23 B. RIEMANN, Sur les Hypothèses qui servent de fondement à la Géométrie, 1854 ; tr. fr. J. Houël, 1870 ; rééd. dans les oeuvres mathématiques de Riemann, Blanchard, 1968.

24 G. SACCHERI, Euclides ab omni naevo vindicatus, Milan, 1733 ; rééd. et tr. anglaise Halsted, Chicago, 1920 ; rééd. American Mathematical Society, 1986.

25 P. TANNERY, Sur l'authenticité des axiomes d'Euclide, Bulletin des Sciences Mathématiques (2), t.8 1884, 162-175.

Pour en savoir plus :

Sur les preuves mathématiques :

N. EFIMOV, Géométrie supérieure, 1978 ; trad. fr. éd. Mir, Moscou, 1981.

Sur les détails de la préhistoire des géométries non euclidiennes :

J-C. PONT, L'aventure des parallèles, Ed. Lang, Berne, 1986.

REMARQUES SUR L'ESSAI DE BAYES EN VUE DE RESOUDRE UN PROBLEME DE LA DOCTRINE DES CHANCES

Jean Pierre CLERO

I.R.E.M. de Rouen

Le nom de Bayes n'est pas seulement associé aujourd'hui à un théorème que tous les probabilistes connaissent comme *"théorème ou principe de Bayes"* -qui d'ailleurs, sous la forme où il est présenté, est plutôt un principe de Laplace[1]- ; il est sous l'aspect d'un adjectif (*"bayesien"*, *"bayesienne"*) souvent lié à une façon d'envisager les probabilités et les statistiques ; et même, plus généralement, selon l'expression de Mme Fagot-Largeault à *"une philosophie de l'entreprise humaine de rationalité (théorique et pratique)"*[2]. C'est ainsi que M. Granger consacre les dernières pages de son livre La Philosophie du Style, qui examine diverses façons fondamentales de faire la science, à la conception bayesienne dans laquelle il voit *"une interprétation de la connaissance comme travail, en nous écartant de l'interprétation qui en ferait le dévoilement d'un être-en-soi caché"*[3].

On assiste depuis l'entre-deux guerres à une sorte de réhabilitation de la méthode bayesienne en calcul des probabilités et en statistiques, après un rejet presque unanime au début du siècle, faisant suite à une mise en doute commencée cinquante ans auparavant. Le nom de Bayes est devenu une sorte de lieu de ralliement des positions dites *"subjectivistes"* en statistiques et en probabilités ; la probabilité devant être regardée comme une *opinion*, comme une *croyance* -dont on peut calculer et discuter le bien-fondé- mais non comme de d'une quelconque forme d'existence objective[4]. On découvre -ou plutôt : on redécouvre- des domaines de prédilection de la conception bayesienne de l'activité pratique et théorique en matière de diagnostic médical, de prévisions en litique, en économie, d'apprentissage par expérience en psychologie, ... bref dans tous les domaines où l'on se trouve devant l'évaluation de cas ou d'événements singuliers -dont on nous répète depuis Aristote qu'il n'y a pas de science[5]-.

Ce mouvement qui s'est peut-être amplifié récemment est assez important (sur le plan quantitatif du volume des publications et sur celui des enjeux idéologiques) pour que la curiosité du philosophe et de l'historien soit alertée, donnant l'envie de regarder l'Essai que Bayes a écrit et qui semble devoir jouer un rôle dans un courant qui se réclame d'un presbytérien du XVIII ème siècle.

Mon propos est très délimité : il consiste à essayer de situer l'origine du problème de Bayes et de sa solution mathématique par rapport à divers courants de pensée du XVIII ème siècle -je veux parler d'un courant nominaliste, d'un courant sceptique, de courants newtoniens et anti-newtoniens, de prises de positions théistes ou déistes concernant la Providence- tous venant confluer dans cette oeuvre mathématique. Mais ce n'est pas parce qu'une oeuvre est mathématique et qu'elle ne contient explicitement aucune ligne de philosophie qu'elle échappe aux conflits idéologiques du temps. Nous voudrions -situant l'Essai de Bayes sur un axe défini, à une extrémité par la philosophie de Berkeley que notre auteur n'ignorait peut-être pas, à l'autre par la philosophie critique de Kant qui a radicalement ignoré l'oeuvre de Bayes- mener notre réflexion sur quatre points :

- la définition de la probabilité et, de manière plus générale, les problèmes du langage posés par l'Essai.
- la prise en compte du temps par les mathématiques bayesiennes.
- ce qu'il en résulte pour la conception de la cause qui se trouve au centre de polémiques au XVIII ème siècle.
- Nous terminerons enfin par la conception de l'activité scientifique que la trentaine de pages de l'Essai semblent supposer.

I. QUESTIONS BIOGRAPHIQUES ET BIBLIOGRAPHIQUES ; PRESENTATION ET EXPLICATION DU PROBLEME DE L'ESSAI DE BAYES.

Il nous faut présenter tout d'abord le personnage de Bayes -ce qui sera vite fait, puisque nous n'en connaissons pas grand chose- ; le problème de l'Essai et le style de sa solution.

Lorsqu'on veut remonter aux origines du courant dit *"bayesien"*, on aboutit à l'Essai en vue de résoudre un problème de la Doctrine des Chances publié en 1763 (alors que son auteur est mort depuis deux ans le 17 avril 1761) dans les Philosophical Transactions de la Royal Society par les soins de Richard Price qui était l'ami de Bayes et qu'il nous faudra également présenter en raison de son implantation, explicite cette fois, dans les courants philosophiques du temps. Cette publication de 1763 a été complétée, toujours grâce à Price et dans les mêmes Philosophical Transactions en 1765, par quelques précisions concernant la démonstration d'une des Règles qui terminent l'Essai[6]. Bayes n'a donc rien publié par lui-même sur la question qui nous intéresse ; cette décision ou indécision de l'auteur ne nous semble pas fortuite : nous essayerons d'en déceler les raisons par l'examen même de l'oeuvre.

Disons tout de suite que, si certains auteurs probabilistes contemporains traitent Bayes en véritable pionnier d'une nouvelle rationalité, la publication des Philosophical Transactions, en dépit des efforts de Price pour souli-

gner l'intérêt du texte en mathématiques, en physique et en philosophie, n'a suscité immédiatement aucune espèce d'attention ; il faut attendre plus d'une quinzaine d'années pour voir les recherches de Bayes prises en compte par Condorcet qui les cite en 1781 dans un texte de l'Histoire de l'Académie Royale des Sciences de Paris[7] et par D'Alembert en 1780[8]. Laplace, qui publie en 1774 son Mémoire sur la Probabilité des Causes par les Evénements, ne semble pas les avoir connues. Il est donc arrivé à ce texte mathématique à peu près la même mésaventure que celle qui est survenue, dans le même courant de pensée d'ailleurs, au philosophe David Hume, une trentaine d'années auparavant[9]. Si l'on prend l'oeuvre de Kant, à la fin du XVIII ème siècle, comme une sorte de lieu de vigilance où viennent se réfléchir les diverses attitudes scientifiques -car l'attitude kantienne se veut critique et, à tort ou à raison, la moins dogmatique possible à l'égard des sciences-, on s'aperçoit qu'il y a en elle une connaissance au moins pratique de la loi des grands nombres de Jacques Bernoulli[10], mais rien sur l'attitude bayesienne pourtant si originale à l'égard du temps, de la cause et plus généralement de la raison dans son activité de connaissance. Kant n'a certainement pas connu le travail de Bayes ou l'a estimé indigne de l'investigation critique.

Mais qui était Bayes ?

Il est né en 1702 ; il reçut une éducation privée, comme il était ordinaire chez les non-conformistes de l'époque. Il n'est pas impossible qu'il ait appris les mathématiques par l'un des fondateurs de la théorie des probabilités, A. de Moivre. Fils d'un Ministre de l'Eglise Presbytérienne à Leather Lane, près de Holborne, il devint lui-même l'un des six Ministres non-conformistes publiquement ordonnés comme tels en Angleterre. En 1742, Bayes a été élu Membre de la Société Royale, soutenu par Stanhope, Martin Folkes, J. Burrow, Cromwell Mortimer et John Eames.

On sait peu de chose sur son amitié avec Price. Les deux hommes se sont probablement rencontrés grâce à John Eames, ami de Newton par lequel il avait été introduit à la Royal Society, qui enseigna les mathématiques à Price à Moorfields ; Bayes avait aussi probablement étudié par ses soins les mathématiques à l'Académie dissidente de Tenter Alley.

On sait presque aussi peu de chose sur son oeuvre.

Il est peut-être l'auteur d'une Défense des Mathématiciens contre les objections de l'auteur de l'Analyste publiée anonymement en 1736 sous le titre Une Introduction à la Doctrine des Fluxions ; du moins la Bodleian Library d'Oxford l'assure-t-elle. Il s'agit d'une réponse aux attaques du philosophe Berkeley qui, en 1734, avait rudement mis en cause, non pas la vérité des résultats obtenus par Newton dans son Traité de la Quadrature des Courbes (de 1676, publié en 1704), mais la rigueur des méthodes qui y mènent[11]. La Défense, présentant une argumentation originale qui met en avant les droits du temps et du devenir en mathématiques, nous semblerait bien porter pour cette raison la marque de Bayes dont nous verrons le souci particulier du temps dans l'Essai.

Il est aussi possible qu'il soit l'auteur, à moins que ce ne soit son père Joshua Bayes, d'un essai sur la Providence intitulé De la Divine Bienveillance ou un Essai de prouver que la fin principale de la divine Providence et du Gouvernement divin est le Bonheur de ses Créatures (publié en 1731). En dépit de l'anonymat de la publication, la Bodleian Library assure que Thomas Bayes en est l'auteur.

Il est, cette fois certainement, l'auteur d'une lettre à John Canton (Secrétaire de la Royal Society) sur les séries asymptotiques, publiée après sa mort en 1763 dans les Philosophical Transactions, l'année même de la parution de l'Essai.

Quant à Price, né en 1723, mort en 1791, il fut Ministre presbytérien à Newington Green et partagea son temps entre ses obligations de prêcheur et ses études de théologie, de mathématiques et de philosophie. En chacun de ces domaines, il marquera la vie intellectuelle de son temps. Il entrera en polémique avec Hume sur des questions liées au fondement de la morale qu'il envisage dans un intuitionnisme rationnel -alors que Hume cherchait à évacuer la raison des principes de la morale ; sur des questions liées d'une part à la Providence, d'autre part aux miracles -dont Hume avait contesté la possibilité, arguant du petit degré de probabilité de leur existence comparé à celui des lois de la philosophie naturelle-. Donc Price connaît les idées de Hume sur la cause et en apprécie la valeur polémique.

De plus, Price est un grand connaisseur du calcul des probabilités, assez pour être expert des questions d'assurances, conseiller d'une grande société ; et même pour être expert des questions financières, en particulier des problèmes d'annuité, sur lesquels il écrivit et qui l'amenèrent à conseiller des gouvernants (Shelburne et Pitt) sur les moyens de rembourser la Dette Nationale.

Il n'est pas impossible que son élection comme Fellow of the Royal Society fut liée à son travail d'édition de l'Essai de Bayes.

Il est temps de présenter le problème de Bayes et de l'expliquer.

L'Essai en vue de résoudre un problème de la Doctrine des Chances s'ouvre sur la demande suivante :

"Etant donné le nombre de fois qu'un événement inconnu s'est réalisé ou a fait défaut (has happened or failed), on demande la chance (chance) que la probabilité (probability) de sa réalisation lors d'une seule épreuve (in a single trial) soit comprise entre deux degrés quelconques (between any two degrees of probability) que l'on puisse assigner (that can be named)".

Avant d'expliquer quelques termes de ce problème, je voudrais prendre un exemple pour faire comprendre la question proposée.

Je joue à un jeu de hasard avec un adversaire qui gagne 8 parties sur 10 qu'on a réellement jouées. Ai-je affaire à un tricheur ? Dois-je continuer à jouer s'il me le propose ? Y a-t-il plus de chances pour que j'aie affaire à un tricheur plutôt que non, sachant que s'il ne triche pas j'ai, comme lui, 1 chance sur 2 de gagner, et que, s'il triche, il a 9 chances sur 10 de gagner ?

En d'autres termes, tout se passe comme si nous disposions de deux urnes remplies chacune de boules noires et de boules blanches, mais en proportions différentes.

Dans la première, la proportion de boules blanches est de 50/100.
Dans la seconde, elle est de 90/100.

On tire de l'une des urnes (on ne sait pas laquelle) un échantillon de 10 boules, dont on examine la composition : 8 sont blanches. Quelle est la probabilité pour que l'échantillon ait été tiré de la première urne, c'est-à-dire de celle qui contient autant de boules de chaque couleur ?

La règle de Bayes donne cette probabilité.

(Plus exactement, car les choses sont un peu plus fines : je me dis que si mon adversaire triche, sa probabilité de gagner doit se situer entre 8/10 et 10/10. Il vient de gagner 8 parties sur 10 réellement effectuées ; quelles sont les chances pour que son score soit compris entre ces deux valeurs lors des parties suivantes ?).

Revenons à l'explication des termes de notre problème.

Un *"événement inconnu"* est un événement dont on ne sait s'il va se produire ou non, ou dont on ne sait s'il s'est produit ou non ; la méconnaissance ne portant pas sur la nature de l'événement, mais sur son éventualité, sur son advenir. Le problème prend donc son sens, non pas du point de vue des événements eux-mêmes, mais du point de vue de ce que l'on sait d'eux. La notion de *"to fail"* (faire défaut) engage à faire la même remarque : Bayes définit "to fail" comme ne pas pouvoir se produire parce que la place est prise par l'événement contraire ; on voit que cette notion n'a de sens que du point de vue d'un sujet qui attend ou suppute (moyennant des indices suffisants) l'existence ou l'inexistence d'un événement dont on fixe mentalement l'essence.

Désormais, c'est en opposition au problème de Bernoulli tel qu'il est traité dans la fin de la IV ème partie de l'Ars Conjectandi (publié en 1713) que nous pouvons éclairer le problème de Bayes. Bernoulli avait établi qu'en multipliant indéfiniment les observations ou les expériences d'événements de même nature, le rapport du nombre des événements favorables au nombre total des événements (favorables et manquants) approche de leur probabilité dans des limites dont l'intervalle se resserre de plus en plus et devient moindre qu'aucune quantité assignable. Le problème traité par Bayes apparaît nettement par différences : tenons-nous en à trois principales.

La première engage une conséquence majeure sur laquelle nous reviendrons. Bayes parle de *"la probabilité de la réalisation d'un événement inconnu lors d'une seule épreuve"*, qu'elle soit la première, la seconde, la troisième ou la n ème. Certes, c'est bien en conséquence du nombre de fois qu'on a vu un événement se passer ou ne pas se passer que l'on calcule cette probabilité, mais elle porte sur la réalisation (ou le manque) d'un événement à venir ou passé qui m'est inconnu. Dès lors, si la probabilité est probabilité d'un événement, elle ne peut plus sans incohérence être définie comme le rapport des cas favorables au nombre total de cas (ce qui est, sinon la définition, du moins la pratique bernoullienne de la probabilité).

La seconde différence tient à ce que le problème de Bayes garde un sens quel que soit le nombre d'informations dont on dispose sur les événements passés semblables à celui dont on examine la probabilité. Simplement, le degré de cette probabilité varie en fonction de ce nombre.

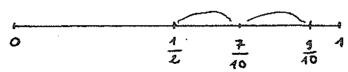
Enfin, si l'on regarde l'exact intitulé de la *"demande"* de Bayes, on remarque que l'objet du problème n'est pas la probabilité même de la réalisation d'un événement, mais *"la chance que la probabilité de sa réalisation soit comprise entre (quelques) deux degrés que l'on puisse assigner"*. Certes, la définition 6 établit que la *"chance"* et la *"probabilité"* veulent dire la même chose ; il n'en demeure pas moins que la recherche porte sur la probabilité pour que la probabilité d'un événement soit comprise entre deux valeurs.

Avant de nous lancer dans les analyses de la *probabilité*, du *temps mathématique* et de la *cause* qui nécessitent une prise en compte presque exclusive de

la I ère Section de l'ouvrage, donnons tout de même un aperçu de la solution du problème qui est proposée dans la II ème Section.

L'idée de la solution est la suivante :

1) Quand on ne sait absolument rien sur la probabilité de réalisation d'un événement, on doit supposer que la chance pour que cette probabilité se trouve entre deux valeurs quelconques est la même pourvu que l'écart entre ces deux valeurs soit conservé.



2) Mais dès que l'on a des informations, la chance n'est évidemment plus la même ; et tout le problème de Bayes consiste à calculer *"ma chance d'avoir raison lorsque, un événement s'étant produit p fois et ne s'étant pas produit q fois, en n épreuves* $(n = p + q)$, *je conjecture que la probabilité de sa production en une seule épreuve soit comprise entre deux degrés de probabilité X et x"*.

L'estimation de cette chance (d'avoir raison) est donnée par le binôme de Newton qui permet de dénombrer les combinaisons des événements réels ou manquants connus, de constituer à partir de là la fonction binomiale $y = Ea^p b^q$ (E représentant le coefficient du terme en $a^p b^q$ du binôme $(a + b)^{p+q}$, a signifiant la probabilité que l'événement se produise, b celle que l'événement ne se produise pas) et de raisonner géométriquement sur des rapports déterminés de surfaces délimitées par la courbe qui correspond à cette fonction.

Ces généralités étant posées, nous voudrions en venir à nos analyses de détail, en commençant par celle de la *"probabilité"*.

II. UNE ETRANGE DEFINITION DE LA PROBABILITE.

Nous avons déjà remarqué que, pour que le problème de Bayes ait un sens et une solution recevable, il fallait recourir à une définition que Price, qui assure la présentation philosophique du texte de son ami, qualifie de *"peculiar"* (d'étrange) ; je lis la 5 ème définition de la I ère Section de l'Essai :

"La probabilité d'un événement quelconque est le rapport entre la valeur (value) à laquelle on devrait estimer une espérance (at which an expectation depending on the happening of the event ought to be computed) dépendant de la réalisation de cet événement et la valeur (value) de la chose espérée s'il se réalise (upon it's happening)".

La probabilité n'est donc pas un rapport de nombres de cas ; elle est un rapport de valeurs (qui n'ont de sens, même si elles peuvent s'exprimer numériquement, que par l'estimation d'un sujet) : d'une part, de la valeur d'une espérance, c'est-à-dire d'un contrat qui accorde un prix si un événement se produit ; d'autre part, de la valeur du prix attaché à cet événement. Ainsi, dans l'ordre des raisons, si l'on ose dire, la notion d'*espérance* précède celle de *probabilité* qui la suppose.

On pourrait voir dans cette définition un héritage de Huygens, de la Logique de Port-Royal, de Pascal peut-être ; mais pour qu'elle soit jugée *"peculiar"* par un spécialiste des probabilités, il faut considérer les choses autrement et se souvenir de l'éventuelle polémique avec Berkeley. Nous verrons tout à l'heure que la façon de répondre à Berkeley consiste moins à rejeter l'objection qu'à infléchir l'activité scientifique de telle sorte qu'elle ne soit plus vulnérable aux attaques de l'évêque de Cloyne : ainsi Bayes voudra-t-il sauver des contradictions la méthode des fluxions par une certaine considération du temps. Bayes -si du moins il est l'auteur du texte- revendiquera la possibilité de faire de la science en tenant compte des objections berkeleyennes. Pourquoi n'en ferait-il pas de même avec le nominalisme de Berkeley ?

Plutôt que de condamner la science qui fait usage d'idées abstraites dont Berkeley montre qu'elles sont contradictoires et impensables[12], ne pourrions-nous tourner l'obstacle en essayant de faire une science qui ait un fondement nominaliste ou du moins qui sache tenir compte des objections nominalistes ? Le danger de l'attaque nominaliste, c'est le *scepticisme*, c'est la condamnation des sciences, même des plus efficaces d'entre elles comme l'est la science de Newton (dont on voit mal comment elle pourrait se passer de l'usage d'idées générales). Or la tentative de Bayes est peut-être de trouver le moyen de faire une science, donc d'éviter l'écueil sceptique, tout en prévenant l'attaque nominaliste.

Sans doute doit-on avouer que la définition bayesienne de la probabilité n'est pas des plus claires ; elle semble même enfermer quelque circularité (quel contenu donner au mot "devrait" de la définition ? Ne faudrait-il pas déjà connaître la règle que veut établir Bayes pour donner un sens à la définition de la probabilité ?). Mais c'est peut-être trop demander à une définition que d'exiger de se la pouvoir représenter[13]. De plus, que l'on réfléchisse un instant à la définition de la probabilité comme rapport du nombre des cas favorables et de celui de l'ensemble des cas quand ces nombres deviennent très grands et inaccessibles à l'expérience : ne forgeons-nous pas pour le coup une idée abstraite susceptible de tomber sous les attaques de Berkeley ? Or il semble que le postulat -à la fois philosophique et de pratique mathématique- de

Bayes soit que *seul existe l'événement singulier* et que si la probabilité doit avoir un sens, ce ne soit pas en prenant en compte un nombre immense d'événements semblables, inatteignable par l'expérience, mais que ce soit en concernant un événement à venir ou passé. *La probabilité n'a de sens qu'en assignant une valeur numérique singulière*[14] *à un événement singulier*. Sa définition comme rapport de deux valeurs singulières permet d'échapper au reproche d'abstraction.

L'idée d'un fondement nominaliste de la science peut être menée beaucoup plus loin qu'au niveau des définitions. Qu'une définition puisse être nominale n'est évidemment pas une découverte du XVIII ème siècle ! Ce qui en revanche est peut-être caractéristique d'une science qui a pris au sérieux les objections berkeleyennes du début du siècle et qui a tenté de les transformer positivement, c'est la promotion du *verbe*, avec ses divers *modes* et ses *temps grammaticaux* particuliers. Le traducteur de l'Essai de Bayes est immédiatement frappé par la richesse grammaticale et sémantique des verbes tout à fait inhabituelle dans un texte mathématique, même lorsqu'il est rédigé -ce qui est très souvent le cas au XVIII ème siècle- avec très peu de formalisme et dans la langue vernaculaire. Nous en verrons un échantillon tout à l'heure lorsque nous étudierons le temps. Sans la grammaire et son subtil système des temps, le texte mathématique s'effondrerait.

Comprenons bien ce dernier point dans ses racines berkeleyennes. La langue mathématique est classiquement formée des noms et du verbe "être" qui semblent travailler ensemble[15] ; si chez Bayes le clavier sémantique et grammatical des verbes, par lesquels on saisit les temps et les rapports de temps, est plus riche que le seul auxiliaire "être" et que le seul présent de l'indicatif, c'est peut-être une conséquence du nominalisme berkeleyen mené au bout de sa logique. La critique des idées abstraites rend très problématique le signifié des mots ; si les mots ne représentent plus les idées générales (en dépit de l'illusion que nous avons de leur existence), que deviennent les signifiés ? Ils ne sont pas autre chose que l'*usage* que nous faisons des mots[16] ; du coup, la représentation perd son caractère fondamental et n'est plus que l'un des usages -d'ailleurs ambigu- que nous faisons des mots. La contestation de la représentation est inévitablement une remise en question de la suprématie du nom dont on imagine, plus facilement que pour tout autre mot, qu'il *représente* quelque chose ; dès lors des mots qu'on ne voyait guère agir dans le discours mathématique, comme les verbes par exemple, semblent désormais pouvoir jouer un rôle capital.

Un temps de verbe ne représente rien à proprement parler ; il constitue plutôt un rapport de situations des événements entre eux et de l'observateur à l'égard de ces événements. Il constitue un univers où le calcul, la prévision etc.

peuvent prendre sens. D'autres fonctions interviennent que la stricte représentation : tous les mots ne sont plus traités comme des noms. Cela est si vrai que le nom même de "probabilité" ne se définit que par le détour d'un verbe au conditionnel[17] et que la fonction du nom est souvent occupée par des gérondifs[18].

La libération du verbe et de ses fonctions est le versant positif d'un nominalisme dont le versant négatif est le scepticisme. Nous allons, dans un instant considérer concrètement ce travail des verbes au niveau de deux propositions et de leurs démonstrations ; mais nous voudrions au préalable inscrire ces considérations dans une recherche sur le temps car une science qui cherche à faire travailler les verbes devient tout naturellement une science du temps.

III. LA PRISE EN COMPTE DU TEMPS DANS L'ESSAI DE BAYES.

Nous voudrions montrer que le temps est pris en compte par l'Essai de Bayes d'une façon très originale par rapport à son traitement newtonien en mathématiques et en mécanique. Nous trouverons plus tard le fondement de cette singulière prise en compte du temps dans une attitude -que nous pourrons qualifier de "dialectique"- du calculateur de probabilités devant la science de son époque.

Nous savons déjà que Bayes a probablement défendu les droits du temps en mathématiques contre les accusations de contradictions lancées par Berkeley à l'égard de la méthode newtonienne des fluxions.

Il faudrait -sur ce point- pénétrer brièvement quelques éléments de la querelle du calcul infinitésimal.

L'idée -brièvement rapportée- du De Quadratura Curvarum, c'est de faire décrire une surface sous une courbe dont on connaît l'équation $y = x^n$ par une ordonnée -c'est-à-dire une ligne- qui se déplace à une vitesse uniforme et de calculer à chaque instant l'accroissement de cette surface par unité de temps. Il faut que cet instant soit de grandeur suffisante pour que, x s'étant accru d'une petite quantité o, y ait à son tour changé ; le rapport des deux accroissements pouvant s'effectuer de telle façon que la simplification par o au numérateur et au dénominateur soit possible. Mais il faut aussi que ce o soit assez petit pour qu'on puisse négliger toutes ses puissances. Cela fait donc un double rôle joué consécutivement par la quantité notée "zéro" : d'une part, elle témoigne du temps nécessaire pour passer d'une étape de la description de la figure à une autre et, à ce titre, elle donne lieu à des

rapports d'accroissements naissants ; mais d'autre part, on lui demande, pour simplifier le rapport d'accroissements, de disparaître et de devenir un zéro et elle peut, à ce titre, donner lieu à des rapports de quantités évanouissantes.

Mêlée à la querelle du calcul infinitésimal, Berkeley a attaqué -c'est du moins un des arguments de l'Analyst- au nom du principe de contradiction, ce double usage de la quantité "zéro" par laquelle il est possible de simplifier le rapport d'une part, et par laquelle il est possible d'autre part, quand la quantité devient égale à zéro, de tenir les termes indésirables dans le développement du résultat de ce rapport pour nuls[19]. Les critiques berkeleyennes à l'encontre du De Quadratura Curvarum ne sont pas sans fondements logiques dès lors qu'elles refusent les deux descriptions dans le sens naissant, puis dans le sens évanouissant, et le découpage du traitement mathématique en deux moments qui permet de faire dans le second une opération contraire à celle qui est produite dans le premier. Or, parmi les défenseurs des droits du temps et du devenir en mathématiques, il n'est pas impossible que l'on trouve Th. Bayes, s'il est l'auteur -comme l'affirme la Bodleian Library- d'une Introduction to the Doctrine of Fluxions sous-titrée Defence of the Mathematicians against the objections of the Author of the Analyst.

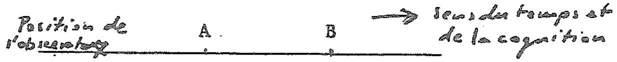
"Supposer que deux grandeurs soient égales et inégales en même temps eût été une inconséquence ; mais les supposer d'abord égales et puis, ensuite les supposer devenir inégales[20] ne présente pas l'ombre d'une difficulté. Et il n'y a rien d'extraordinaire à ce que, à partir de la supposition que ces deux choses se passent l'une après l'autre, je puisse déduire une conclusion qui ne dérive pas de l'une ou de l'autre supposition prise séparément ; la justesse de la conclusion que j'ai tirée apparaîtra à quiconque considèrera ce qui est distinctement établi à chaque étape de la démonstration"[21].

Là où Berkeley dénonce un usage du temps qui permet les paralogismes nécessaires à la production de résultats dont il ne conteste que la rigueur, l'auteur de la Defence accepte dans une démonstration mathématique des étapes, un devenir mettant en scène des grandeurs qui changent de valeur au cours du temps. La vérité qu'elle établit n'est pas indépendante du cheminement qu'elle a suivi pour se poser. Le temps de l'épanouissement (de la figure, de la fluente) n'est pas celui de l'évanouissement ; et la contradiction n'apparaît qu'à celui qui refuse le temps en mathématiques, en particulier ce double passage du temps. Mais il faut reconnaître que cette défense est philosophique et que le statut newtonien du "zéro" est parfaitement contradictoire[22].

L'idée de Bayes est peut-être de donner un statut plus méthodologique, plus mathématique, à un temps qu'il a philosophiquement défendu vingt cinq ans plus tôt. Mais alors que chez Newton le temps *"vrai et mathématique""coule uniformément"*, que la distinction d'un avenir et d'un passé est radicalement inutile, que seule est prise en compte la *"durée"*, Bayes ose dans l'<u>Essai</u> un discours mathématique qui comprenne la distinction du futur, du présent et du passé.

Il nous faut, pour nous en convaincre, lire trois propositions fondamentales de la I ère Section. Elles traitent toutes les trois des probabilités conditionnelles (ou "suppositionnelles").

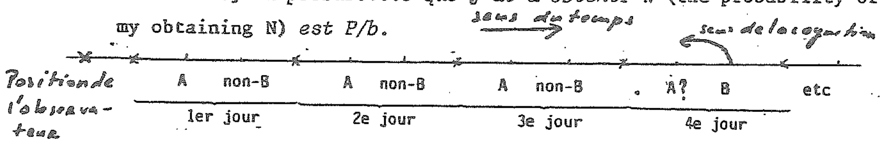
Proposition 3 : La probabilité pour que deux événements subséquents se produisent tous les deux dans l'avenir (will both happen) *est en rapport composé de la probabilité du premier et de la probabilité du second en supposant que le premier se produise.*



$$P\ (A\ \&\ B) = P\ (A) \times P\ (B \text{ si } A).$$

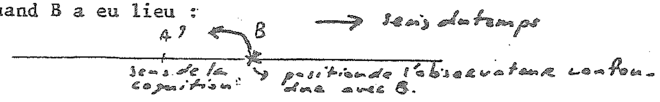
Le corollaire se contente de remarquer que $P\ (B \text{ si } A) = \dfrac{P\ (A\ \&\ B)}{P\ (A)}$

Proposition 4 : Si deux événements subséquents sont de telle sorte qu'ils puissent être déterminés chaque jour (c'est-à-dire sont tels que chaque jour, chacun ne puisse échapper à cette alternative : il a eu lieu ou il n'a pas eu lieu), *si pour chaque jour la probabilité du second est b/N et celle des deux ensemble P/N, et si je dois recevoir N à condition que les événements arrivent tous les deux dès le jour où le second événement se produit, je dis que, dans ces conditions, la probabilité que j'ai d'obtenir N* (the probability of my obtaining N) *est P/b.*



$$P(A \text{ si } B) = \frac{P(A\ \&\ B)}{P(B)}$$

Le corollaire se contente de souligner que la probabilité ne change pas quand B a eu lieu :



Proposition 5 (qui n'est autre que la conclusion de la démonstration du corollaire) : *Soient deux événements consécutifs, soit b/N la probabilité du second et soit P/N la probabilité des deux ensemble ; lors de la découverte en premier lieu du second événement, je conjecture à partir de là que le premier événement s'est aussi produit avec une probabilité que je suis en droit d'évaluer à P/b.*

Cette lecture appelle plusieurs remarques.

Le temps n'est pas la simple forme intuitive -qu'elle sera chez Kant- qui permettrait de loger de diverses façons, et pour lui donner divers sens, la même forme logique -à savoir : la formule de la probabilité conditionnelle-. En ce cas, le temps ne serait qu'une sorte d'habillage philosophique (et sans grand intérêt, avouons-le) d'une condition traitée logiquement ; le texte deviendrait un essai philosophique déguisé en caractères mathématiques.

Le temps fait intimement partie de la structure de la probabilité conditionnelle. Dans l'esprit de Bayes, cette formule n'est en aucune façon une définition comme elle est prise par exemple dans le traité de Kolmogorov sur les probabilités[23] ; elle fait l'objet de théorèmes différents qui la démontrent dans diverses orientations (de l'observateur et des événements) par rapport au temps. Kolmogorov n'a plus aucun égard à l'orientation temporelle dans laquelle on procède pour définir et calculer une probabilité conditionnelle. Du coup, ce formalisme logique peut indifféremment accompagner un intuitionnisme philosophique du temps qui prétendrait lui donner sens ; cet intuitionnisme n'intéresse en aucune façon le mathématicien comme tel.

Telle n'est pas la position bayesienne qui est plus proche, quoique étant parfaitement mathématique, de la pensée empirique du temps. Ce paradoxe, rarement souligné, que les mathématiques qui ne sont pas elles-mêmes empiriques peuvent servir la cause de l'empirisme doit être rendu parfaitement clair.

Le temps chez Bayes n'est pas détachable des événements autrement que par abstraction ; raisonner sur le passé, ce n'est pas raisonner sur un continu homogène, c'est opérer sur des événements passés. Et la démonstration de la valeur de la probabilité conditionnelle d'événements liés dans le passé ne s'effectue pas de la même façon que celle concernant les événements présents ou à venir. Il ne s'agit nullement de détacher une forme logique d'une intuition temporelle -ce qui fait à la fois le lit d'un intuitionnisme, d'un formalisme et de tous les conflits qui peuvent résulter de cette séparation- ; il s'agit de montrer au contraire que la forme des raisonnements par laquelle on établit la valeur de P (A si B) ou de P (B si A) est inséparable des événements en position d'être passés, présents ou à venir. La formule établie par Bayes est une sorte de schème du temps.

D'une certaine façon, nous pourrions dire que l'irréversibilité du temps -qui n'est aucunement prise en compte par Newton et qui sera chez Kant un caractère de la forme a priori du temps[24] (que seule donc la philosophie peut considérer)- est à l'oeuvre dans l'argumentation de cette première Section. Nous verrons tout à l'heure l'importance de cette remarque pour l'expression mathématique de la cause -que le kantisme rejette par principe-.

Voyons dans leur principe simplement les façons différentes de démontrer en direction de l'avenir et en direction du passé.

Commençons par la démonstration de la Proposition 3.

"*Supposons que si deux événements se produisent ensemble, je doive recevoir N* (I am to receive N), *que la probabilité pour que les deux événements se produisent à l'avenir ensemble soit P/N* (that the probability both will happen is P/N), *que la probabilité du premier événement futur soit a/N* (l'anglais dit simplement : that the first will is a/N), *(et par conséquent (N − a)/N celle de sa non-existence future), et que celle que le second arrivera en supposant que le premier arrive soit b/N. Alors (par la déf.5) P sera la valeur de mon espérance qui deviendra égale à b si le premier événement se produit.* (Deux remarques : 1) Nous sommes au coeur de la démonstration : il s'agit de regarder comment croît l'espérance que les deux événements se produisent quand A se produit. 2) "the value of the expectation" dont il est parlé dans le texte est un condensé de l'expression consignée dans la définition 5 : "the value at which an expectation depending upon the happening of the event ought to be computed". Autrement dit, le nom ici vaut pour une formule entière qui contient un verbe au mode conditionnel). *Par conséquent, si le premier se produit, le gain que j'obtiens par son moyen est b − P, et s'il ne se produit pas, ma perte est P. C'est pourquoi, par la précédente proposition, a/N est à (N − a)/N, autrement dit a est à N − a comme P est à b − P.* (Le reste de la démonstration nous amène par un simple calcul au résultat : P (B si A) = P (A & B) / P (A).)

La Proposition 3 utilise une propriété démontrée précédemment (en Proposition 2) selon laquelle :

$$\frac{P\ (A)}{P\ (-A)} = \frac{\text{Perte en cas de } -A}{\text{Gain en cas de A}}$$

soit : P (A) x Gain (en cas de A) = P (−A) x Perte (en cas de −A).

(On retrouve une équation qui est dans la <u>Logique de Port-Royal</u>).

L'idée de la démonstration est la suivante : gain et perte sont gain et perte d'espérance (liés à la production des événements A et B).

Dans la situation où je dois calculer P (B si A),

observateur —— A —— B ——>

j'utilise l'équation ainsi :

Probabilité que A se produise	x	Quantité d'espérance que j'ai gagnée (que B se produise si A se produit)	=	Probabilité que A ne se produise pas	x	Espérance que je risque de perdre (espérance de voir les événements A & B subséquents se produire)
$\frac{a}{N}$	x	(b − P)	=	$\frac{N-a}{N}$	x	P

De ce raisonnement, Bayes tire la valeur de P (B si A).
Le sens de cette démonstration est que mon espérance de voir B arriver si A est arrivé <u>grandit</u> quand A est arrivé.

Il eût été facile à Bayes de traiter uniformément le cas de la Prop.3 et ceux de la Prop.4, de son corollaire et de la Prop.5. Il a été pris d'un scrupule dont nous avons trop souvent perdu la clé aujourd'hui lorsque la définition des probabilités conditionnelles nous semble indépendante du fait que le temps cognitif s'écoule dans le sens du temps ou à contresens du temps réel. Certes Bayes ne donne aucune explication à ce sujet ; mais ce qui est sûr, c'est qu'il change de démonstration selon que le temps est considéré en prédiction ou en rétrodiction ; et que la Prop.4 est un effort pour ménager une médiation entre la Prop.3 et le corollaire de la Prop.4, dont la Prop.5 n'est qu'une conclusion. (Dans la Prop.4, l'observateur regarde vers le passé, mais A et B sont encore à venir, comme dans la Prop.3. Cette situation, fort étrange quand on cherche à lui donner un contenu empirique, est un intermédiaire entre la situation temporelle de la Prop.3 et la situation de la Prop.5, plus "naturelle").

Voyons de plus près la Proposition 4 qui aboutit au résultat symétrique de celui de la Proposition 3, cette fois pour le calcul de P (A si B). Lisons-la et dévoilons son principe.

"En effet, si ce n'est pas le cas, soit x/N la probabilité que j'obtienne N et soit le rapport de y à x comme celui de N – b à N. Alors, puisque x/N est la probabilité que j'obtienne N (par la définition 5), x est la valeur de mon espérance. De plus, parce que, compte tenu des conditions précédentes, j'ai, dès le premier jour l'espérance d'obtenir N si les deux événements se produisent ensemble, ce dont la probabilité est P/N, la valeur de cette espérance est P. De la même façon, si la coïncidence ne se produisait pas, j'aurais l'espérance de retourner à ma situation initiale ("Likewise, if this coincident should not happen I have an expectation of being reinstated in my former circumstances") *; en d'autres termes, j'aurais l'espérance de recevoir ce qui a x pour valeur et dépend de la non-réalisation du second événement dont la probabilité (par le corollaire de la Proposition 1) est (N – b)/N ou y/x, parce que y est à x comme N – b à N. Par conséquent, puisque x est la chose espérée et y/x la probabilité de l'obtenir, la valeur de cette espérance est y. Mais ces deux dernières espérances réunies sont évidemment identiques à mon espérance originelle, dont la valeur est x, et par conséquent, P + y = x. Mais y est à x ce que N – b est à N. De là, x est à P ce que N est à b, et x/N (la probabilité que j'ai d'obtenir N) est égal à P/b".*

L'idée est que, dès le premier jour et avant de savoir si B s'est produit, je puis analyser la situation ainsi : trois cas peuvent se présenter. Ou : A et B se réalisent tous les deux et je gagnerai N. Ou : B ne se réalise pas et je conserverai l'espérance de gagner N, dont la valeur est x. Ou : B se réalisera, mais non A, auquel cas je perdrai mon enjeu x.

Pour effectuer le calcul d'espérance de A si B, il faut se rapporter d'abord à la Prop. 1 qui suppose l'additivité de l'espérance liée à des événements incompatibles. En effet, l'espérance d'obtenir N par B ou par non-B est égale à la somme de ces espérances. L'espérance d'obtenir N (ou ce qui revient au même, celle que l'événement A se soit produit) si B a eu lieu -et l'on ne tient compte de A que s'il a eu lieu conjointement à B- est donc égale à la somme de l'espérance d'avoir A et B et à celle d'avoir A quand non-B a eu lieu. Cela donne :

$$\begin{array}{ccccc} Esp._{(A \text{ si } B)} & = & Esp._{(A \text{ et } B)} & + & Esp._{(A \text{ si non-}B)} \\ x & = & P & + & y \end{array}$$

Le calcul de y est celui de l'espérance que je puis avoir (de la réalisation de A ou de l'obtention de N) quand B ne s'est pas produit. y est une partie de x (qui représente l'espérance que j'ai de constater que A s'est réalisé quand B s'est produit) obtenue par la multiplication de x et de la probabilité d'obtenir cet x (ou de sauvegarder cet x). Cette probabilité de la non-réalisation de B est $1 - \frac{b}{N}$.

Donc : $x = P + x(1 - \frac{b}{N})$.

$$x = \frac{PN}{b}$$

D'où :

$$\frac{x}{N} = \frac{\frac{P}{N}}{\frac{b}{N}} \qquad \text{CQFD.}$$

Il y aurait beaucoup à dire sur les stratagèmes de ces démonstrations. Ainsi les schèmes grammaticaux permettent-ils des médiations entre des situations qui apparaissent fort discontinues. Par exemple, la démonstration obtenue dans le sens de l'avenir (A et B étant futurs, le sens de la cognition allant dans celui du cours du temps) joue-t-elle un rôle dans la démonstration prise dans le sens du passé, en ce que la grammaire nous rend possible de projeter un regard rétrospectif dans l'avenir (A et B étant tous deux futurs, mais la cognition allant à contresens du cours du temps). La grammaire permet des glissements que ne permettrait pas la considération abrupte des cas.

Il n'est pas impossible même que quelques particularités de la langue anglaise permettent des glissements que d'autres langues (en particulier le français) ne permettent pas. La concordance des temps ne fonctionne pas en anglais comme elle fonctionne en français ; il n'est pas impossible par exemple que le retour toujours plus prompt en anglais, après un futur ou après un conditionnel au présent de l'indicatif favorise l'embrayage du discours mathématique sur la structure grammaticale. Nous trouvons le cas dans la démonstration de la Proposition 4 ; nous pourrions en montrer de semblables dans celle de la Proposition 5.

Mais sur le point où nous commençons à voir avec intérêt des arguments chez Bayes qui permettent à l'empirisme de résister victorieusement, avec intelligence, et en sauvegardant l'accord avec les mathématiques, contre la critique qui sera celle de Kant[25] quelques années plus tard, il nous faut convenir de graves lacunes de l'Essai, probablement à l'origine de sa non-publication par l'auteur. Price n'a peut-être pas vu les points faibles de l'ouvrage ; si faibles pourtant qu'on se voit rejeté, sans l'avoir voulu, du côté de la théorie a priori et non-mathématique de la cause que l'on trouve chez Kant.

Les faiblesses tiennent à la topique du temps lorsqu'on cherche à la restituer dans sa totalité. Sur l'ensemble des cas possibles de configurations du I guess, de A et de B (soit 10 cas)[26], Bayes n'en envisage que 3, voire 4 (car la prise en compte, dans les démonstrations, d'étapes intermédiaires peut faire office d'une inspection de cas) ; laissant implicitement à chacun le soin de raisonner sur les autres cas pour trouver soit des formules identiques, soit des formules symétriques[27]. Mais le jeu de situations qu'il étudie suffit à nous laisser soupçonner que l'intention qu'il poursuit ne semble pas aboutir. Ce qu'il veut très honnêtement et très courageusement obtenir, c'est l'identité ou la symétrie des formules des probabilités conditionnelles, quel que soit le sens temporel dans lequel elles sont prises; cela non par la voie courte et arbitraire de la définition, mais par démonstrations. Or si chacune des démonstrations de la Prop.3, de la Prop.4 et de son corollaire, de la Prop.5, peut être acceptée séparément, elles sont inacceptables quand on les rapproche car elles supposent la considération de cas intermédiaires qui donnent des valeurs d'espérance et de probabilité fort différentes de celles des cas envisagés en leur entier. Le double statut de certaines étapes intermédiaires compromet le caractère exhaustif du "schématisme" bayesien. L'auteur n'a pas pu ne pas s'apercevoir de ces paralogismes ; il est vraisemblable qu'il ait abandonné l'idée de publier un essai aussi *fondamental* dans ses vues qu'*incertain dans ses résultats*.

IV. LA CAUSE ET LA LOI.

Une rapide incursion dans l'idée de cause achèvera de nous convaincre que l'Essai aurait pu être fondamental si Bayes était allé au bout de sa tâche et avait tenté de redresser les paralogismes qui barrent l'accès au but qu'il s'était fixé. L'échec bayesien devrait être médité autrement qu'en le fuyant -soit à la façon formaliste de Kolmogorov, soit par la transformation, à la manière de Kant, du principe de causalité en un *dogma* (c'est-à-dire en un principe non-mathématique)-.

Kant a distingué la causalité des principes mathématiques (issus des catégories de l'unité, de la pluralité, de la totalité, de la limite,...) ; et là encore, il nous permet paradoxalement de remarquer que les mathématiques peuvent -jusqu'à un certain point- être en affinité avec une conception empiriste de la cause. S'il est par principe impossible de mathématiser la causalité kantienne, on peut trouver chez Bayes, dans certaines limites, un équivalent mathématique de la conception humienne de la cause. Montrons-le.

Hume, quelques décennies plus tôt, avait critiqué l'idée de cause -traditionnellement envisagée comme connexion nécessaire entre les événements- dans le Traité de la Nature Humaine (1739) et dans l'Enquête sur l'Entendement Humain (1748). Price souligne d'ailleurs que le texte de Bayes *"concerne de très près la Philosophie Expérimentale par le sujet qu'il développe"* et qu'il *"doit nécessairement être considéré par quiconque voudrait clairement rendre compte de la force du raisonnement par analogie ou inductif"*. Quels sont les points com-

muns de l'Essai avec l'analyse humienne de la cause ?

L'analyse de la cause avait mené Hume à distinguer en elle l'antécédence, la contiguïté et la conjonction constante (plus ou moins parfaite) d'événements semblables (la perfection tenant au nombre de cas observés et, dans ce nombre de cas, aux cas favorables ou défavorables). La probabilité d'une subséquence, compte tenu des subséquences semblables déjà passées, ne remplit-elle pas exactement le même rôle qu'une causalité ainsi analysée ?

La causalité humienne ne prend son sens que par l'appréhension subjective des événements ressemblants ou, plus exactement, des conjonctions d'événements ressemblantes. Il en va de même dans l'Essai, les événements et leurs subséquences n'ayant de sens que *du point de vue du sujet* (du *"I guess"*) qui, moyennant ses expériences, s'efforce de conjecturer sur le compte de cas semblables. Certes, on pourra toujours objecter que Bayes ne parle jamais de *cause* ; toutefois, nous devons constater qu'il met un soin particulier dans sa Première Section à donner au problème qu'il veut résoudre une ampleur qui lui permette de valoir non seulement pour des événements singuliers semblables, mais encore pour des subséquences semblables d'événements ; ce qui donne évidemment un intérêt expérimental à son problème.

Même si la causalité s'applique à des cas singuliers (encore inédits ou dont je n'ai pas été témoin de l'édition), elle ne prend son sens que si l'on sort d'une conjonction unique pour considérer la constance de cette conjonction dans d'autres cas semblables. Le geste bayesien qui calcule la probabilité d'avoir raison en conjecturant un événement à venir (ou passé) à partir d'événements (ou de subséquences d'événements) semblables est analogue.

Nous pourrions aller jusqu'à dire que, par bien des côtés, Bayes accomplit un voeu humien, celui de transformer en nombres ces fameux coups de crayon qui, à chaque nouvelle expérience, viennent conférer une vivacité supplémentaire à ton de l'idée (d'une subséquence). L'exemple des bateaux dans le Traité de la Nature Humaine n'est pas sans analogie avec l'approche bayesienne de la probabilité[28].

Enfin -et c'est sur ce point que nous envisageons pleinement l'enjeu de cette conception de la cause-, Hume attaquait, en un sens sceptique, l'universalité et la nécessité de la loi physique (newtonienne), insistant sur le fait que la raison ne pouvait pas prouver, ni l'expérience nous assurer, que l'événement qui allait venir serait conforme aux événements semblables passés. De la même façon, rien ne garantit plus chez Bayes que l'événement futur (ou l'événement passé dont on n'a pas été témoin) sera certainement conforme aux événements semblables du passé ; il ne l'est plus qu'avec un degré de probabilité seulement, c'est-à-dire qu'avec un certain degré de croyance (sans tomber dans la crédulité d'une part, dans le dogmatisme de la loi d'autre part).

Lorsque Price écrit à Canton que la question de Bayes et sa résolution permettent d'établir *"la raison que nous avons de croire* (reason we have for believing) *qu'il y a dans la constitution des choses des lois fixes selon lesquelles les événements se produisent, et que, par conséquent, la structure du monde doit être l'effet de la Sagesse et de la Puissance d'une cause intelligente ; l'argument téléologique en faveur de l'existence de la Divinité se trouvant ainsi confirmé"*[29], il faut bien peser cette affirmation ; en particulier l'expression *"raison de croire"*.

Si la démarche de Bayes est la moins dogmatique possible et la seule réellement convaincante pour montrer qu'il y a des lois, puisqu'elle n'en présuppose pas l'existence (formant ses inductions des cas particuliers à un cas particulier, mais jamais à la loi générale), il faut dire aussi -comme s'en est bien aperçu Price- que nous ne saurons jamais par ce moyen s'il y a des lois et que nous n'aurons jamais que des *"raisons d'espérer"*[30] qu'il y en a ou que l'événement suivant existera plutôt qu'il n'existera pas.

C'est pourtant paradoxalement là que le point de rupture devient manifeste avec Hume. Price voyait dans le texte de Bayes une réponse à la conception humienne de l'induction[31]. Chez Hume, le mouvement qui me porte, à partir de la mémoire des événements semblables passés et de leur proportion de réussites et d'échecs, à inférer l'existence d'un événement semblable pour l'avenir n'est pas *rationnel*. Il est *naturel*. Il ne saurait y avoir au sens strict de *raison de croire*[32], mais une habitude qui, assumant le poids de l'expérience passée, s'ouvre vers l'avenir en influant sur notre esprit.

Or le calcul bayesien est un calcul des *raisons* de croire ou d'espérer. Du décompte des réussites et des échecs passés, je prétends tirer des raisons de croire pour l'avenir. On voit aussitôt que cette façon de penser tombe sous les coups de la critique humienne au même titre que la conception rationaliste traditionnelle de la cause ; car il n'y a pas de raison de croire, mais simple croyance que l'avenir sera conforme au passé dans sa répartition déjà expérimentée des réussites et des échecs. Certes, la rationalité humienne est en affinité avec celle du calcul des probabilités, mais la notion de *"raison de croire"* sonne comme une opposition à Hume sur le terrain de la causalité et devrait, si l'on voulait à tout prix lui donner un sens humien, au moins recevoir une interprétation *"naturelle"* (même s'il y a, dans le Traité de la Nature Humaine, des règles qui limitent l'usage du principe de causalité). Hume ne s'est nullement laissé séduire par la rationalité des probabilités qui est pourtant très proche de sa représentation du monde ; de son point de vue, les

mathématiques bayesiennes ne concernent en aucune façon le monde à venir, mais décrivent simplement notre façon de l'envisager, compte tenu de son passé; leur valeur est plus psychologique que transcendantale.

Traçons en guise de conclusion plus positivement la conception du savoir qui semble à l'oeuvre dans l'Essai.

V. UN STYLE DE RATIONALITE.

Le nominalisme qui pousse Bayes à calculer le degré de probabilité d'un cas individuel, la conception empiriste de la cause sont peut-être des efforts pour répondre positivement au scepticisme. Il ne paraît pas impossible de faire une science sans être obligé de croire aux idées générales ; on peut traiter la cause comme une subséquence et en faire tout de même un usage à la façon de Hume dans le domaine politique ou économique.

Bien entendu, on nous rappellera qu'il s'agit d'un texte mathématique auquel le présentateur seul (et non l'auteur) s'avise de donner un sens philosophique. Mais le texte de Bayes ne relève-t-il pas, par la nature de son problème et par la rédaction de sa solution, d'une science plus *réfléchissante* que déterminante ?

Les mathématiques bayesiennes ne fournissent aucun instrument pour mettre en forme des phénomènes selon des lois, comme le font les mathématiques newtoniennes dont les quadratures de courbes, par exemple, peuvent directement servir en mécanique. Bayes calcule des raisons de croire à la reproduction d'une conjonction d'événements à partir d'une certaine quantité de répétitions. Ce n'est pas une science qui infléchit la représentation que l'on se fait des choses ; elle agit au niveau de l'évaluation, de la supputation, de la pondération de la représentation. Connaître une chose, c'est savoir évaluer les raisons de croire en un jugement la concernant. Si la science de Newton relève d'une conception conquérante et déterminante à l'égard du réel (des phénomènes), la science de Bayes est une réflexion "en retrait" de la détermination. Elle tend à substituer à la loi objective la règle du croire, susceptible d'une objectivité d'ailleurs, mais qui n'est pas celle des choses.

M. I. Hacking, dans The Emergence of Probability, a montré l'existence d'un courant permanent situé en marge de la *high science* qui découvre les lois depuis la Renaissance. La *low science*, qui prend en compte les probabilités, vit en rapport de réflexion, voire d'attaque sceptique, à l'égard de la haute science qui triomphe en physique notamment et elle amplit les terrains qui échappent à l'activité législatrice de cette haute science (la médecine, la biologie, l'économie, la politique).

Cette science bayesienne qui fait sienne l'inquiétude du XVIII ème siècle à l'égard des lois, de leur portée de détermination des phénomènes, de leur égale stabilité dans le passé et dans l'avenir, se fait aussi beaucoup plus *"dialectique"* que la science des lois. Nous entendons ici *"dialectique"* en un sens très voisin de son acception aristotélicienne d'activité de recherche préalable à la science ou par laquelle il est possible de discuter les thèses de cette science. La dialectique est une activité intermédiaire entre la simple opinion et la science. La science bayesienne est dialectique en ce qu'elle s'efforce de saisir des réalités que la haute science ne parvient pas à exprimer. Donnons en exemple l'irréversibilité du temps que la science bayesienne prend au moins davantage en compte que la science newtonienne qui la traite comme une sorte d'illusion. De façon plus générale, le caractère "dialectique" de cette recherche est très visible en ce qu'il serait bien difficile de citer une découverte qui ait été faite par la méthode bayesienne. Ne vit-elle pas d'une activité de médiation entre la *high science* qui parle universellement des phénomènes et le cas singulier (ou la conduite singulière à tenir) qui seul(e) intéresse le médecin, l'économiste, le politique ?

C'est pourquoi si le philosophe est un moment satisfait de voir des notions philosophiques recevoir une sorte de fonctionnement à l'intérieur de la science, il doit se demander si la science n'a pas plus à gagner d'une séparation plus nette avec la philosophie. Bayes a échoué dans sa tentative de schématisme (temporel) de la probabilité conditionnelle, peut-être pour avoir cru qu'une science réfléchissante était possible en mathématiques, soit en dehors de la philosophie.

N O T E S

(1) Ce principe donne la probabilité pour que, un événement B s'étant produit et pouvant se produire de plusieurs manières différentes (s'excluant mutuellement) par la cause A_1, A_2, ... A_n, ce soit précisément à cause de A_1 qu'il s'est produit :

$$P(A_1 \text{ si } B) = \frac{P(B \text{ si } A_1) \times P(A_1)}{\sum_1^n P(B \text{ si } A) \times P(A)}$$

Le §2 du Mémoire sur la Probabilité des causes par les événements de 1774 énonce le "Principe" en ces termes : *"Si un événement peut être produit par un nombre n de causes différentes, les probabilités de l'existence de ces causes prises de l'événement sont entre elles comme les probabilités de l'événement prises de ces causes, et la probabilité de l'existence de chacune d'elles est égale à la probabilité de l'événement prise de cette cause, divisée par la somme de toutes les probabilités de l'événement prises de chacune de ces causes".*

(2) Actes de la journée d'études du 15 décembre 1979, Médecine et Probabilités, Institut de Recherche universitaire d'Histoire de la Connaissance, Université Paris-Val de Marne, Paris, Didier-Erudition, 1982, p.79.

(3) op. cit., A. Colin, Paris, 1968, p.294.

(4) Dans la préface de son ouvrage sur la Théorie des probabilités (Theory of probability, trad. par A. Machi et A. Smith, New York, Wiley, 1974), De Finetti écrit : "LA PROBABILITE N'EXISTE PAS" et ajoute que *"la probabilité, si on la regarde comme quelque chose de douée d'une certaine forme d'existence objective, est ... un concept erroné, une tentative illusoire d'extérioriser ou de matérialiser nos véritables croyances probabilistes"*

(5) *"Il n'y a de science que du général, d'existence que du particulier"*, selon la célèbre formule d'Aristote ; il n'en demeure pas moins que *"ce n'est pas la santé que considère le médecin, mais la santé de l'homme, et peut-être même plutôt la santé de tel homme, car c'est l'individu qu'il soigne"* (Ethique à Nicomaque, I, 6, 1097 a 10).

(6) Il est difficile de dire ce que le texte publié par R. Price en 1765 (Philosophical Transactions, 54, pp.296-325) sous le titre *"A demonstration of the second rule in the Essay towards the solution of a problem in the Doctrine of Chances"* doit à Bayes.

(7) op. cit., pour 1778 (1781), pp.43-6.

(8) Opuscules mathématiques, T.VII, 1780, p.60.

(9) Hume écrit dans une de ses lettres que *"l'ouvrage tomba mort-né des mains de la presse"*.

(10) Comme en témoigne l'alinéa introductif de l'Idée d'une Histoire Universelle au point de vue cosmopolitique (1784).

(11) The Analyst ; or a discourse addressed to an infidel mathematician (1734), § 20 : *"I have no controversy about your conclusions, but only about your logic and method...... I consider the geometrical analyst as a logician, i.e. so far forth as he reasons and argues ; and his mathematical conclusions, not in themselves, but in their premises ; not as true or false, useful or insignificant, but as derived from such principles, and by such inferences".*

(12) Pour le montrer, Berkeley, dans ses Principes de la Connaissance Humaine, (Int., § 13), s'était appuyé sur un texte célèbre de Locke dans lequel le philosophe montrait que l'idée de "triangle" était impensable psychologiquement (quand je pense au triangle, il faut bien qu'il ait quelques caractéristiques : qu'il soit obliquangle, ou rectangle, ou équilatéral, ou

isocèle etc... mais pas tout cela à la fois) et monstrueuse sur le plan logique (il faut qu'un triangle soit ou rectangle ou équilatéral, par exemple, mais pas les deux à la fois). Locke avait écrit en effet dans l'Essai sur l'entendement humain (L. IV, Ch.VII, §9) : *"Ne faut-il pas de la peine et de l'habileté pour former l'idée générale de triangle : car elle ne doit être ni obliquangle, ni rectangle, ni équilatérale, ni isocèle, ni scalène ; mais à la fois tout cela et rien de tout cela ? En réalité, c'est quelque chose d'imparfait qui ne peut exister : c'est une idée où sont rassemblés des éléments de plusieurs idées différentes et incompatibles"*. Mais de façon extrêmement contradictoire, Locke n'en continuait pas moins à soutenir la thèse du rapport de représentation entre les mots et les idées générales. Berkeley va faire éclater cette contradiction et en tirer toutes les conséquences.

(13) Berkeley avait très bien dit dans les Principes de la Connaissance Humaine (Intr., §18, 1710) : *"C'est une chose que de garder constamment à un mot la même définition, autre chose que d'en faire le représentant d'une même idée : l'une est nécessaire, l'autre est inutile et irréalisable"*.

(14) Kant indique bien la synthèse singulière effectuée par le nombre quoique son usage soit général (Critique de la Raison Pure, PUF, Paris, 1963, p. 166).

(15) La théorie du verbe se trouve dans la II ème Partie de la Logique de Port Royal, au ch.II (qui reproduit le ch.XIII de la Grammaire Générale). Il s'agit de ramener tous les verbes à un seul : le verbe "être" qui relie entre eux des mots ayant fonction d'être des noms.

(16) Comme l'a montré Mme Brykman dans sa thèse, Berkeley, Philosophie et Apologétique, (Vrin, Paris, 1984).

(17) Voir la définition de la "probabilité" au début du §II.

(18) idem. Le français escamote fatalement ce mode verbal de désignation dans la langue anglaise.

(19) The Analyst, §§ 13-16.

(20) Ce qui est le cas pour la surface à laquelle on ajoute une petite quantité qui va ensuite devenir égale à zéro ; en toute rigueur, cela fait deux surfaces, tantôt égales, tantôt inégales.

(21) Op. cit., (London, J. Noon, 1736), pp.39-40.

(22) Notre sujet n'est pas ici de montrer comment sous l'impulsion de D'Alembert, ce statut sera dépassé par l'idée de "limite".

(23) Nous parlons bien entendu des Grundbegriffe des Wahrscheinlichkeitrechnung de 1933. Dans la traduction de N. Morisson, Foundations of the Theory of Probability, 2d edition, Chelsea Publishing Company, New York, 1956, voir I, §4, (5), pp.6-7.
De Finetti s'insurgera contre ce statut de définition donné à la formule des probabilités conditionnelles et pensera, comme Bayes, qu'elle doit être l'objet de théorèmes. (Voir : Probability, Induction and Statistics - The Art of Guessing-, (John Wiley & Sons, London, New York, Sydney, Toronto, 1972), p.81).

(24) Kant distingue par ce caractère le temps de l'espace -par ailleurs deux formes a priori de la sensibilité susceptibles de la même analyse-. Le philosophe critique même l'usage spatial que la physique (ou même déjà la phoronomie) fait du temps. (Voir la Critique de la Raison Pure, op. cit., p.63).

(25) Très dangereuse sous ses apparences de prudence par ses coupures du concept et de l'intuition, de la forme et de la matière ; très exclusive aussi sous ses apparences d'accueil du travail scientifique.

(26) On peut faire le diagramme suivant des configurations possibles :

A	B	B si A	A si B
Futur	Futur	A B cas de la Prop.3	A B cas de la Prop.4
Présent	Futur	A B moment de la Prop.3	A B
Passé	Futur	A B	A B
Passé	Présent	A B	A B Cas du corollaire de la Prop.4 et Prop.5.
Passé	Passé	A B	A B

Nota : La croix représente la position présente de l'observateur.

(27) Du moins avons-nous cherché à le montrer dans un article intitulé Temps et Langage dans la Première Section de l'Essai en vue de résoudre un Problème de la Doctrine des Chances de T. Bayes qui sera peut-être bientôt publié par le Centre d'Analyse et de Mathématiques Sociales de l'Ecole des Hautes Etudes.

(28) *"Supposez par exemple que j'aie trouvé à la suite d'une longue observation que, de vingt bateaux qui prennent la mer, dix-neuf seulement reviennent Supposez que je voie actuellement vingt bateaux qui sortent du port : je transfère à l'avenir mon expérience passée et me représente dix-neuf de ces bateaux comme revenant sans dommages et l'un d'eux comme périssant"*. (Traité de la Nature Humaine, trad. Leroy, Aubier-Montaigne, Paris, 1946, p.218).

(29) Pearson & Kendall, Studies in the History of Statistics and Probability, (Griffin, London, 1970, p.135).

(30) Bayes parle aussi de *"reason to reckon"*, de *"reason for thinking"*.

(31) Du moins M. Hacking l'affirme-t-il dans The Emergence of Probability (Cambridge, 1975).

(32) Ces deux mots de *"reason"* et de *"believe"* sont constamment écartés l'un de l'autre par Hume dans le domaine de la réflexion sur la cause.