La conclusion paraît donc claire :

- l'idée d'une probabilité PVI est attrayante et, à première vue, parfaitement saine
- sa mise en œuvre effective n'a été essayée que dans un cadre plutôt étrié, qui devrait pouvoir servir dans les situations usuelles de la pratique mais qui est sans doute insuffisant pour «juger» d'une théorie scientifique
- cette mise en œuvre met en évidence que la probabilité logique PVI, qui par nature devrait s'imposer à chacun et donc être univoquement déterminée, est entachée en fait d'un degré d'arbitraire tel que son utilisation effective est impossible, même dans des cas relativement simples.

Nul, semble-t-il, n'a jamais défendu l'idée d'une «machine logique» qui permettrait d'inventer de nouvelles théories scientifiques ; mais les théoriciens de la probabilité VI maintiennent, logiques avec eux-mêmes, la possibilité d'une machine logique qui calculerait le degré de confirmation d'une hypothèse ou d'une théorie, compte tenu des observations disponibles [17]. Le moins qu'on puisse leur demander est d'en donner un exemple explicite : le fait est qu'ils n'en ont jamais donné un seul. Au vu de l'analyse ci-avant, on peut bien douter qu'ils le feront jamais.

4,3. *Paris et fréquences*

4,31. Les axiomes du formalisme PVI, ainsi que les propriétés supplémentaires invoquées pour restreindre autant que faire se peut l'indétermination de la fonction Φ , devraient, en principe, être justifiés de la même façon que les fondements de la logique déductive classique. En fait, leur seule justification tant soit peu explicite est qu'ils sont intuitivement satisfaisants et que leurs conséquences paraissent parfaitement acceptables. On est un peu étonné, toutefois, de voir parfois invoquer, pour cette justification, des systèmes de paris analogues à ceux qu'emploient les subjectivistes. On se demande en effet en quoi la logique, même inductive, peut être invoquée pour déclarer que (si l'on décide de parier) il est «rationnellement contraignant» d'accepter l'enjeu a contre A pour la vérité d'une proposition X dès lors que $PVI(X, d) > a/(a + A)$. Il y a là un élément (d'ailleurs non élucidé) de pragmatique dont l'intrusion dans une question de logique n'est explicable que par le désir de justifier à tout prix des axiomes dont on ne sent que trop bien la fragilité.

À propos du pari, la position des subjectivistes est assurément plus solide que celle des logiciens. Par contre, ceux-ci échappent au reproche de magie: dans la mesure où une relation logique (en particulier, une implication partielle) est quelque chose d'intrinsèque, il n'est pas anormal de décider d'un comportement à partir d'une donnée d (supposée vraie) et d'une relation $PVI(h, d) = p$. On peut même accepter que de telles règles de comportement sont «rationnelles» (encore qu'elles dussent l'être davantage si la relation PVI était moins entachée d'arbitraire); mais dans quelle mesure sont-elles «efficaces»? cette dernière question a-t-elle même un sens?

4,38. Il est bien digne de remarque que les théories épistémiques des probabilités (LF ou VI), qui sont, en droit, applicables à toute situation d'incertitude, ne sont en fait appliquées, par ceux qui les proposent et les défendent, qu'aux situations que l'on appelle aléatoires, celles où l'on définit une probabilité MS quand on croit à l'existence du hasard physique. En particulier, le classique problème de l'extraction de billes d'une urne est pris comme paradigme par tous les probabilistes «épistémiques», sous deux formes au moins.

4,321. Voici une urne, dont la composition est connue : N billes, dont B sont bleues et M mauves ($B + M = N$) : cet énoncé est la donnée, d . Voici la proposition h : «la bille que l'on va extraire sera bleue». Combien vaut $PVI(h, d)$? La réponse généralement fournie est celle-ci : pourvue que la donnée soit bien limitée à l'énoncé d ci-dessus – en d'autres termes : pourvue que *rien* ne soit dit du mode d'extraction –, $PVI(h, d) = B/N$ (4). Cette réponse (dont la justification «logique» est des plus entortillée) est extrêmement surprenante. D'une part, en effet, on veut se garder d'une description explicite du mode d'extraction, car il faudrait l'incorporer à d , et le caractère «purent logique» de (4) serait quasiment impossible à préserver. D'autre part, chacun sait combien il est malaisé de réaliser en pratique un échantillonnage représentatif; mais la réponse (4) revient à dire qu'il suffit de ne rien savoir pour que tout soit bien : cette imitation des autruches est absurde.

4,322. Voici une urne, qui contient N billes, B bleues et M mauves, B et M étant *inconnus*. On extrait n billes de l'urne, dont b sont bleues et m mauves : tout cela constitue d . La proposition h énonce une certaine valeur pour B . Combien vaut $PVI(h, d)$? La théorie des méthodes λ (parag. 4,25 ci avant) fournit comme réponse

$$PVI(B = b, d) = (b + \lambda)/(n + \lambda),$$

λ étant quelconque. Non seulement elle est indéterminée, mais elle est intuitivement déraisonnable (car PVI ne peut alors pas être plus petite que b/n); en outre, aucune valeur n'est attribuée aux nombres $PVI(B = k, d)$ lorsque $k \neq b$. C'est pourquoi les avocats de ces méthodes dans la pratique statistique traitent ce problème au moyen du théorème de Bayes, en partant d'une «probabilité a priori» [la fonction Φ]

$$PVI(B = k, \Phi) = 1/N \quad (k = 1, \dots, N)$$

qui, prétendent-ils, est la représentation *logique* de l'ignorance que décrit d . Ce dernier énoncé n'est rien d'autre qu'un postulat entièrement arbitraire. En outre, l'objection faite ci-dessus subsiste dans ce cas-ci : le mode d'extraction n'étant pas décrit, rien ne permet en vérité d'attribuer une quelconque probabilité VI à un résultat de cette extraction. La théorie des probabilités MS y arrive parce qu'elle accepte de conjecturer un modèle théorique de la situation, préala-

blement décrite avec suffisamment de détails. Elle fait donc appel à des faits, ce qui lui permet de conclure à d'autres faits, au risque de se tromper. Drapée dans sa pureté logique, la théorie PVI prétend tout déduire sans conjecture ni risque d'erreur: elle ne produit que des mots.

Université de l'État à Liège

H. BRENLY

RÉFÉRENCES

- [1a] Carnap, R.: *Logical foundations of probability*. Chicago, 1950.
- [1b] Carnap, R., Stegmüller, W.: *Induktive Logik und Wahrscheinlichkeit*. Wien, 1959.
- [2a] Jeffreys, H.: *Theory of probability*. Oxford, 1939 (304).
- [2b] Ellis, L.: On the foundations of the theory of probabilities. *Trans. Camb. Phil. Soc.*, 8, 1844 (2).
- [3] Popper, K.: The propensity interpretation of probability. *Brit. J. Phil. Sc.*, 10, 1959 (25).
- [4] Jacobi Bernoulli *Ars conjectandi* ... Basileæ MDCCXIII (7,20).
- [5] Laurent, H.: *Traité du calcul des probabilités*. Paris, 1873 (35).
- [6] Thomas Aquinas: In Aristotelis librum De memoria et reminiscencia, lect. 1 (éd. Pirotta, n° 305).
- [7] Popper, K.: *The logic of scientific discovery*. London, 1959 (173 n° 1, 361).
- [8] Popper, K.: The propensity interpretation of the calculus of probability and the quantum theory. In: Körner, S. (ed.): *Observation and interpretation in the philosophy of physics*. London, 1957 (88ss).
- [9a] Ramsey, F.: Truth and probability (1926). In: Kyburg, H., Smokler H.: *Studies in subjective probability*. New York, 1963 (63).
- [9b] de Finetti, B.: La prévision, ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré*, VII/1 (1937).
- [10] Mansion, P.: *Leçons de calcul des probabilités*. Gand, 1916 (1).
- [11] Jacquard, A.: *Les probabilités* (Coll. «Que sais-je» 1571). Paris, 1974 (11,14).
- [13a] Landé, A.: *From dualism to unity in quantum physics*. Cambridge, 1960.
- [13b] Popper, K.: Quantum mechanics without «the observer». In: M. Bunge (ed): *Quantum theory and reality*. Berlin, 1967 (7-44).
- [13c] Bunge, M.: *Philosophie de la physique*. Paris, 1975 (114-160).
- [14] Parkinson, R.: *The Auk, Victor at Alamein*. London, 1977 (199-211).
- [15a] Carnap, R.: *Logical foundations of probability* (2d ed.) Chicago, 1962.
- [15b] Carnap, R., Jeffrey, R.: *Studies in inductive logic and probability*. Berkeley, 1971 (33-165).
- [16] Hintikka, J.: A two-dimensional continuum of inductive methods. In: Hintikka, J., Suppes P.: *Aspects of inductive logic*. Amsterdam, 1966 (113-132).
- [17] Carnap, R.: *Les fondements philosophiques de la physique*. Paris, 1973 (40).

- [18a] von Mises, R.: *Wahrscheinlichkeit. Statistik, und Wahrheit*. Berlin, 1928.
- [18b] Ville, J.: *Étude critique de la notion de collectif*. Paris, 1939 (138-140)
- [19a] Fréchet, M., Halbwachs, M.: *Le calcul des probabilités à la portée de tous*. Paris, 1924
- [19b] Cramér, H.: *Mathematical methods of statistics*. Princeton, 1946 (137-150).
- [19c] Feller, W.: *Probability theory and its applications, I*. New-York, 1950 (1-7).