

Leclerc-Olive, Michèle, Ms, Acad, Social Science, France, Paris, « *Délibération et décision : l'espace public disqualifié par le calcul économique- Une approche critique de la théorie probabiliste classique* » -I

Michèle Leclerc-Olive, CEMS-CNRS, EHESS, 54 boulevard Raspail , 75 006 Paris
Notes en vue d'une communication au colloque PEKEA, novembre 2005.

« Aux questions particulières, il faut des réponses particulières ; si la série de crises dans laquelle nous vivons depuis le début du siècle peut nous enseigner quelque chose, c'est, je crois, le simple fait qu'il n'existe pas de normes générales pour déterminer infailliblement nos jugements, ni de règles générales sous lesquelles subsumer les cas particuliers avec un certain degré de certitude. »

Hannah Arendt (1906-1975)

« La justice est une expérience de l'impossible. Une volonté, un désir, une exigence de justice dont la structure ne serait pas une expérience de l'aporie, n'aurait aucune chance d'être ce qu'elle est, à savoir juste appel de la justice. Chaque fois que les choses se passent ou se passent bien, chaque fois qu'on applique tranquillement une bonne règle à un cas particulier, à un exemple correctement subsumé, selon un jugement déterminant, le droit y trouve peut-être et parfois son compte, mais on peut être sûr que la justice n'y trouve jamais le sien. Le droit n'est pas la justice. Le droit est l'élément de calcul, et il est juste qu'il y ait du droit, mais la justice est incalculable, elle exige qu'on calcule avec de l'incalculable ; et les expériences aporétiques sont des expériences aussi improbables que nécessaires de la justice, c'est-à-dire de moments où la décision entre le juste et l'injuste n'est jamais assuré par une règle. »

Jacques Derrida, Force de loi, 1994.

Reinhard Stelten, prix Nobel d'économie en 1994, écrivait récemment : « Le principal courant théorique moderne en économie se fonde sur une image irréaliste de la prise de décision par l'homme. Les agents économiques sont décrits comme des maximisateurs bayésiens de l'utilité subjective pleinement rationnels. Cette conception de l'économie n'est pas étayée sur des preuves empiriques, mais plutôt sur l'axiomatisation de l'utilité et de la probabilité subjective. Dans le livre fondamental de Savage (1954), les axiomes sont des exigences de cohérence (consistency) des actions. On ne peut qu'admirer la structure imposante construite par Savage. Elle exerce une forte attirance intellectuelle comme concept de rationalité idéale. Toutefois, il est faux de supposer que les hommes se conforment à cet idéal¹ ».

L'autorité de la pensée économique se soutient aujourd'hui de décisions épistémologiques, principalement de la possible modélisation probabiliste de pans entiers du monde et de l'activité humaine – notamment du jugement – ainsi que de la calculabilité des décisions. La maniabilité technique de l'outil probabiliste pousse à en élargir abusivement le champ d'application. En effet, hypothèses sous-jacentes aux modèles stochastiques les plus usités (le mouvement brownien, par exemple) supposent une anthropologie implicite – rationalité, indépendance des acteurs, etc. – et postulent une possible réduction du temps au présent immédiat², bien éloignées de l'expérience quotidienne partagée. Surtout, en feignant de croire que le jugement ou la décision découlent, par pure déhiscence, de l'enquête qui les documentent, elle invite chacun à se défaire de sa compétence de citoyen au profit de l'expertise technique. Elle nous conduit à rabattre nos projets sur des anticipations élaborées par ces derniers et tente de nous convaincre que toutes nos préférences sont prédéterminées. Elle dissuade au fond

¹ G. Gigerenzer, R. Stelten, *Bounded Rationality : The Adaptive Toolbox*, Cambridge, MIT Press, 2001, p. 13. Cité par A. Berthoz, *La décision*, Odile Jacob, 2003, pp. 11-12.

² J. Baschet, « L'histoire face au présent perpétuel. Quelques remarques sur la relation passé/futur », dans F. Hartog et J. Revel (dir), *Les usages politiques du passé*, Éditions de l'EHESS, 2001. Voir aussi F. Hartog, *Régimes d'historicité. Présentisme et expériences du temps*, Seuil, 2003.

d'ouvrir un espace public pour délibérer de notre avenir. Cette pensée économique renforce le sentiment que l'espace politique est un espace « par défaut », en creux, pure condition de possibilité de l'action des acteurs de la société civile pour faire valoir des intérêts particuliers, espace colonisé par la technisation des problèmes, rétif aux pratiques d'assemblées et aux délibérations publiques, plus enclin aux négociations furtives, souvent discrètes, qui trouvent leur forme achevée dans le *lobbying* généralisé. En revanche, nous verrons que, reconnaître l'impossibilité de « calculer » l'avenir du monde, prendre acte que décider en situation d'incertitude relève du jugement ou du pari, peut contribuer à restituer à l'espace politique public sa place fondamentale. Examiner plus en détail ce qui distingue les conceptions de la démocratie centrée sur la notion de société civile de celles centrées sur la notion d'espace public, ne peut être entrepris ici. Notons seulement que les premières mettent l'accent sur la négociation et le lobbying de groupes constitués sur la base d'une identité ou d'un intérêt partagé, alors que les secondes privilégient la visibilité des actions et la délibération publique. Ces deux idéaux-types de démocratie se combinent de manière à la fois singulière et fluctuante dans nos sociétés occidentales. Ils informent de manière différente les programmes d'appui aux réformes démocratiques dans les pays du Sud³.

Le pouvoir de séduction exercé par le calcul des probabilités est tel, qu'on peut lire dans un ouvrage récent : « Si le président Kennedy s'est montré si ferme dans la crise cubaine, c'est que les ordinateurs américains, digérant des montagnes de statistiques ont affirmé qu'il y avait une probabilité considérable pour que les Russes abandonnent au dernier moment et s'ils ont effectivement abandonné, c'est probablement parce que leurs ordinateurs avaient fait les mêmes calculs et donné les mêmes résultats⁴ », alors que A. Schlesinger, ancien conseiller de J. F. Kennedy, confirmait très récemment encore « que les généraux soviétiques présents à Cuba étaient prêts à utiliser armes nucléaires tactiques, en 1961, si les États-Unis avaient attaqué. Or, c'est précisément ce que les chefs d'état-major et les responsables du renseignement conseillaient à Kennedy. JFK rejeta leurs recommandations, afin de poursuivre – et de mener à son terme – une action diplomatique⁵ ». Cette anecdote est révélatrice à plus d'un titre. La décision de JFK va à l'encontre des recommandations « calculées » de l'État major et les Russes sont prêts à attaquer contrairement à ces mêmes calculs. Les calculs sont faux quant aux intentions des Russes et la décision ne découle pas du calcul. Cette anecdote illustre tout à la fois, l'ascendant qu'exerce le calcul des probabilités, dès qu'il s'agit de commenter une situation aléatoire, et la confusion entre jugement moral et jugement probabiliste. Cette communication veut montrer l'absence de fondement de cette identification en s'appuyant sur la généalogie du calcul des probabilités. Avant d'entreprendre cette enquête d'épistémologie historique, il n'est pas sans intérêt pour notre propos, de présenter rapidement une autre enquête, menée par B. Mandelbrot sur les marchés financiers⁶. Cette dernière, qui prend en compte la longue durée – et qui, en cela, est extérieure et globale – complète notre enquête sur le jugement et la décision appréhendée à la troisième personne du singulier⁷.

Probabilités et décisions sur les marchés financiers

Aujourd'hui, la théorie des probabilités, non seulement assume le rôle de légitimation des politiques économiques qu'elle a joué jusqu'ici, mais ses récents développements techniques – notamment l'utilisation à partir des années 70, du « calcul de Ito » (une intégrale stochastique) pour exprimer le bénéfice d'une série d'achats et de ventes de certaines quantités d'actifs – en font le noyau central de nouvelles pratiques financières qui ont, de l'avis de nombreux observateurs, « bouleversé la rationalité économique⁸ ». Ces stratégies financières comportent évidemment des prises de risque, mais celui-ci est dispersé dans une multitude de gestes dans des lieux différents, et globalement, ce risque a une structure fortement dissymétrique (forte probabilité d'un gain faible, faible probabilité d'une perte énorme, type de risque « catastrophe »). À cette dissymétrie, sont liés les risques de cataclysmes financiers. « L'un des points les plus controversés des produits dérivés est l'effet de levier : une somme d'argent limitée permet d'acheter une exposition au 'risque du marché' très importante⁹ ».

³ Voir par exemple, M. Leclerc-Olive, « Les nouvelles collectivités territoriales, entre aide internationale et pouvoirs locaux: légitimité instrumentale ou légitimité politique ? » dans Y. Lebeau, B. Niane, A. Piriou, M. de Saint Martin, (édits), *État et acteurs émergents en Afrique*, Karthala, 2003

⁴ J. L. Boursin, *Les structures du hasard. Les probabilités et leurs usages*, Seuil, 1986, p. 133.

⁵ *Le Monde*, 12 avril 2001.

⁶ B. Mandelbrot et R. L. Hudson, *Une approche fractale des marchés financiers. Risquer, perdre et gagner*, O. Jacob, 2005.

⁷ Il faudrait pour être complet, convoquer des comptes rendus rapportés à la première personne. Voir par exemple, Pierre-Antoine Fabre, « Espace mystique et espace politique : la décision ignatienne », dans P. A. Fabre, P. Gruson, M. Leclerc-Olive (dir), *L'actualité du XVII^e siècle*, à paraître.

⁸ N. Bouleau, *Martingales et marchés financiers*, Odile Jacob, 1998, p. 15.

⁹ J. Scheinkman, « Produits dérivés et effets de levier », dans B. Jacquillat et J. M. Lasry, *Risques et enjeux des marchés dérivés*, PUF, 1995. Cité par N. Bouleau, *op. cit.* p. 103.

Ces calculs qui, fondés sur l’outil probabiliste pour l’évaluation des risques, déterminent les décisions prises sur ces marchés, semblent donner des assurances. Mais, d’une part, ces assurances ont un horizon absolument court (quelques mois), d’autre part elles présupposent que les cours des produits sont des phénomènes « dérivables », c’est-à-dire qu’un certain type de régularité informe leurs variations, autorisant une relative prévisibilité. L’observation des cours boursiers sur de longues périodes, données qui constituent un formidable corpus de décisions empiriques agrégées, conduit à renoncer à cette modélisation par trop simplificatrice. Recourir aux modélisations standard utilisées sur les marchés financiers suppose que l’on fasse un certain nombre d’hypothèses implicites sur les phénomènes modélisés. Outre les hypothèses théoriques qui nous intéressent principalement ici, il convient de rappeler les hypothèses anthropologiques partagées par la plupart des économistes : 1) les gens sont rationnels et ne cherchent qu’à s’enrichir, 2) tous les investisseurs se ressemblent. Le point de départ de ces modèles standard, c’est la proposition de Louis Bachelier (1900) de décrire les cours de la bourse comme une « marche aléatoire ». Tout l’édifice repose sur l’idée que la courbe en cloche (la loi « normale ») est pertinente pour décrire la distribution des cours.

Il faut évidemment s’interroger sur le choix de cette famille de modèles mathématiques qui suppose qu’il est légitime de considérer que le phénomène étudié – les cours dans notre cas – varie de manière « continue » (continue au sens commun du terme, c’est-à-dire en fait « dérivable »). Plus précisément (et je suis ici au plus près des analyses de B. Mandelbrot) modéliser les variations des cours à l’aide du mouvement brownien, suppose que : 1) les variations de prix sont indépendantes l’une de l’autre (de la précédente mais aussi de celles qui l’ont précédée). Cela signifie que toute l’information est contenue dans les cours actuels et qu’il est inutile d’étudier les historiques des cours – que la condition « de martingale » est satisfaite : la meilleure estimation du cours de demain est le cours d’aujourd’hui. 2) Les processus qui engendrent les variations de cours sont invariables dans le temps. 3) Les variations de cours épousent une distribution normale (la plupart des variations sont faibles, très rares sont les variations d’amplitude importante). En termes mathématiques : « les mouvements de cours forment une séquence de variables aléatoires indépendantes et identiquement distribuées ».

L’observation des cours réels montre que ces hypothèses ne peuvent être, raisonnablement, tenues bien longtemps : l’observation de « queues épaisses » (qui contredisent le bien-fondé d’une approximation normale) et de diverses formes de dépendance (des variations de grande amplitude arrivent « en rafales », suivies de variations faibles elles aussi « en paquets ») ont été pendant de nombreuses années écartées comme étant des « anomalies¹⁰ ». Mais les théoriciens ont eu à rendre compte peu à peu, de ces anomalies. Pourquoi, dès lors, s’obstiner ? « Habitude et confort », répond Mandelbrot.

L’introduction de la notion de rugosité d’une courbe (propriété globale et non ponctuelle comme celles de continuité et de dérivabilité), issue de la théorie des fractales, permet d’exprimer à la fois l’intensité des variations (relative aux phénomènes de discontinuité), et leur dépendance (relative à l’ordre).

La position de surplomb (toujours extérieure et globalisante) du mathématicien qui modélise ainsi les longues séries des cours du coton ou du taux de change du yen, semble pouvoir dialoguer avec d’autres lignes argumentatives issues de perspectives ou d’horizons radicalement différents, notamment celle que je vais développer maintenant, et qui concerne la formalisation du jugement de l’acteur – à la troisième personne du singulier –, à la lumière des formalisations contemporaines du jugement.

Analyser les usages descriptifs et prescriptifs des probabilités requiert d’évaluer les effets induits par le postulat d’additivité¹¹ sur lequel repose le style de raisonnement qui consiste à « quantifier l’incertain ». Et pour ce faire, il est nécessaire de saisir ce style de raisonnement au moment de son émergence. Ce détour historique permettra de distinguer les différentes opérations intellectuelles impliquées dans le calcul des probabilités et la place de cette additivité qui signifie notamment que la probabilité d’un énoncé détermine la probabilité de l’énoncé contraire, selon un principe, sans reste, de vases communicants. Certes les économistes distinguent prudemment le risque et l’incertitude. Le risque pour les situations probabilisables, l’incertitude pour les autres. Mais il reste que l’absence de probabilité disponible n’efface pas pour autant le postulat implicite d’additivité comme matrice de pensée obligée, qui se traduit notamment par la confusion entre « croire que non p » et ne pas croire que p ».

Émergence des probabilités

Le calcul des probabilités qui s’ébauche à la fin du XVII^e siècle acquiert rapidement une robustesse technique qui lui confère rapidement le monopole de la formalisation de l’aléatoire.

¹⁰ Au fond, c’est le même processus que celui qui valut au système ptolémaïque sa longévité, faite de complexifications répétées d’un modèle inadéquat.

¹¹ D’autres postulats soutiennent également l’édifice : la possibilité d’une description exhaustive des possibles et évidemment l’indépendance des actions ou des jugements individuels, qui concerne leur enchaînement, ainsi que la stabilité des comportements. Indépendance et stabilité requièrent, pour être examinées que l’on ait affaire à des séries de données comme dans l’analyse de Mandelbrot présentée plus haut.

Pour L. Daston, le domaine spécifié de phénomènes dont traitent les probabilités est constitué par les « jugements des hommes éclairés, jugements formés en situation d'incertitude ». La théorie classique des probabilités, nous dit l'auteur, prétend répondre à une série de questions : est-il rationnel de jouer à cette loterie, de croire en Dieu, d'acheter une rente viagère, de condamner un accusé sur la base des informations disponibles ? « La fortune de la théorie classique des probabilités dépendait de la conformité de ses résultats aux intuitions des hommes raisonnables¹² ». Certes, le terme raisonnable est lui-même relativement indéterminé, notamment à la fin du XVIIe siècle. L. Daston note, par exemple, que la figure de l'homme raisonnable n'est pas la même pour Nicolas et Daniel Bernoulli. Pour le premier, l'homme raisonnable est celui qui défend un certain sens de l'**équité**, tel qu'il se donne à voir notamment dans les contrats aléatoires : c'est le juge désintéressé qui fait valoir ce qui est conforme à la justice et à l'égalité. Pour Daniel Bernoulli, qui tire ses exemples du monde des affaires et du commerce, le calcul des probabilités traduit la **prudence** commerciale du marchand avisé. Retenons de ceci la proximité, voire l'adéquation, entre sciences morales et calcul des probabilités que ces conceptions présupposent, et donc entre description et prescription, entre jugement et probabilité. « Les probabilistes classiques persistaient à croire que le raisonnable était univoque, malgré les débats interminables sur sa propre définition¹³ ».

L'objet de ce texte est proposer un commentaire critique de cette proximité entre calcul probabiliste et jugement, en introduisant une grille d'analyse permettant d'une part, de distinguer les opérations élémentaires de quantification qui, rassemblées, inaugurent le « calcul des probabilités » à la fin du XVIIe siècle, et, d'autre part, d'analyser les enjeux théoriques et politiques liés à la « probabilisation » du monde, enjeux qu'on peut mettre en perspective, à partir des formalisations du jugement en situation d'incertitude que nous offrent certains travaux récents, en Intelligence Artificielle, préoccupés de la mise au point de « systèmes experts ». Probabilité et jugement sont au début étroitement liés. Ils le resteront jusqu'à la fin du XIXe siècle comme en témoignent par exemple les travaux de Poisson, même si Condorcet, par exemple, distingue conceptuellement probabilité et « motif de croire ».

Si on suit L. Daston, la conception des probabilités qui prévaut au XVIIIe siècle qui n'est mise en question réellement qu'au XXe, repose sur l'idée que ce calcul n'est qu'une formalisation mathématique du jugement de l'homme raisonnable.

Que font les mathématiciens quand ils probabilisent le monde ?

Commençons par tenter d'identifier les modes opératoires élémentaires qui constituent l'innovation des années 1660. Il me semble qu'il faut distinguer, d'abord, deux opérations fondamentales, souvent intriquées mais pas toujours, et en tout cas, mêlées dans des proportions différentes selon les auteurs, les situations ou les problèmes à résoudre. Il s'agit, d'une part, de la production de nombres qui veulent signifier le poids ou la force d'un argument, la confiance en une opinion, l'intensité d'une croyance, à partir de ce qui l'étaye, de ses « éléments d'évidence » et, d'autre part, de l'établissement de proportions entre ces poids, ou entre ces degrés de confiance. Une troisième opération, possible uniquement si l'événement considéré est lui-même quantifiable (un gain, par exemple, ou un bien non « organique » selon l'expression de Keynes, c'est-à-dire atomique et sommable) est la pondération de ces gains par ces poids relatifs, pour « évaluer l'espérance ».

Soit une situation simplifiée, celle où l'on a affaire à deux événements ou énoncés p et q . 1) Il se peut que je puisse énumérer *tous* les éléments d'évidence qui supportent p , et que, de plus, ils soient *incompatibles* et entraînent *nécessairement* p . Si je connais le poids de chacun de ces éléments, je pourrai calculer le poids de p ¹⁴. De même pour q . Admettons que le poids de p soit N_1 et celui de q , N_2 . C'est là le premier mode opératoire. 2) Supposons à présent que q est l'énoncé *non* p . Alors dans ce cas, et par comparaison des poids respectifs de ces deux énoncés contraires, je pourrai « faire la balance » et dire que la probabilité de p est $N_1 / (N_1 + N_2)$, quotient qui représente la part de p dans le poids total. Il est clair que si q n'est pas *non* p , on ne pourra pas envisager de calculer une probabilité, on pourra tout au plus observer que le poids de p (par exemple) est supérieur au poids de q . Il se peut que, dans certains cas, ceci soit suffisant pour prendre une décision. Mais on ne pourra pas cependant parler de probabilité.

Ceci mérite d'être comparé à ce qui se passe lors d'un procès judiciaire¹⁵ « Lors d'un procès en justice, toute inférence est relative, ou, du moins, conditionnée par les éléments d'évidence dont dispose la cour » et on peut se trouver dans le dernier cas évoqué où une décision peut être prise sans que pour autant on ait évalué le poids de

¹² L. Daston, « L'interprétation classique des probabilités », *Annales ESC*, 44^{ème} année, n° 3, mai-juin 1989, p. 716.

¹³ L. Daston, *op. cit.* p. 726.

¹⁴ S'il apparaît que le principe d'« indifférence » s'applique, on pourra se contenter de les dénombrer.

¹⁵ Voir I. Hacking, *L'émergence de la probabilité*, Seuil, 2002, p. 133.

toutes les alternatives¹⁶. C'est cette dernière opération, la balance, le calcul de proportions (qui n'est possible que lorsqu'on a évalué le poids de toutes les éventualités), qui constitue l'innovation essentielle, même si l'opération précédente, évaluation du poids ou de la confiance est parfois le problème principal à résoudre¹⁷. 3) L'autre innovation (la troisième opération), consiste, là où c'est possible, à pondérer un avantage ou un danger par ce poids relatif, ce qui conduira à la notion d'espérance. Longuement discutée au cours des trois derniers siècles, elle est relative au versant des probabilités orienté directement vers l'action et le jugement.

La Logique de Port-Royal, dans le livre IV, met chacun en garde : « Il faut désabuser tant de personnes qui ne raisonnent guère autrement dans leurs entreprises qu'en cette manière : Il y a du danger en cette affaire donc elle est mauvaise, il y a de l'avantage dans celle-ci, donc elle est bonne ; puisque ce n'est ni par le danger, ni par les avantages, mais par la proportion qu'ils ont entre eux qu'il faut en juger ». On n'a pas attendu Pascal pour procéder ainsi lorsque la comparaison est aisée ? Certes¹⁸. Ce qui est neuf, c'est la tentative de quantifier cette comparaison. Ayant réussi à le faire dans certains cas particuliers, on tentera dès lors de le faire ailleurs, d'étendre le champ de pertinence de ce mode opératoire. Lorsque les auteurs de *La Logique* veulent convaincre du bien-fondé de la méthode, ils n'hésitent pas à avancer des chiffres qui ne correspondent à aucune donnée connue. « Si c'est le seul danger de mourir par le tonnerre qui leur cause cette appréhension extraordinaire, il est aisé de leur faire voir qu'elle n'est pas raisonnable ; car de deux millions de personnes, c'est beaucoup s'il y en a une qui meure de cette manière, et on peut dire même qu'il n'y a guère de mort violente qui soit moins commune¹⁹. » Comme le souligne Hacking, ce nombre de deux millions est purement rhétorique. Lorsqu'en revanche, on dispose de fragments de tables numériques (nombre de morts selon la maladie, par exemple), on sera tenté de les traiter de cette manière et de faire apparaître des fréquences, cette notion relevant en général des deux types d'opérations que j'ai distingués, mais parfois du second seulement, lorsqu'on ne fait que dénombrer des cas.

Pour traiter certains problèmes, il se peut que l'on ne compare pas des fréquences mais les nombres eux-mêmes. Jusqu'au milieu du XVII^e siècle, la probabilité est un attribut de l'opinion. Une opinion est probable lorsqu'elle est approuvée par une autorité, quand les livres anciens en témoignent. Ce qui manque jusque là, et qui constitue l'invention majeure c'est la notion d'« évidence factuelle » ou d'« élément d'évidence²⁰ ». « L'évidence factuelle pointe vers autre chose qu'elle-même, mais pas exactement à la manière du test, vers une hypothèse, mais vers une autre chose, un autre état physique, plutôt à la manière du diagnostic²¹ ». Jusqu'au milieu du XVII^e siècle, on ne peut concevoir d'« élément d'évidence » que dans le cadre d'une théorie. « Il n'existe pas de critère *épistémique indépendant* (je reprends ici la théorie de I. Hacking). C'est seulement quand des critères épistémiques peuvent être appréhendés indépendamment de ce que la théorie dit des causes, que peuvent émerger probabilité et usage des statistiques ». En témoigne par exemple le raisonnement de Graunt pour analyser les causes de la contagion de la peste. « La contagion de la peste dépend plus de disposition de l'air que des effluves émanant des corps des hommes. Ce que l'on prouve [...] par les brusques écarts faits par la peste, sautant en une semaine de 118 à 927, reculant de nouveau de 993 à 258, pour passer encore à 852 dès la semaine suivante²² ». L'argument porte sur la variabilité du nombre de cas qui s'apparente bien plus aux variations de la météorologie et dont il paraît difficile de rendre compte sous la seconde hypothèse. On se trouve ici dans le deuxième mode opératoire examiné plus haut, qui permet de comparer des poids, mais pas de calculer directement des probabilités, au sens de Bernoulli. Un tel raisonnement, fondé sur des critères *purements épistémiques*, ignore les « éléments d'évidence » qui relèveraient d'un raisonnement proprement théorique. La première opération élémentaire, qui consiste à évaluer la force probante ou le poids d'un événement ou d'un argument, est ici absente.

Que ces modes opératoires aient à la fin du XVII^e siècle une relative autonomie, en témoignent les travaux de J. Bernoulli. Dans certains travaux de Bernoulli, la notion de probabilité en tant que quantification de la comparaison des poids d'arguments est franchement passée au second plan, au profit d'une probabilité conçue

¹⁶ L'opération de comparaison n'a pas cependant complètement disparu puisqu'il subsiste, même si elle n'est pas quantifier, des opérations d'évaluation de la cohérence entre les divers énoncés ultimes prononcés.

¹⁷ Voir par exemple les tentatives de formalisation de ce calcul par J. Bernoulli.

¹⁸ Voir par exemple, Al-Ghazali, *Vivification des sciences de la foi*, III, 365, I, 19. Ce texte est cité par Miguel Asin Palacios, dans *Huellas del islam*, 1920.

¹⁹ Arnaud et Nicole, *La logique... op. cit.* p. 430.

²⁰ Ces éléments d'évidence correspondent à ce qu'Arnauld distingue comme circonstances intérieures. « J'appelle circonstances intérieures celles qui appartiennent au fait même et extérieures celles qui regardent les personnes par le témoignage desquelles nous sommes portés à le croire ». Arnauld et Nicole, *La logique ou l'art de penser*, p. 349. Il s'agit de l'« évidence » qu'une chose ou un événement transmet à une autre chose ou à un autre événement. C'est une des significations du terme anglo-saxon « evidence » qui n'a pas qu'équivalent simple en français.

²¹ I. Hacking, *L'émergence... op. cit.* p. 71

²² Cité par I. Hacking p. 150.

comme évaluation du poids, de la force probante d'un argument, comme degré de certitude. Cet usage révèle clairement l'autonomie relative des deux opérations mises en pratique dans le calcul des probabilités naissant (la quantification du poids d'un argument d'une part, la proportion qu'ont ensemble les poids de toutes les éventualités d'autre part, et que Bernoulli distingue en la nommant *probabilité absolue*). « Si outre les arguments favorables à la chose à approuver, s'offrent d'autres arguments purs par lesquels le contraire est persuadé, les arguments des deux genres sont à pondérer en particulier au moyen des précédentes règles, afin de pouvoir établir la raison qui intervient entre la probabilité de la chose et la probabilité du contraire. On doit noter, ici, que si les arguments apportés dans les deux parties sont assez forts, il peut se faire que la probabilité des deux parties dépasse notablement la moitié de la certitude, c'est-à-dire rende probable chacun des deux contraires, quoiqu'une plaidant relativement à l'autre soit moins probable que l'autre ; ainsi il peut se faire que quelque chose ait 2/3 de certitude, tandis que son contraire possède 3/4, et de cette façon les deux contraires seront probables, et cependant le premier moins probable que son contraire, et cela dans la raison de 2/3 à 3/4, soit 8 à 9²³. »

Mais, et c'est, me semble-t-il, le point crucial, lorsqu'on parle de la force probante d'un argument on attribue une force à une *inférence* et non à une simple *croissance*. La pertinence de cette distinction apparaît dès lors que l'on veut composer des arguments entre eux, comme l'a bien vu Bernoulli. C'est précisément cette confusion qui, d'une part, permet d'espérer que le calcul des probabilités puissent non seulement informer l'action (ou le jugement), mais la déterminer, et qui, d'autre part, est levée dans les formalisations récentes de l'incertitude.

Comment les jugements de vraisemblance évoluent dans le temps ?

Il faut ici accorder une attention toute particulière à l'évaluation de la vraisemblance d'une *inférence*, et, de manière plus générale, à l'évolution des degrés de croyance (de jugement, de décision) dans le temps. La plus grande partie des résultats de la théorie des probabilités n'est pas affectée par la diversité de ses interprétations (probabilité subjective, objective, épistémique, quantitative, etc.). Néanmoins, une exception au cœur de ce consensus : des divergences apparaissent lorsqu'il s'agit de formaliser la manière dont les probabilités évoluent avec le temps, lorsqu'il s'agit de donner forme à une « cinématique » des probabilités²⁴. La conception bayésienne, largement admise, héritière du postulat d'additivité, réduit ce processus à une pure conditionalisation.

Van Fraassen, en distinguant deux concepts de rationalité qu'il nomme avec humour, le concept prussien et le concept anglais, illustre une forme non bayésienne (bien que possiblement additive quant au fond) de dynamique de la croyance. « Ce qu'il est rationnel de croire, est exactement ce que l'on est rationnellement contraint de croire » est la version prussienne²⁵ ; « ce qu'il est rationnel de croire inclut tout ce qu'on n'est pas rationnellement contraint de refuser » est la version anglaise²⁶. Suivons le commentaire de l'auteur : « La première consiste à dire qu'il existe certaines règles pour bien diriger sa raison, des canons de la logique au sens large, qu'il est nécessaire de suivre et qui dictent l'unique manière permise de s'adapter à ce que l'expérience nous fournit de nouveau. Cette conception est celle à laquelle les écrits de Newton sur l'induction faisaient allégeance, du moins officiellement, et c'est aussi celle que professe strictement le bayésien orthodoxe. La seconde conception consiste à dire que nous créons et inventons de nouvelles hypothèses et théories et que nous agissons non seulement par inclination, mais aussi de manière rationnelle lorsque nous les adoptons après avoir pris en considération le témoignage des données qui leur sont favorables, même si elles vont bien au-delà de tout ce que nous aurions pu espérer constituer en mélangeant adaptation aux données et opinions antérieures. Cette conception de la libre entreprise rationnelle de l'esprit (qui requiert un concept permissif de rationalité) a été défendue de manière adéquate par les pragmatistes américains²⁷. » Van Fraassen cite W. James²⁸ pour qui « lorsque nous formons une opinion, nous poursuivons deux buts principaux : croire le vrai et éviter l'erreur. L'intensité avec laquelle nous poursuivons chacun de ces buts, qui nous entraînent dans des directions différentes, doit être pour une part affaire de choix (...). » « James illustre ainsi clairement l'une des conceptions traditionnelles de la théorie de la connaissance, qui admet l'existence de sauts rationnels légitimes laissant loin derrière la rassurante sécurité²⁹ de notre mélange de témoignages-des-données-plus-opinions-antérieures. Comment aurions-nous pu, sinon, adopter jamais rationnellement la haute opinion dans laquelle nous tenons

²³ J. Bernoulli, *L'art de conjecturer*, traduction de J. Peyroux Blanchard, 1998, p. 157.

²⁴ Voir par exemple B. Van Fraassen, *Lois et symétries*, Vrin, 1994, p. 452.

²⁵ Par analogie avec la conception prussienne de la loi : tout ce qui n'est pas explicitement permis est interdit.

²⁶ Bas van Fraassen, *Lois et symétries*, Vrin, 1994, p. 279.

²⁷ *Ibid.* p. 280.

²⁸ W. James, « The will to believe » dans *Essays in Pragmatism*, Hafner, New York 1948.

²⁹ Il faut entendre ici la mécanique bayésienne.

aujourd'hui les théories de Darwin, d'Einstein et de Bohr (...) ?³⁰ » W. James et, après lui, Van Fraassen préconisent de traiter un énoncé et son contraire³¹ de manière en partie autonome (« les enquêtes nous entraînent dans des directions différentes »). Partant, une formalisation additive de la cinématique des jugements s'en trouve disqualifiée.

Les chercheurs du XIXe siècle avaient l'espoir de pouvoir placer les sciences morales sous l'autorité des raisonnements mathématiques, mathématique étant entendu dans le sens très restrictif de méthodes numériques précises. C'est la visée des travaux de Poisson, par exemple, qui intitule son traité, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*. Poisson renonce à toute évaluation des probabilités de base à partir d'éléments d'évidence. C'est la loi des grands nombres appliquées à des observations *a posteriori* qui lui permettra d'estimer la probabilité qu'un juge a de ne pas se tromper³². Poisson quantifie les moments 2 et 3 de notre matrice pour « calculer » une décision. Poisson, comme beaucoup d'autres, au XIXe siècle, exprime un fort préjugé en faveur de l'additivité des modélisations de l'incertitude et pour la similitude (et donc l'indépendance) des expériences répétées³³.

Malgré le recours, de plus en plus fréquent au XIXe siècle, aux statistiques et probabilités pour déterminer les paramètres des assurances, il est intéressant de noter que l'assurance maritime expriment encore des réticences à ces modes de décision.

Au XIXe siècle, l'assurance maritime (bien qu'elle ait été la première forme d'assurance) réside encore à caler ses décisions sur des données statistiques.

Je prendrai un exemple qui n'est probablement pas généralisable mais qui a le mérite de forcer les différences entre les trois moments de l'approche du « probable » : 1) son approche à l'aide des éléments d'évidence, en termes de croyance, (qui ont quelque chose à voir avec le concept de passé chez Mead), 2) son approche à l'aide du calcul de proportions (présent) et enfin 3) son approche en termes de recommandations, d'inférences, de jugements (futur).

Cet exemple est emprunté à l'ouvrage de L. Boiteux, ancien président de l'Académie de Marine, intitulé *La fortune de la mer. Le besoin de sécurité et les débuts de l'assurance maritime*.

L'auteur expose notamment la manière dont on a déterminé, entre le XVIe et le XIXe siècle, la rémunération des assureurs. Celle-ci est en général composée d'un intérêt sur le capital, et d'un « prix du péril ». Celui-ci est fixé au cas par cas, en fonction de « l'espèce de la marchandise, le plus ou moins de danger qui paraît sur la mer, si c'est en paix ou en guerre, si on dit qu'il y a beaucoup de pirates ou qu'il n'y en a pas, la saison, la longueur de la navigation, la distance des lieux, la pays où se fait la navigation, le plus ou moins d'expérience que le maître du navire pouvait avoir des mers et des côtes, etc. » Les primes d'assurance n'ont alors rien de fixe ni de réglé. Si bien que les marchands cherchent les assureurs qui leur font le meilleur parti : c'est ainsi le marché qui régule, *in fine*, la hauteur des rémunérations.

Un assureur du XIXe siècle, Alfred de Courcy, exprime sa méfiance et son désintérêt pour une approche du risque en termes statistiques (et donc purement rétrospectifs). Il écrit, en 1862, dans un ouvrage intitulé *Essai sur les lois du hasard*, « le tâtonnement, la discussion, l'inspiration individuelle, l'exemple des assureurs influents, les bénéfices acquis où les pertes à récupérer, le degré de confiance attribué à tel navire ou à tel capitaine, l'état de la température, la concurrence, la proportion de l'offre et de la demande, voilà tout ce qui concourt à diriger les rapides gageures des assureurs ». L'auteur considère que l'assurance maritime est autant une science *expérimentale* (elle requiert du bon sens et des connaissances variées et étendues) qu'une science *augurale*, dans laquelle le « flair » tient une place non négligeable.

L. Boiteux le cite : « J'entends les théoriciens se récrier contre tant de barbarie ; faites une bonne statistique des naufrages, diront-ils et vous calculerez ensuite avec précision vos primes. (...) À quoi un homme de métier répondra : « quand vous m'apporterez la statistique la plus exacte des naufrages elle n'aurait pour moi qu'un intérêt de curiosité ; car les circonstances de la navigation, ses dangers, ses chances changent tous les jours. Le temps que vous aurez passé à rassembler et à coordonner vos documents, l'activité commerciale l'aura employé à détruire leur autorité... ». Notant les grands changements qui affectent la navigation à cette époque (ouverture du canal de Suez, introduction de l'hélice, de la vapeur, etc.) il ajoutait : « Convenez qu'en présence de tous ces grands changements, vos laborieuses investigations du passé ont peu de valeur. L'activité des hommes ne s'arrêtera pas et votre statistique sera éternellement en retard. Le commerce est là, avec ses besoins urgents, qui m'interroge, qui me presse, qui sollicite une réponse immédiate³⁴... ». Bien sûr, comme le note L. Boiteux, les

³⁰ *Ibid.* p.281.

³¹ Qui fait écho, on l'aura vu, au couple liberté positive, liberté négative, ou encore au deux sortes d'erreur que l'on commet lorsqu'on applique un test statistique.

³² S. D. Poisson, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*, Éditions Jacques Gabay, 2003 (1837), p. 368.

³³ Le débat entre Bienaymé et Poisson est à cet égard très éclairant.

³⁴ L. A. Boiteux, *La fortune de la mer. Le besoin de sécurité et les débuts de l'assurance maritime*, SEVPEN, 1968.

assureurs maritimes ne se prononcent pas comme les assureurs « non-marine » qui ont beaucoup plus tôt eu recours aux « computers ».

Je retiens de cette citation un peu longue, que le manque de confiance dans les statistiques concerne le niveau « horizontal », celui qui cherche des informations dans la comparaison entre des données, niveau purement épistémique (c'est-à-dire hors théorie) au sens où J. Hacking l'a employé à propos du raisonnement de Graunt rappelé plus haut. L'assureur, en revanche, retient comme informations fiables, même si elles ne permettent pas aux yeux de l'assureur de faire un « calcul », les indices, les éléments d'évidence qui participent au poids (ou à la probabilité au sens de Bernoulli) d'un événement ou d'un énoncé. Il souligne aussi le caractère inévitablement rétrospectif des calculs statistiques, qui en affaiblit la valeur « augurale ». Je voudrais souligner également que cet exemple donne à voir que les éléments qui sont à prendre en compte pour évaluer le « prix du péril » ne sont pas forcément identiques à ceux qui se dégageront de l'analyse statistique des naufrages. Si ceux-ci se produisent principalement au cours de la traversée du golfe de Gascogne, en quoi servent-ils à évaluer la prime pour un aller-retour entre Bordeaux et Ostende ? Alors que dans la théorie des probabilités tous les arguments sont mixtes (ils sont à mettre au dossier des deux propositions, p et $non\ p$), on voit que dans la pratique du jugement, il paraît raisonnable de considérer que des arguments peuvent être purs³⁵, qu'ils portent à charge de p exclusivement, sans rien dire en faveur de $non\ p$. On n'a pas forcément ici l'équivalence entre « croire que $non\ p$ » et « ne pas croire que p ».

En résumé, la grille d'analyse des approches de l'incertitude proposée comporte trois « moments » :

- 1) Ce qui relève de la formation et de la confiance dans des connaissances, des hypothèses, des croyances, et qui porte le regard sur ce qui accrédite ces croyances. Il faut évidemment, I. Hacking nous y a fortement invités, veiller à ne pas confondre ce qui relève d'un lien causal ou explicatif (une problématique théorique) et ce que B. Winters appelle des « considérations de vérité » : les témoignages et les éléments d'évidences.
- 2) Ce qui relève d'un travail de comparaison, de balance, entre des poids, des occurrences. C'est un mode opératoire que l'on voit émerger notamment dans le style d'argument, portant sur des effectifs, inauguré par Graunt à propos des épidémies. La première forme achevée se trouve dans la Logique de Port-Royal.
- 3) Ce qui relève de l'action, de la décision, de la conduite. Le premier instrument mathématique de l'inférence, du jugement pour l'action, pour tenter de déterminer ce qu'il convient de faire est la notion d'espérance. On connaît les controverses qu'elle a suscitées, notamment autour du paradoxe de Saint Pétersbourg.

La formule paradigmatique du « pari mondain³⁶ » peut être représentée par la règle suivante :

Le joueur J est placé devant le jeu suivant³⁷ : s'il « gage³⁸ » il gagne a avec une probabilité p , il perd b avec une probabilité q (évidemment $p+q=1$).

Gain	Perte
a	b
p	q

Si a surpasse plus b que q surpasse p , c'est-à-dire si $ap > bq$, on peut « gager ». Ce que nous apprend l'axiome d'Eudoxe³⁹, c'est que si a et b sont de *même ordre*, même si a est très grand et b très petit, il existe des nombres p et q tels que $ap < bq$. Autrement dit, même si la promesse de gain est très grande et la perte possible petite, il se peut que « la proportion que toutes ces choses ont ensemble » recommande de ne pas jouer. « Il est de la nature des choses finies de pouvoir être surpassée, quelques grandes qu'elles soient, par les plus petites, si on les multiplie souvent, ou que ces petites choses surpassent les grandes en vraisemblance de l'événement, qu'elles n'en sont surpassées en grandeur. Ainsi le moindre petit gain peut surpasser le plus grand qu'on puisse imaginer, si le plus petit est souvent réitéré, ou si ce grand bien est tellement difficile à obtenir, qu'il surpasse moins en grandeur que le petit en facilité. Et il en est de même des maux que l'on appréhende, c'est-à-dire que le moindre petit mal peut-être plus considérable que le plus grand mal qui n'est pas infini, s'il le surpasse par cette proportion⁴⁰. » C'est donc là précisément l'invention majeure de cette fin du XVIIème siècle, cette règle de décision pour les affaires mondaines.

³⁵ D. Dubois et H. Prade, *Théorie des possibilités*, Masson, 1988, p. 17.

³⁶ Par opposition au « pari existentiel » du fragment Infini-Rien des *Pensées* de Pascal.

³⁷ Un exposé plus détaillé de cette approche anthropologique du pari est présenté dans M. Leclerc-Olive, *Gestion de l'incertitude et temporalités*, à paraître. En toute rigueur, ce tableau ne devrait pas être présenté en termes de probabilités, mais de « chances », ce qui est fait dans ce dernier texte.

³⁸ Par exemple en misant b dans un jeu où, si J gagne, il touche $a+b$.

³⁹ Que Pascal attribue à Archimède. J.-L. Gardies, *Pascal entre Eudoxe et Cantor*, op. cit.

⁴⁰ A. Arnaud et P. Nicole, *La logique...* p. 430.

Un nouveau probabilisme ?

En écrivant les *Provinciales*, Pascal s'attaque au « probabilisme » qui forme un ensemble de maximes que de nombreux casuistes (et parmi eux surtout des jésuites) utilisent pour énoncer les recommandations et les conduites à tenir dans les différents cas de conscience qui leur sont soumis. Parmi ces cas de conscience figurent notamment les questions d'usure et de prêts à intérêts. Peu de temps après la publication de ces *Lettres* et de la *La logique de Port-Royal*, les autorités pontificales en viennent à condamner les maximes élaborées par les casuistes pendant près d'un siècle, maximes dont certaines, aujourd'hui, nous sidèrent par leur laxisme moral. La place manque ici pour retracer toutes les péripéties de la signification du terme « probable ». Contentons-nous de noter qu'il passe d'une acception qui entretient la plus grande proximité avec celle de certitude, (tout proche est le glissement de probable vers approuvé puis approuvable par une autorité reconnue⁴¹ (les Pères de l'Église, puis d'autres savants) à celle toute différente de « qui peut se produire ». Le point de basculement, pour la signification du terme, et d'ouverture pour les opinions et actions autorisées, paraît être une maxime de Medina (1577) dont l'expression excède les intentions de l'auteur : « Si une opinion est probable, il est permis de la suivre, quand même est plus probable l'opinion opposée ». Cette maxime a un double effet : elle introduit la conception moderne de la probabilité (elle est susceptible de degré) bien qu'en fait l'auteur ait en vue une acception proche de la conception aristotélicienne, 2) elle libère le sujet de son statut de créature (dont l'action ne peut que se conformer à une règle qui la détermine) au profit de décisions qui peuvent prendre en compte des « intérêts » échappant à la juridiction religieuse.

En recommandant de calculer une bonne décision (la matrice de Port-Royal) – en dehors bien sûr de ce qui relève de la foi, (cf le fragment des *Pensées* infini-rien) –, Pascal et Port-Royal font plusieurs choses à la fois. D'une part, ils consacrent le concept « moderne » de probable, autorise la séparation du champ des affaires « terrestres » d'avec celui de la foi, ils dégagent le croyant de l'obligation de s'en remettre aux autorités ecclésiastiques pour prendre une décision, et résolvent le « scandale » induit par le laxisme du probabilisme, en introduisant une « règle » de décision fondée sur la « pesée croisée » des actions et de leur probabilité (moment 2 de la probabilisation). On peut représenter la matrice du probabilisme sous la forme d'un tableau où la recherche de l'intérêt ne semble plus contrainte par rien.

Probable	Plus probable
Intérêt plus grand	Intérêt

La maxime de Medina, dans sa lecture laxiste, ignore au fond la première ligne du tableau, à la différence de la maxime de Port-Royal. Celle-ci réintroduit une autorité (celle du calcul, de la méthode) là où semblait pouvoir proliférer des actions sans limite. Elle définit ce qui est raisonnable, au lieu de dicter ce qui est bon. À ce titre, elle contribue à fonder une nouvelle rationalité de l'action.

Si l'émergence du « probabilisme » percute de plein fouet le dogmatisme religieux, lequel n'avait pas réussi au cours des siècles précédents à formuler des maximes claires et précises permettant de fixer avec certitude les conduites à suivre en toutes circonstances (notamment dans le nouveau contexte marchand et commercial), la règle de Port-Royal reconfigure une règle et par là une autorité qu'il conviendra de suivre. Aujourd'hui, le dogmatisme du savoir réapparaît précisément à travers l'autorité du calcul probabiliste, tendant à confiner le sujet à nouveau dans un rôle de simple créature, de créature du marché.

Le postulat d'additivité

La théorie « classique » des probabilités propose de confondre les trois moments, de rassembler les trois modes opératoires encore distinguables au moment de l'émergence de ce style de raisonnement, en une seule opération intellectuelle qui, du coup, étend, par contamination, le principe d'additivité à l'ensemble du champ.

Dès le début du XXe siècle, des réserves sont formulées. Dans sa thèse, soutenue en 1906, qui s'intitule *A treatise on probability*, Keynes soutient que l'additivité ne concerne que des situations très particulières. D'une part, en cette période de pensée utilitariste, il défend l'idée que la notion de bien elle-même n'est pas intégralement additive (une part des biens est de caractère « organique » et non atomique et sommable) et, surtout, que les degrés de probabilité sont pas toujours numériquement mesurables (« Normal ethical theory at the present day, if there can be said to be any such, makes two assumptions :first, that degrees of goodness are

⁴¹ Cette acception du terme probable est particulièrement ébranlée par la « découverte du Nouveau Monde » et la révolution copernicienne. La controverse de Valladolid et la première publication de la conception copernicienne sont contemporaines. Le sens ancien de probable, lié à une autorité incontestable ne peut plus servir de prémisse assurée, on ne le retrouvera que comme conséquence, certifiée par le « syllogisme probabiliste », .

numerically measurable and arithmétique additive, and second, that degrees of probability also are numerically measurable⁴². ») Ces hypothèses, au fondement de la théorie éthique « normale » qu'il rejette, forment une condition nécessaire pour introduire la notion d'espérance (déjà critiquée de manière pertinente notamment par D'Alembert au XVIIIe siècle⁴³). Mais, si Keynes souligne les limites de pertinence de l'additivité⁴⁴, il développera néanmoins une théorie additive de la formation de la croyance rationnelle, là où elle lui semble légitime, qui sur un principe d'indifférence. Il en précise lui-même les limites : il s'agit d'une théorie qui, faisant fond sur des connaissances certaines, appelées propositions primaires, rend compte du poids que l'on peut accorder à des propositions secondaires. Pour Keynes, cette manière de rendre compte de la formation des connaissances est conforme à l'expérience courante. Néanmoins, il souligne que sa théorie devrait être modifiée si, à la différence de son hypothèse de travail, les prémisses primaires ultimes des arguments ne sont que des « connaissances probables⁴⁵ ». Le procédé n'est donc pas itératif, car à la seconde étape, on n'a plus de certitude sur les propositions secondaires.

Notons cependant que (et c'est, me semble-t-il, le point crucial) lorsqu'on parle de la force probante d'un argument on attribue une force à une *inférence* et non à une simple *croyance*. La pertinence de cette distinction apparaît dès lors que l'on veut composer des arguments entre eux. C'est précisément le fait de traiter l'*inférence* comme une simple *croyance* qui permet d'espérer que le calcul des probabilités puissent non seulement informer l'action (ou le jugement), mais la déterminer. Bernoulli avait attiré l'attention sur ce problème, mais ces travaux qui ne postulaient pas une additivité *a priori*, ont longtemps été écartés de l'histoire de la pensée probabiliste. C'est à cet endroit précis que les modélisations récentes de l'incertitude introduisent une mécanique inférentielle dont je voudrais exposer quelques traits. J'utilise l'expression « mécanique inférentielle » et non calcul, même si l'outil informatique permet *in fine* de produire des résultats numériques, parce que ceux-ci ne peuvent faire l'objet d'une simple arithmétique, et oblige à reprendre à nouveaux frais l'analyse du rapport entre calcul et jugement.

Deux présupposés implicites conditionnent en fait les approches probabilistes de l'incertitude⁴⁶ : d'une part les événements (ou l'ensemble des propositions formulables, des éventualités) peuvent être décrits de manière exhaustive⁴⁷. D'autre part on postule l'additivité de la « probabilité » en tant qu'application, c'est-à-dire que la probabilité de $p \vee q$, si p et q sont incompatibles (p et q ne peuvent être satisfaites en même temps) est la somme des probabilités de p et de q . Ce second postulat est lié au premier, lequel doit être examiné de manière minutieuse selon la nature de la situation concernée. Notons que ces conditions pourraient bien n'être que rarement satisfaites, réduisant ainsi l'extension du champ d'application légitime des probabilités. K. Popper a tenté d'engager un programme de recherche sur cette question dès le début des années trente. Il avait convenu avec R. Carnap, qu'il fallait « distinguer clairement, d'une part, les probabilités telles qu'elles sont utilisées en physique, en particulier en mécanique quantique, et qui satisfont aux principes du « calcul des probabilités » ; et d'autre part, ce que l'on appelle la « probabilité des hypothèses », leur degré de confirmation, ou encore (...) leur « degré de corroboration ». K. Popper écrit « Nous nous accordions sur ceci qu'il *ne* fallait *pas* présupposer, sans, du moins, de sérieux arguments, que le degré de confirmation (ou de corroboration) d'une hypothèse satisfait aux principes du calcul des probabilités. Nous étions d'accord pour considérer, au vu des arguments développés dans la *Logik*⁴⁸, que ce problème demeurait ouvert, et qu'il constituait à vrai dire *le* problème central⁴⁹. » Comme nous le savons, R. Carnap n'a jamais donné suite à cet engagement : dans l'ouvrage⁵⁰ qu'il publie quinze ans plus tard il affirme, sans réellement argumenter, selon K. Popper, que le degré de confirmation d'une hypothèse est une probabilité au sens du calcul des probabilités.

Modéliser la formation du degré de confirmation d'une hypothèse, sans imposer les postulats du calcul des probabilités, peut éventuellement s'inspirer des théories des possibilités qui prennent en charge une des intuitions de Bernoulli, restée en friche faute d'outils mathématiques suffisamment perfectionnés.

⁴² J. M. Keynes, *A treatise on probability*, Macmillan for the Royal Economic Society, 1973, p. 343.

⁴³ *Ibid.* p. 347.

⁴⁴ *Ibid.* pp. 29-30.

⁴⁵ *Ibid.* p. 18.

⁴⁶ C'est le travail d'axiomatisation des probabilités de Kolmogorov (1933) qui met au jour ces présupposés.

⁴⁷ Van Frassen inclut même l'équiprobabilité parmi ces exigences.

⁴⁸ Publié en français sous le titre *La logique de la découverte scientifique*, Payot, 1992.

⁴⁹ K. Popper, *Un univers de propensions*, L'éclat, 1992, p. 24.

⁵⁰ R. Carnap, *Logical Foundations of Probability*, Chicago, 1950.

Des modélisations de l'incertitude non additives

Le premier pas de la formalisation consiste à poser un axiome qui a prétention à être un axiome " naturel " : la fonction g qui associe une mesure de confiance⁵¹ (ou un " poids ") à un argument, un énoncé p ou un événement A ⁵², est " monotone ", c'est-à-dire que :

si $p \Rightarrow q$ (ou si $A \subseteq B$) alors $g(p) \leq g(q)$ (ou $g(A) \leq g(B)$).

Cet axiome ne constitue pas un geste définitoire. C'est la formulation (au sens ethnométhodologique) d'une règle inscrite dans une forme de vie. Pour autant, cet axiome « naturel » ne porte que sur un ensemble restreint d'énoncés, duquel sont exclus les énoncés contenant des verbes potentiels ou des contrefactuels⁵³.

Ce postulat induit un certain nombre de propriétés, notamment :

$g(p \wedge q) \leq \min (g(p), g(q))$ et $g(p \vee q) \geq \max (g(p), g(q))$

Mais celles-ci ne suffisent pas cependant à déterminer un calcul.

D'autres gestes doivent être faits pour restreindre la classe des fonctions g , ces traits sont trop vagues pour permettre la poursuite du travail de formalisation. Il faut imposer des propriétés supplémentaires qui permettront d'engager une mécanique inférentielle. Or, comment sont choisies les propriétés que l'on impose à g ? Elles sont déterminées très largement par la configuration attribuée aux connaissances disponibles. Ce geste définitoire désigne donc en creux la sous-détermination du modèle, mais en même temps l'indexicalité de la procédure de raisonnement elle-même.

Pour illustrer ceci il est nécessaire d'entrer plus avant dans le compte rendu formalisé de l'argumentation.

Les connaissances peuvent être représentées par un ensemble de propositions. Notons p_i une proposition focale (ou E_i un élément focal) c'est-à-dire une proposition qui peut être imprécise mais à laquelle l'état de nos connaissances permet d'associer $m(p_i)$, une " part de croyance " (laquelle ne prétend à aucune propriété, en particulier ce n'est pas une mesure de confiance) et F l'ensemble de ces propositions focales. Nous ne perdons aucune généralité à supposer que :

$$m(\emptyset) = 0 \quad \text{et} \quad \sum m(p_i) = 1$$

$$p_i \in F$$

On associe alors à une proposition q les mesures de confiance suivantes :

$$g_1(q) = \sum_{p_i \Rightarrow q} m(p_i) \quad \text{et} \quad g_2(q) = \sum_{p_i \not\Rightarrow \text{pas non } q} m(p_i)$$

$$\text{ou } g_1(A) = \sum_{E_i \subseteq A} m(E_i) \quad \text{et} \quad g_2(A) = \sum_{E_i \cap A \neq \emptyset} m(E_i)$$

1) Si les propositions focales sont toutes des propositions élémentaires (c'est-à-dire que chaque proposition focale est incompatible avec une proposition qu'elle n'entraîne pas), ou, dit autrement, que tous les éléments focaux sont des événements élémentaires (c'est-à-dire que tout événement est une réunion d'événements focaux), les mesures de confiance g_1 et g_2 sont égales et on démontre qu'il s'agit de probabilités. La mécanique inférentielle qui peut être mis en œuvre est celle du calcul des probabilités.

2) Si les propositions focales sont deux à deux incompatibles (ou les éléments focaux deux à deux disjoints), la mesure de confiance $g_1(A)$ (notée $P_*(A)$) est calculée en considérant tous les éléments focaux qui rendent nécessaire l'occurrence de A , $g_2(A)$ (notée $P^*(A)$) est calculée en considérant les éléments focaux qui ne rendent pas A impossible. Dans ce cas, on considère qu'une probabilité existe mais que les connaissances dont on dispose permettent seulement de proposer un intervalle qui la contient, $[P_*(q), P^*(q)]$, défini par :

$$P_*(q) = \sum_{p_i \Rightarrow q} m(p_i) \quad \text{et} \quad P^*(q) = \sum_{p_i \not\Rightarrow \text{pas non } q} m(p_i)$$

ou encore :

$$P(A) \in [P_*(A), P^*(A)]$$

$$\text{où } P_*(A) = \sum_{E_i \subseteq A} m(E_i) \quad \text{et} \quad P^*(A) = \sum_{E_i \cap A \neq \emptyset} m(E_i)$$

⁵¹ D. Dubois et H. Prade, *Théories des possibilités*, Masson, 1988, p. 13.

⁵² Classiquement, on a des événements (associés à un corps de connaissances imprécises et incertaines), considérés comme des sous-ensembles d'un référentiel Ω dit " événement certain " ($g(\Omega) = 1$). l'ensemble vide \emptyset est identifié à l'événement impossible ($g(\emptyset) = 0$). Notons que $g(A) = 1$ ne signifie pas que A est certain. Par ailleurs ces fonctions d'ensemble ne sont pas des mesures au sens mathématique du terme (elles ne sont en général pas additives).

⁵³ M. Leclerc-Olive, « Probabilités et formalisations du jugement », Communication aux journées d'étude *L'actualité du XVIIe siècle*, Maison des Sciences de l'Homme, 27 et 28 novembre 2003

On les appellent parfois improprement probabilités inférieure et supérieure, bien qu'il ne s'agisse pas de probabilité (P_* et P^* ne sont pas additives).

3) si les éléments focaux forment une suite d'ensembles emboîtés $E_1 \subset E_2 \subset E_3 \subset \dots \subset E_n$ (ce qui correspond à une situation fréquente où les connaissances sont plutôt concordantes et non dispersées comme dans les cas précédents), les mesures de confiance g_1 et g_2 sont respectivement les mesures de nécessité N et de possibilité Π de la théorie des possibilités⁵⁴, dont la machinerie inférentielle est définie par :

$$\Pi(p \vee q) = \max(\Pi(p), \Pi(q)) \quad \text{et} \quad N(p \wedge q) = \min(N(p), N(q))$$

4) enfin, dans le cas le plus général où on ne dispose d'aucune propriété particulière supplémentaire sur la structure de F relative à la configuration des connaissances, les mesures de confiance g_1 et g_2 sont respectivement la mesure de crédibilité C_r et la mesure de plausibilité P_1 de Shafer⁵⁵ mais on ne peut obtenir dans ce cas de contraintes inférentielles suffisantes qui définiraient à elles seules un calcul puisqu'elles prennent la forme suivante :

$$C_r(p \vee q) \geq C_r(p) + C_r(q) - C_r(p \wedge q)$$

$$P_1(p \wedge q) \leq P_1(p) + P_1(q) - P_1(p \vee q)$$

Notons que la formalisation, telle que nous venons de la présenter, articule des conditions "sémantiques" et "syntaxiques" sans qu'ils soient possible de les séparer.

Quelques commentaires permettent de "contrôler" en quoi ces définitions sont conformes au sens commun.

Prenons l'exemple des mesures de possibilité et de nécessité.

D'abord, $N(A) \leq \Pi(A)$, ce qui est conforme à l'intuition selon laquelle il faut qu'un événement soit possible avant d'être nécessaire.

Ensuite, alors que $P(A) + P(\text{non}A) = 1$ dans le cas des probabilités (postulat d'additivité),

$$N(A) + N(\text{non}A) \leq 1 \quad \text{et} \quad \Pi(A) + \Pi(\text{non}A) \geq 1$$

La possibilité (ou la nécessité) d'un événement et celle de son contraire ne sont que faiblement liées. En particulier, pour caractériser l'incertitude relative à un événement ou une proposition on a besoin de deux nombres, $\Pi(A)$ et $N(A)$, liés par les relations : $N(A) < 0 \Rightarrow \Pi(A) = 1$ et $\Pi(A) < 1 \Rightarrow N(A) = 0$

S'il y a lieu de parler de calcul, c'est en sens beaucoup plus large que ce que l'on peut entendre par calculabilité "humainement effective", selon Gödel, puisque

- 1) la mise en œuvre de la mécanique inférentielle requiert une technologie électronique sophistiquée,
- 2) les entités sur lesquelles cette mécanique s'applique sont le plus souvent des couples de nombres
- 3) le résultat de la procédure se donne également sous la forme d'un couple, (même si dans certains choix de formalisation peut être considéré comme un intervalle).

C'est aussi un "calcul" qui, en général, n'est pas "vérifonctionnel" (ou purement syntaxique). Un calcul est dit vérifonctionnel lorsque la valeur de vérité d'un énoncé ne dépend que des valeurs de vérité des éléments qui le composent et non des éléments eux-mêmes. Un calcul qui n'est pas vérifonctionnel présente une forme particulière de dépendance au contexte. Plus précisément, lorsqu'il s'agit de formaliser les procédures de jugement en situation d'incertitude, c'est la configuration du contexte qui oriente le choix des définitions des procédures de calcul et, partant, qui détermine la syntaxe ultérieure de la formalisation, c'est-à-dire la machinerie inférentielle permettant de "modéliser" le jugement.

Enjeux théoriques et politiques liés à ces décisions épistémologiques

Pour clarifier ce point, il faut revenir sur « ce sur quoi fait fond » le jugement épistémique (au sens large), le premier moment de notre grille. C'est ce que désigne le terme anglais « *evidence*⁵⁶ », ce que Bernoulli appelle « cas » et que je nommerai ici « indice » : ce qui paraît soutenir l'indexicalité, la dépendance au contexte. Ces indices ne sont pas eux-mêmes toujours complètement assurés (à la différence du modèle keynésien requérant des connaissances certaines) : ils font individuellement l'objet d'une confiance inégale, partielle, instable. De plus, leur configuration peut prendre des formes différentes selon les situations (être simplement juxtaposés, fortement liés par des liens logiques, emboîtés, etc.). Par ailleurs, la nature du lien entre l'indice et l'énoncé peut être direct, détourné, etc. (Bernoulli en a donné des exemples). Ensuite, le plus souvent, le travail d'examen de la

⁵⁴ D. Dubois et H. Prade, *Théorie des possibilités...*, op. cit.

⁵⁵ G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

⁵⁶ Traduit par « mise en évidence » ou « évidence factuelle » par exemple dans l'ouvrage de I. Hacking, *L'émergence de la probabilité*, Seuil, 2002.

confiance que l'on peut accorder à un énoncé v se mène « en soupesant le pour et le contre », c'est-à-dire que, à la manière d'un procès judiciaire⁵⁷, l'on constitue un dossier à charge et un dossier à décharge. On enquête simultanément sur v et $non\ v$. Enfin, l'enquête sur la validité des énoncés v et $non\ v$ peut ne pas pouvoir être menée jusqu'au bout : l'urgence de la décision, par exemple, peut l'interrompre. Si on accepte à présent de formaliser l'aboutissement de cette démarche, au moment de la suspension de l'enquête, on aura accordé un poids, un degré de confiance à chacun de ces deux énoncés : p_1 exprimera la confiance accordée à v , p_2 à $non\ v$ ⁵⁸ si bien qu'en général, c'est un couple $(p_1, 1-p_2)$ qui exprimera notre confiance dans un énoncé v . Si nous avons à comparer la confiance que l'on peut accorder à deux propositions différentes, v et v' , il y a évidemment des cas où clairement le couple $(p_1, 1-p_2)$ pourra être dit supérieur au couple $(p'_1, 1-p'_2)$. Mais R^2 ne peut pas être ordonnée de manière compatible avec les opérations usuelles, si bien que dans le cas général, ces deux couples peuvent très bien ne pas pouvoir être ordonnés. Il y a donc à « décider », faire un choix, un pari, certes informé par ces données, mais qui ne s'y réduit pas. C'est dans cet espace de la décision qu'apparaît l'illusoire prétention à calculer le monde. Un tel effort de formalisation de la production du jugement, même en ayant simplifié drastiquement la description des bases de connaissance, invite à restituer à la décision politique partagée toute sa place : l'expertise ne peut remplacer, par le calcul, la responsabilité politique.

Les théories des possibilités – je nomme ainsi improprement l'ensemble des modélisations de l'incertitude moins exigeantes au niveau axiomatique que la théorie des probabilités, et dont la mécanique calculatoire est donc plus complexe – forment aujourd'hui cette avancée théorique au plus près, aujourd'hui, des procédures effectives de jugement : leur « mise à l'épreuve » par des enquêtes empiriques promet d'être féconde. Mais, surtout, ce qu'il importe de noter ici, est que ce remaniement n'est pas un simple élargissement de la perspective. C'est une profonde mutation de point de vue.

Dans le cadre de la théorie des probabilités, on considère implicitement que l'on peut distinguer ce qui est un avantage de ce qui constitue un danger⁵⁹, et qu'on peut leur affecter un poids en fonction de leur degré de probabilité ou de vraisemblance. Calculer l'espérance dans ce cas signifie que l'on prétend adopter une position de **surplomb** à partir de laquelle la décision s'impose⁶⁰. Nous ne sommes pas loin du paradoxe. L'invention du calcul des probabilités qui voulait apporter au dogmatisme de la causalité universelle les correctifs nécessaires, se mue ici en un nouveau dogmatisme, éclipsant à son tour la complexité des problèmes.

Les théories des possibilités ouvrent à nouveau l'espace de décisions que la maxime de Port-Royal avait fermé. Un même énoncé ou événement peut être envisagé à la fois comme un avantage **et** un danger, un obstacle **et** une circonstance opportune. Aucun calcul ne dispense d'un jugement. Les quantifications élaborées ne sont plus qu'une aide à la décision. Celle-ci ne découle plus simplement de la délibération issue des « documents » qui l'étayaient. La décision n'est plus l'application de recommandations déduites des expertises techniques, elle est irrémédiablement un pari, requalifiant ainsi le politique, et plus encore l'espace public.

D'une part, on ne peut plus être totalement assuré que les préférences que l'on s'apprête à faire valoir, toujours fondées sur des informations incomplètes, resteront les meilleures à nos yeux. Elles peuvent évoluer à l'écoute d'informations complémentaires (apportées par d'autres) et de l'exposé des préférences des autres : on a dès lors affaire à des citoyens et non à des individus. D'autre part, aucune position privilégiée, aucun calcul ne garantit la décision issue de la délibération. À la position de surplomb se substitue l'obligation de l'**engagement**, ouvrant sur un espace public d'expérimentation politique et de responsabilité.

⁵⁷ La métaphore du procès judiciaire a ses limites : une argumentation pour gagner, pour faire valoir son point de vue et une argumentation réglée sur la recherche de la vérité, ne sont pas réglées sur les mêmes visées.

⁵⁸ Sans perte de généralité, on peut supposer que ces deux nombres sont compris entre 0 et 1.

⁵⁹ Arnault et Nicole, *La logique ou l'art de penser*, Flammarion, 1970, p. 427.

⁶⁰ Notons que cette opération de quantification de l'incertain est celle qui a permis de découpler peu à peu l'éthique de l'action économique.