

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

Herausgeber: Vereinigung Schweizerischer Versicherungsmathematiker

Band: 60 (1960)

Artikel: Sur les fondements de l'actuariat et plus généralement des applications
des probabilités

Autor: Féraud, Lucien

DOI: <https://doi.org/10.5169/seals-966778>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 04.12.2024

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Sur les fondements de l'actuariat et plus généralement des applications des probabilités

Par Lucien Féraud, Genève

Pour autant que les propositions des mathématiques se rapporteront à la réalité elles ne seront pas certaines et pour autant qu'elles sont certaines elles ne se rapporteront pas à la réalité. Einstein

Résumé

Les fondements des applications des probabilités ont toujours été l'objet d'études et même de controverses auxquelles ont participé philosophes, mathématiciens, techniciens de diverses spécialisations. L'auteur reprend d'abord, en la développant, la présentation introduite dans deux articles antérieurs. En partant d'une « expérience » et de la distribution associée à ses résultats, on ne peut obtenir une conclusion ayant un intérêt pour la conduite de l'expérimentation qu'en ayant recours — explicite ou implicite — à la « convention terminale de Cournot ». Celle-ci doit être énoncée avec précaution et accompagnée de sa « règle d'application ».

De la convention de Cournot peut être tiré le principe de Bernoulli et par suite la définition de la « valeur d'une opération aléatoire ».

La convention et le principe sont l'objet d'un examen critique qui fait une place aux objections qui leur ont été opposées.

Dans une troisième partie, sont de nouveau considérés les arbitraires qui interviennent dans la convention de Cournot et il est montré que la connaissance de ceux-ci équivaut à la connaissance des « règles empiriques » auxquelles font appel toutes les techniques. Ainsi apparait la possibilité de transmettre par une méthode rationnelle les règles empiriques considérées souvent comme le produit d'une accumulation individuelle d'expériences antérieures qui échapperait à toute analyse et serait par suite essentiellement intransmissible.

Des exemples montrent que ces considérations épistémologiques ne sont pas étrangères au domaine de l'actuariat mais qu'on les rencontre immédiatement dès que l'on s'interroge sur la valeur et la portée des méthodes classiques.

L'actuariat peut paraître, à première vue, un recueil de formules qui auraient la propriété, magique, de s'appliquer à l'assurance. On dit et l'on enseigne que ces formules sont fondées sur le calcul des probabilités et, pour préciser un peu, sur la loi des grands nombres. Quelque réflexion appuyée sur une connaissance élémentaire de la théorie des probabilités montre que tout n'est pas dit en si peu de mots. On arrive

vite à mettre en cause l'application des probabilités non seulement à l'actuariat mais encore à toute activité, à toute expérimentation, en d'autres termes on est conduit à étudier les fondements des applications des probabilités. Ces fondements ont été l'objet de nombreuses analyses — dont nous rappelons quelques-unes — et même de nombreuses controverses. Nous voudrions montrer que l'étude de ces fondements n'a pas seulement un intérêt épistémologique, ou même philosophique, mais qu'elle peut avoir des répercussions sur le domaine essentiellement pratique qu'est l'actuariat.

Le calcul des primes comme le calcul des réserves mathématiques ne sont autre chose que des calculs d'espérances mathématiques. Nous montrerons que le recours à l'espérance mathématique trouve sa justification dans le «principe de Bernoulli». On remarque toutefois que le principe de Bernoulli conduit à un jeu équitable et par suite interdirait à toute personne de souscrire une assurance qui est toujours un jeu inéquitable. Un raisonnement dû à Daniel Bernoulli permet de sortir de cette difficulté en introduisant la notion d'«utilité».

Du point de vue de l'assureur le principe de Bernoulli n'est pas suffisant: il faut de plus que l'assureur n'ait pas à craindre des écarts qui mettraient sa trésorerie en difficulté. Le problème de la «ruine du joueur» montre précisément que si le jeu est équitable, le joueur courra à sa ruine s'il joue indéfiniment (la probabilité de ruine tend vers 1 lorsque le nombre des parties augmente indéfiniment). Ainsi d'un côté l'assuré devrait exiger que le jeu soit équitable et de l'autre l'assureur ne peut accepter qu'un jeu inéquitable en sa faveur. Cet apparent paradoxe se résout facilement et l'assureur devra s'arranger pour que sa probabilité de ruine — dont la définition doit être précisée — soit suffisamment petite pour qu'il puisse la considérer comme nulle en appliquant la «convention de Cournot» qui lui permet d'exclure, de l'ensemble des possibles, l'éventualité de la ruine.

Ces deux fondements principaux, convention de Cournot et principe de Bernoulli, étant rappelés nous serons amenés, par notre analyse du raisonnement probabilitaire, à serrer de plus près la pratique en envisageant les règles empiriques qui constituent une part importante de la technique de l'actuariat comme de toute technique. Nous nous efforcerons de montrer que ces règles, qu'elles soient administratives, commerciales, financières, ne sont souvent en dernière analyse que des applications de la convention de Cournot ou du principe de Bernoulli.

Nous ne chercherons pas à les énoncer systématiquement en termes probabilitaires — ce qui serait fastidieux et souvent sans intérêt — mais, en sens inverse, nous essaierons de montrer que ces règles permettent de préciser les indéterminées qui subsistent dans la convention de Cournot. Il peut toutefois y avoir un grand intérêt à donner de l'une ou l'autre de ces règles un énoncé probabilitaire lorsqu'il s'agit de les comparer ou de les discuter en vue de les rendre plus précises. Enfin et surtout ces règles n'apparaîtront plus comme le résultat d'une expérience individuelle intransmissible et l'on disposera d'une méthode rationnelle pour conserver et propager ce qu'ont appris de nombreuses expériences antérieures.

Le présent travail ne vise pas à épuiser l'étude des fondements des applications des probabilités mais seulement à indiquer une voie par laquelle elle paraît pouvoir être approchée. Et encore dans cette voie, qui sera strictement limitée par diverses restrictions, il ne s'agira ici que des deux étapes qu'il paraît opportun de considérer tout d'abord, l'une étant la convention de Cournot, l'autre le principe de Bernoulli. Ce n'est que lorsque ces positions auront été conquises et consolidées que l'on pourra aborder l'étude, beaucoup plus subtile, des problèmes d'un autre type, ceux qui relèvent de l'inférence statistique.

I.

La convention terminale de Cournot

Depuis quelques années on trouve fréquemment mention dans les recherches sur les principes et les fondements des probabilités ¹⁾ du principe ou du paradoxe de Cournot ou encore de Buffon-Cournot.

¹⁾ Pour ne citer que quelques œuvres collectives :

- i) Enquête internationale relative à l'estimation statistique des paramètres. Proceedings of the International statistical conferences. Vol. III. Part A. International statistical Institute. Washington 1947.
- ii) *Dialectica* vol. 3, n° 9/10, en particulier p. 69.
- iii) Actes du Congrès international de Philosophie des sciences (1949). Rapport général sur les travaux de la section du Calcul des probabilités par Maurice Fréchet. (Ce rapport est reproduit pp. 205-241 dans l'ouvrage du même auteur publiée aux Presses Universitaires de France, Paris, sous le titre «Les mathématiques et le concret».)
- iiii) Quatrièmes entretiens de Zurich 1953. *Dialectica* nos 28, 29 et 30.

D'autre fois, la même substance est donnée sans être désignée par un nom d'auteur. De tout ceci il résulte d'une part que les probabilistes ne sont pas tous convaincus de l'intérêt de recourir à un principe ayant en substance le contenu dont il s'agit et d'autre part que parmi ceux qui sont favorables à son adoption des divergences subsistent quant à son énoncé ¹⁾.

Ayant pour ma part, dans un article ²⁾ paru en 1948, fait appel à un énoncé du même principe – sans le désigner par un nom d'auteur – et ayant touché à quelques-unes des difficultés que soulève son application je m'efforce, dans le présent travail, de reprendre l'examen de la question, d'analyser plus complètement les modalités d'application du principe notamment en précisant les précautions qu'il est indispensable de prendre. J'aboutis à un énoncé différant à peine de celui que j'avais écrit dans l'article précédent, mais je lui adjoins une «règle d'application» et j'appelle le tout la «convention terminale de Cournot».

1° Préliminaires

Dans les Principes de la philosophie, quatrième partie, nos 205 et 206, Descartes distingue deux sortes de certitude et définit l'une d'elle en ces termes: «La première est appelée morale, c'est-à-dire suffisante pour régler nos mœurs, ou aussi grande que celles des choses dont nous n'avons point coutume de douter touchant la conduite de la vie bien que nous sachions qu'il se peut faire, absolument parlant, qu'elles soient fausses.» Dans cette phrase on voit déjà apparaître l'idée qu'une conclusion peut être fort utile tout en étant fausse, «absolument parlant»; on sait, et on verra dans la suite, qu'il en est ainsi des conclusions que l'on tire des raisonnements probabilitaires ³⁾.

La valeur des conclusions probabilitaires n'a cessé d'être un objet de préoccupations pour les probabilistes. Il suffit ici de citer quelques noms bien connus d'auteurs qui ne sont pas récents: Jacob Bernoulli (1654-1705), Nicholas Bernoulli (1687-1759), Daniel Bernoulli (1700-1782), Bayes (décédé en 1761), Buffon (1707-1788), Hume (1711-1776), Laplace

¹⁾ Voir plus loin pp. 79-80.

²⁾ Revue de Métaphysique et de Morale, avril 1948, pp. 113-138, voir en particulier p. 122.

³⁾ L'adjectif probabilitaire sera constamment employé, par exemple l'expression «raisonnement probabilitaire» est mise pour «raisonnement fondé sur les probabilités».

(1749–1827), Cournot (1801–1877). Dans leurs œuvres revient constamment l'analyse de la notion de probabilité et de son utilité; la discussion de «paradoxes» est symptomatique d'imperfections dans l'exposé de la théorie.

Au XX^e siècle on a réussi à établir le calcul des probabilités sur une base axiomatique: les promoteurs ont été d'une part von Mises (Kollektivtheorie) et d'autre part Kolmogoroff.

L'axiomatisation de la théorie des probabilités a montré que les difficultés auxquelles on s'était achoppé n'étaient pas du domaine des mathématiques puisqu'il était possible de construire d'une manière parfaitement cohérente une théorie mathématique des probabilités. En dépit de ce résultat fondamental la question était loin d'être résolue. Pendant le même temps la statistique¹⁾ avait progressé notamment pour s'appliquer aux cas où le nombre des observations est petit et par suite de nouveaux doutes surgissaient, de nouvelles discussions étaient entamées, une théorie de l'«inférence statistique» prenait naissance. On fut ainsi conduit à introduire une nouvelle axiomatique dont le point de départ se trouvait plus près des applications pratiques: les travaux poursuivis dans cette direction sont l'objet d'une analyse approfondie dans le livre de Léonard-J. Savage²⁾, qui en donne également une riche bibliographie.

Cet aperçu fort sommaire apporte pleine confirmation à une des conclusions de A. Lalande³⁾.

«La notion de probabilité reste donc une des plus obscures et des plus mal définies au point de vue du rapport entre la logique et la pratique.»

En si peu de mots on ne saurait mieux résumer la difficulté essentielle qui préoccupe les probabilistes depuis plusieurs siècles.

Il peut être intéressant de remarquer que la même difficulté se manifeste sous une autre forme si l'on propose de rechercher ce qui caractérise l'objet des raisonnements probabilitaires. H. Poincaré⁴⁾ a

¹⁾ Il paraît impossible, et de plus inutile, de distinguer les applications des probabilités de celles de la statistique; elles seront donc conjointement l'objet du présent travail.

²⁾ The foundations of Statistics. Wiley, New York 1954.

³⁾ Vocabulaire technique et critique de la philosophie. Cinquième édition. Presses universitaires de France, 1947: à l'article «Probabilité».

⁴⁾ Calcul des probabilités, Gauthier Villars, Paris 1912, 2^e édition. Voir l'Introduction.

cherché la définition du hasard dans la petitesse des causes, c'est-à-dire dans une disproportion entre l'effet et la cause ou même dans la multiplicité des causes ou encore dans une division du monde en tranches que la faiblesse de notre esprit impose mais que la réalité ne respecte pas; H. Bergson ¹⁾ voit dans le hasard l'entrée en jeu d'un intérêt humain; pour J. Piaget et B. Inhelder ²⁾ le hasard est essentiellement irréversible. Ces divergences montrent bien, de nouveau, que la question est délicate, que le sujet n'est pas épuisé et nous aurons à y revenir.

Il ne faudrait pas que ce qui précède puisse donner l'impression que l'étude des fondements de la statistique n'a qu'un intérêt philosophique, épistémologique et n'a pour but qu'un exposé plus satisfaisant à l'esprit, peut-être même qu'une présentation plus élégante ³⁾.

Fréchet ⁴⁾ formule en ces termes la question que l'on ne peut manquer de se poser:

«Comment expliquer la vitalité, la puissance de cette théorie quand on la voit s'édifier sur une base si contestée, si fragile, où l'on ne s'accorde même pas sur la signification de ce que l'on calcule?»

Il ajoute immédiatement:

«C'est là une question qui ne peut manquer de frapper le grand public mais qui ne devrait surprendre aucun philosophe, aucun chercheur. L'un et l'autre savent que si les résultats des sciences sont solides leurs fondements sont partout sujets aux plus violentes critiques.»

Il est clair que les «usagers» de la statistique, ceux qui ont pour mission d'en tirer des applications n'ont pas en principe à se soucier des discussions sur les fondements. Il me semble toutefois que les applications des probabilités sont souvent si simples, d'apparence si intuitives, si intimement mêlées à la recherche scientifique et même à la vie de tous les jours que les divergences sur les fondements se réper-

¹⁾ Les deux sources de la morale et de la religion, Editions Skira, Genève, p. 142.

²⁾ La genèse de l'idée de hasard chez l'enfant. Presses universitaires de France, 1951. Chap. X. Conclusions.

³⁾ Il n'en reste pas moins qu'une des difficultés majeures de cette étude tient à ce qu'il n'est guère possible de distinguer ce qui n'a qu'un intérêt purement pédagogique de ce qui, touchant aux bases scientifiques, peut avoir des répercussions pratiques, conduire à des nouvelles applications et améliorer celles qui sont déjà pratiquées.

⁴⁾ Les mathématiques et le concert. Presses universitaires de France, 1955, p. 157. Voir aussi tout le chapitre II de cet ouvrage consacré au calcul des probabilités et à ses applications.

cutent jusqu'au stade des applications pour engendrer ce sentiment bien connu d'insécurité, de doute, de scepticisme qui se manifeste si fréquemment à l'égard de la statistique. Il n'y a pas si longtemps que l'on prônait encore le «flair du statisticien» et l'on n'a pas renoncé à faire appel au sens commun. Je crois que le recours à des facteurs vagues est une gêne plus ou moins consciente pour la plupart de ceux qui ont à se servir de la statistique, surtout lorsqu'ils s'en servent au niveau des applications élémentaires.

Il y a plus encore. Le prodigieux développement des applications des probabilités depuis quelques dizaines d'années n'est pas allé sans soulever des discussions, des polémiques, sur la validité de telle ou telle méthode nouvelle ¹⁾. On peut même dire que la signification qu'il convient d'attribuer à certaines de ces méthodes ne saurait être exposée sans que soit précisé le point de vue que l'on adopte quant aux fondements: telle conclusion exacte peut devenir fausse si l'on déplace le point de vue. On ne saurait trop préciser les conditions et les limites dans lesquelles une méthode est applicable surtout lorsque ceux qui l'appliqueront n'auront pas à connaître toute la théorie qui la justifie — ce qui est le cas le plus fréquent.

Le plus souvent l'étude des fondements se confond avec la recherche de la meilleure manière de définir la probabilité. On sait que cette recherche a donné lieu à de nombreux travaux ²⁾ dont l'aboutissement actuel est la coexistence de différents points de vue (objectif, subjectif, nécessaire) entre lesquels chacun paraît libre de se décider soit pleinement soit avec des nuances et des réserves.

N'ayant pas l'intention de traiter de la définition de la probabilité il ne me paraît pas indispensable de faire de prime abord un choix entre les différents points de vue et je vais m'efforcer au contraire de montrer que ce choix peut être différé.

En d'autres termes je voudrais montrer que l'on peut conduire l'étude des fondements des applications des probabilités de telle sorte

¹⁾ Cf. Savage, loc. cit. p. 1.

²⁾ Avec les ouvrages de Fréchet et Savage déjà cités voir de nombreuses publications de de Finetti et en particulier: La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, vol. VII, 1937, pp. 1-68. — Le vrai et le probable. *Dialectica*, Zurich, vol. 3, n° 9/10, pp. 78-92. — Recent suggestions for the reconciliations of theories of probability. *Proceedings of the Second (1950) Berkeley Symposium on Mathematical statistic and Probability (J. Neyman editor)*, University of California Press 1951, pp. 217-226.

que toute discussion subtile sur le concept de probabilité soit repoussée *in fine* et par conséquent ne présente plus toute l'importance qu'il faut nécessairement lui attribuer si on l'aborde au point de départ.

2° Expérience et ensemble probabilisé

L'objet du présent travail n'étant ni la théorie des probabilités ni un quelconque des chapitres de la statistique mais le passage de ceux-ci aux applications nous nous placerons de prime abord dans le cadre de l'expérimentation ce qui veut dire que l'on se préoccupe de procurer un auxiliaire utile à l'expérimentateur en faisant appel au raisonnement probabilitaire. On peut penser, si l'on préfère, à l'expérimentation proprement dite, c'est-à-dire à un ensemble coordonné d'expériences provoquées, mais on pourra lui assimiler l'observation dirigée, qui ne s'en distingue pas essentiellement.

Par expérience, on désigne tantôt un événement entièrement précisé, tantôt tous les événements qui se sont produits, qui se produiront ou même qui peuvent se produire dans certaines conditions; dans le second cas le concept d'expérience est déterminé par l'ensemble des conditions imposées; on peut même dire qu'il équivaut à cet ensemble de conditions. Nous emploierons «expérience» au second des sens qui viennent d'être indiqués et, quand nous aurons à revenir au premier sens, nous dirons «une réalisation effective de l'expérience».

Ayant pris une expérience pour point de départ il est naturel, et même en principe indispensable, de commencer par procéder à une énumération complète des résultats possibles ¹⁾. A l'expérience est maintenant associé l'ensemble de ses résultats possibles. Le recours aux probabilités consiste à affecter une mesure à chacun des sous-ensembles de l'ensemble des résultats possibles et à adopter cette mesure pour définir la probabilité du sous-ensemble ²⁾: cette opération peut encore être dite la probabilisation de l'ensemble des résultats possibles. De la sorte on associe à l'expérience considérée l'ensemble de ses résultats possibles avec sa probabilisation ce qui forme un ensemble probabilisé ou en d'autres termes une distribution de probabilités.

¹⁾ Cf. «de faire partout des dénombrements si entiers et des revues si générales que je fusse assuré de ne rien omettre». Discours de la Méthode.

²⁾ Selon un usage universellement adopté la probabilité de l'ensemble des résultats possibles lui-même est égal à l'unité.

Jusque là on n'a fait qu'associer à l'expérience considérée un appareil mathématique mais rien n'annonce une conclusion susceptible de jouer un rôle dans l'expérimentation c'est-à-dire pouvant être utile à l'expérimentateur.

3° Problèmes du type C

En premier lieu il me paraît essentiel et même indispensable de distinguer deux types de problèmes : dans le premier que nous envisagerons (type C) les probabilités seront complètement connues, la probabilisation sera achevée et dans le deuxième (type A) au contraire, la probabilisation sera inachevée ou même complètement indéterminée, les probabilités ne seront pas complètement connues ou même seront entièrement inconnues et l'on s'efforcera soit de les déterminer soit de s'en passer.

Il est clair qu'en supposant les probabilités connues on supprime une difficulté – et même la difficulté majeure – mais il est tout à fait opportun de commencer par traiter le cas le plus facile. Il ne s'agit pas toutefois de problèmes hypothétiques car il existe bien des problèmes réels dans lesquels on peut considérer que les probabilités sont connues : par exemple dans les jeux de hasard où tout le mécanisme matériel a été aménagé, avec grand soin, pour que les probabilités soient bien déterminées. Aux jeux de hasard peuvent être assimilés tous les problèmes de recherche expérimentale – et ils sont nombreux – que l'on traite en introduisant un schéma emprunté à un jeu de hasard. On pourrait même dire d'une manière beaucoup plus générale que dans tout problème on cesse, à partir d'un certain moment, de discuter la probabilisation et par conséquent on entre dans une phase qui relève exclusivement du premier des deux problèmes qui ont été distingués : par exemple, en actuariat, le choix de la table de mortalité à adopter est en général l'objet de longues discussions, voire de recherches spéciales mais tôt ou tard, si l'on veut aboutir à une conclusion on arrêtera le choix d'une table de mortalité qui cessera d'être discuté.

La discussion que nous établissons entre les deux types de problèmes s'impose si l'on cherche à dégager dans le raisonnement probabilitaire ce qui joue un rôle analogue à celui que joue une loi dans le raisonnement certain. On peut établir une analogie satisfaisante entre l'application d'une loi certaine et la conclusion que l'on peut obtenir

dans un problème du type C. Dans un problème du type A au contraire la «loi» n'est pas connue et l'on doit dans le domaine certain comme dans le domaine probabilitaire recourir à une induction amplifiante. Ce point de vue a été plus longuement exposé dans un article antérieur ¹⁾).

Il est bien clair cependant qu'il sera indispensable ultérieurement par un travail de synthèse de considérer simultanément les problèmes des deux types et que la distinction que nous suggérons ne peut avoir qu'un caractère provisoire. Il paraît tout de même avantageux, au moins pour commencer et d'un point de vue didactique, de distinguer et même d'opposer deux démarches de la pensée: selon l'une on part de probabilités connues et l'on cherche à tirer quelque chose de cette connaissance, selon l'autre on est à la recherche d'une ou de plusieurs probabilités encore inconnues.

Le présent travail a trait exclusivement aux problèmes du type C, c'est-à-dire que l'on y supposera constamment que l'on part d'une probabilisation complètement déterminée.

4^o Convention de Cournot et sa règle d'application

Si l'on veut cependant aboutir à une conclusion dont l'expérimentateur pourra se servir, qui devra pouvoir être formulée avant une réalisation de l'expérience, qui devra permettre de faire l'économie de cette réalisation, il faut introduire quelque chose de nouveau. Ce sera la «convention de Cournot» que nous énoncerons ainsi:

«On considère qu'un sous-ensemble – que l'on distingue des autres – de l'ensemble des résultats possibles auquel est associée une probabilité „suffisamment petite“ est vide, c'est-à-dire qu'aucun résultat n'apparaîtra dans ce sous-ensemble.»

Cet énoncé atteint le but que nous nous étions proposé: la conclusion qu'il fournit – aucun résultat dans le sous-ensemble – peut être utile dans la conduite de l'expérimentation.

Nous devons évidemment examiner les critiques dont il peut être l'objet, dont certaines apparaissent à première vue, et c'est ce que nous allons faire dans les paragraphes qui suivent. Nous commencerons par l'assujettir à une «règle d'application».

¹⁾ Revue de métaphysique et de morale, avril 1948, p. 113 et suivantes.

On aperçoit immédiatement que l'application répétée de cette convention conduirait à déclarer vides des unions de sous-ensembles auxquelles seraient associées des probabilités qui ne seraient pas petites — on pourrait même souvent aller jusqu'à recouvrir tout l'ensemble des résultats possibles. Il faut donc, de toute nécessité, limiter l'application de la convention de Cournot.

Dans le travail antérieur¹⁾ déjà cité, l'auteur du présent article a relevé le caractère «terminal» de la dite convention à laquelle on ne doit recourir que tout à fait à la fin du raisonnement et une seule fois. Il semble bien que Destouches atteint le même but en parlant d'une «règle d'application»²⁾. En utilisant cette notion on devrait adjoindre à la convention de Cournot sa règle d'application: «*La Convention de Cournot ne doit être appliquée que tout à fait à la fin du raisonnement et une seule fois.*»

Dans cet énoncé «tout à fait à la fin du raisonnement» signifie que l'application de la convention de Cournot conduit à une conclusion que l'on est libre d'accepter ou de rejeter mais que si on l'accepte on ne doit tirer aucune autre conclusion de l'ensemble probabilisé considéré jusque là. Nous reviendrons au n° 10 sur le caractère de la conclusion que l'on tire de la convention de Cournot.

Ainsi, l'application de la convention de Cournot ne présente plus d'inconvénient si l'on ne perd pas de vue son caractère terminal, c'est-à-dire sa règle d'application. Toutefois l'énoncé de cette convention et surtout la manière de l'appliquer sont si peu conformes aux habitudes qu'elles ne peuvent manquer de surprendre et de soulever de multiples critiques — auxquelles nous allons nous efforcer de répondre dans les paragraphes qui suivent.

5° Autres énoncés

Dans le Rapport Général sur les travaux de la section du calcul des probabilités, déjà cité³⁾, Fréchet donne du principe de Buffon-Cournot les trois énoncés suivants:

A. Un événement dont la probabilité est très petite n'a lieu que très rarement.

¹⁾ Revue de Métaphysique et de Morale. Avril 1948, pp. 112–124. Voir aussi *Dialectica*, vol. 8, n° 2, 1954, p. 141.

²⁾ Cf. *Dialectica*, vol. 8, n° 2, 1954, p. 133.

³⁾ Voir page 71.

- B. Un événement de probabilité très petite est un événement «pratiquement impossible» – on refusera de croire à sa réalisation.
- C. Quand un événement est de probabilité extrêmement petite il convient d'agir comme s'il ne devait pas se produire.

Le défaut d'un accord unanime sur un énoncé est précisément un des motifs qui nous a poussé à reprendre l'étude de la question. Un autre motif est le danger que ce principe soit considéré comme évident par une confusion verbale qui serait en quelque sorte un jeu de mots.

En effet si en s'appuyant sur la correspondance étroite, le parallélisme, qui existe entre le langage de la théorie et le langage de la pratique, on confond

«un événement dont la probabilité (sens scientifique) est très petite»,
avec

«un événement dont la probabilité (sens vulgaire) est très petite»

et si l'on remarque que dans le langage courant la dernière proposition équivaut à

«un événement qui ne se produit que rarement» on arrive à

«un événement dont la probabilité (sens scientifique) est très petite ne se produit que rarement» et même à

«on peut admettre qu'un événement dont la probabilité (sens scientifique) est très petite ne se produit pas».

Cette dernière proposition a un sens mais elle n'est fondée sur rien autre que la confusion des deux sens scientifique et vulgaire du mot probabilité: on applique ainsi sans s'en rendre compte la convention de Cournot.

6° Caractère conventionnel

Si l'on objectait le caractère conventionnel il suffirait de répondre qu'il est à peu près évident que de l'ensemble probabilisé on ne peut tirer que les probabilités de tels ou tels de ses sous-ensembles mais jamais une conclusion de la nature de celle que nous cherchons permettant de prévoir le résultat de l'expérience et par conséquent de faire l'économie de celle-ci. Il est donc aisé d'admettre que pour atteindre notre but il est nécessaire d'adjoindre à l'ensemble probabilisé quelque chose qui n'est pas imposé par l'expérience et qui par conséquent aura un caractère conventionnel.

7° Deux arbitraires figurent dans l'énoncé

Dans l'énoncé de la convention de Cournot figurent deux arbitraires: l'un le choix du sous-ensemble que l'on distingue et l'autre la probabilité qui sera considérée «suffisamment petite».

Ces arbitraires sont inévitables mais souvent on ne les met pas aussi nettement en évidence qu'il a été fait ci-dessus et ils passent inaperçus. En effet dans de nombreux cas on ne va pas jusqu'à appliquer la convention de Cournot, et l'on se contente de tout préparer pour son application: on met en évidence une famille de sous-ensembles dont les probabilités sont de plus en plus petites et on s'arrête là, laissant à l'expérimentateur le soin de mettre le point final à l'argumentation, au niveau qu'il désire, c'est-à-dire au degré de probabilité qu'il lui plaît d'adopter et dont on lui laisse la responsabilité (presque tous les calculs sont conduits pour un degré de probabilité non spécifié et chacun est libre de lui donner la valeur qu'il juge opportune). De même le choix d'un sous-ensemble que l'on distinguera ne peut être soumis à aucune règle logique, il ne peut être guidé que par l'intérêt de la conclusion que fournit la convention de Cournot: on choisira le sous-ensemble qui conduit à la conclusion la plus intéressante pour l'expérience dont il s'agit.

L'existence de ces deux arbitraires caractérise le raisonnement probabilitaire et le distingue essentiellement du raisonnement non probabilitaire. A la différence de celui-ci, celui-là n'aboutit pas à partir de prémisses données à une conclusion, unique, nécessaire, contraignante; il s'ensuit que dans le domaine probabilitaire les conclusions n'admettent pas leurs réciproques et que l'on ne peut remonter des conclusions aux prémisses que par les méthodes de l'«inférence statistique» sur lesquelles un accord unanime est loin d'être réalisé.

8° Contradiction

La convention de Cournot introduit en fait une contradiction: en effet le sous-ensemble que l'on distingue pour le déclarer vide appartient par définition à l'ensemble des résultats possibles et la convention a précisément pour but de l'en exclure. Ce caractère contradictoire ne saurait être nié. Il a été nettement mis en évidence aux Quatrièmes

entretiens de Zurich notamment par G. Hirsch ¹⁾. On peut se demander si cette difficulté n'est pas plus générale et la manifestation d'une «difficulté plus profonde, de principe: celle de conjuguer la théorie à ses applications» que mentionne F. Gonseth dans ses Remarques préliminaires ²⁾. Cette question, évidemment, appartient à la théorie de la connaissance et nous nous contenterons ici de la poser. Nous remarquerons toutefois que la présence d'une contradiction exige que des précautions soient prises et par conséquent explique et justifie l'aspect formel qui a été donné à la convention de Cournot et à sa règle d'application.

On pourrait se demander s'il ne serait pas préférable de ne pas commencer par introduire le sous-ensemble que l'on excluera ensuite. La réponse est négative car ce sous-ensemble joue un rôle (provisoire) important en rendant plus facile la probabilisation. Il est d'ailleurs d'un usage constant en science de faire intervenir provisoirement des éléments qui disparaissent ensuite.

9° Morcelage

L'énoncé de la convention de Cournot tel qu'il a été donné au n° 4 implique la connaissance de l'«ensemble probabilisé des résultats possibles». Il est bien clair que la notion de résultat possible, la définition de l'ensemble des résultats possibles et la probabilisation de celui-ci soulèvent une foule de questions dont l'examen constituerait déjà une partie essentielle de l'analyse des fondements. Pour le moment, dans le présent travail, nous pouvons laisser de côté l'examen de ces questions en raison des importantes limitations que nous avons introduites, à savoir:

- au n° 2 ci-dessus nous avons admis, que la connaissance de l'ensemble des résultats possibles, était comprise dans la connaissance des conditions de l'expérience,
- au n° 3 ci-dessus nous avons décidé de nous borner systématiquement aux problèmes du type C dans lesquels la probabilisation est par hypothèse entièrement déterminée.

¹⁾ Sur un aspect paradoxal de la théorie des probabilités. *Dialectica*, vol. 8, n° 2, pp. 125 et suivantes.

²⁾ *Dialectica*, vol. 8, n° 9, p. 36.

Il est toutefois une objection que l'on ne peut éluder parce qu'elle vise directement les conclusions que l'on obtient en appliquant la convention de Cournot. Le concept d'expérience introduit au n° 2 et qui est un de nos points de départ essentiels implique immédiatement une fragmentation de la réalité. On peut dire que chaque expérience est une « parcelle » où sont groupés des « faits » qui constituent l'ensemble des résultats possibles ; bien entendu ces parcelles peuvent avoir des parties communes ou même s'emboîter, elles peuvent être spatiales ou temporelles ou spatio-temporelles.

La Convention de Cournot qui apparaît ensuite au n° 4 ne peut évidemment être appliquée qu'à l'intérieur d'une parcelle, son caractère terminal ne saurait être admis qu'ainsi limité.

Ainsi à deux reprises nous avons été amené à introduire une décomposition en parcelles, un « morcelage ».

Il faut admettre qu'à première vue cette constatation est fort inquiétante. A la réflexion elle le devient beaucoup moins.

Tout d'abord l'idée du morcelage est implicitement admise lorsque l'on nomme « univers » (ainsi qu'on le fait souvent) l'ensemble des résultats possibles.

En outre quelques citations vont montrer que d'excellents auteurs y ont eu recours :

De H. Poincaré ¹⁾ :

« Notre faiblesse ne nous permet pas d'embrasser l'univers tout entier et nous oblige à le décomposer en tranches. »

...

« deux mondes, généralement étrangers l'un à l'autre. »

Dans A. Lalande ²⁾ se trouve le terme « morcelage » et une discussion de ses caractères et de ses conséquences.

M. Merleau-Ponty ³⁾ fait appel à des « totalités partielles » :

« Nous sommes obligés d'introduire, dans notre image du monde physique, des totalités partielles sans lesquelles il n'y aurait pas de lois et qui sont précisément ce que nous entendions plus haut par forme. »

¹⁾ Loc. cit. p. 11.

²⁾ Loc. cit. p. 210-211.

³⁾ La structure du comportement. Presses Universitaires de France, 1942, p. 185.

Les «small worlds» introduits par J.-L. Savage ¹⁾ sont également à rapprocher de la notion de parcelle.

Il est bien certain que la décomposition de la totalité des faits en parcelles, en d'autres termes, en expériences, ne peut être assujettie à aucune règle. La construction des parcelles dépend donc d'un énorme arbitraire beaucoup plus large que les deux arbitraires mis en évidence au n° 6 ci-dessus. Ceci ne saurait surprendre lorsque l'on a remarqué que le choix d'une parcelle fait partie de l'opération de «conceptualisation» ²⁾ et par conséquent dépend d'une initiative de l'esprit ³⁾.

Comme toute opération de conceptualisation le choix d'une parcelle n'est guidé que par la préoccupation d'atteindre une conclusion aussi efficace que possible pour la recherche que l'on poursuit.

Enfin tout l'arbitraire laissé à la décomposition en parcelles est fort avantageux du point de vue de la souplesse de l'argumentation et de la richesse des applications. Nous citerons encore H. Poincaré ⁴⁾:

«En résumé c'est notre esprit qui fournit une catégorie à la nature. Mais cette catégorie n'est pas un lit de Procuste dans lequel nous contraignons violemment la nature, en la mutilant selon que l'exigent nos besoins. Nous offrons à la nature un choix de lits parmi lesquels nous choisissons la couche qui va le mieux à sa taille.»

10°

Les paragraphes qui précèdent tendent à faire accepter la nécessité d'un morcelage à laquelle conduit le point de vue que nous avons adopté.

On est ainsi conduit, en résumé, aux notions et termes suivants:

on part d'une «expérience» donnée par ses «conditions» et dont on peut définir «l'ensemble des résultats possibles»;

on passe à une «parcelle» en probabilisant l'ensemble des résultats possibles de telle sorte qu'une parcelle est un ensemble de résultats possibles d'une expérience, aux éléments duquel sont associées des probabilités;

¹⁾ Loc. cit. p. 82.

²⁾ Cf. Lalande, loc. cit. p. 210-214 et pour le cas des probabilités: L. Féraud «Le raisonnement fondé sur les probabilités». Revue de Métaphysique et de Morale, 53^e année, n° 2, avril 1948.

³⁾ «On étend sur toute la nature un réseau d'idées qui l'assimilent à l'esprit.» Lalande, loc. cit. p. 211.

⁴⁾ Des fondements de la géométrie, 1921, p. 64.

l'ensemble ainsi considéré et sa probabilisation définissent une distribution qui est l'«image» de la parcelle (deux parcelles différentes peuvent avoir la même image).

les opérations du calcul des probabilités – ou du calcul des distributions – consistent précisément dans le passage d'une image à l'autre.

A l'aide de ces prémisses le rôle de la convention terminale de Cournot peut être envisagé comme suit :

on peut
ou bien appliquer dans une parcelle la convention terminale de Cournot et ultérieurement considérer comme certaine la conclusion ainsi obtenue,
ou bien refuser ce qui précède, englober la parcelle dans une autre parcelle et se trouver pour cette nouvelle parcelle devant la même alternative.

Ainsi apparaît le véritable caractère de la convention terminale de Cournot dont on ne peut tirer qu'une conclusion provisoire et plus précisément dialectique car on doit toujours être prêt à l'abandonner pour recourir à l'inclusion dans une parcelle plus étendue. Bien entendu cette inclusion, toujours possible, ne doit pas être un simple artifice, elle doit procurer quelque avantage soit du point de vue théorique soit du point de vue des vérifications expérimentales.

On a donc vu que le passage du «probable» au «certain» ne peut être que provisoire, toujours susceptible d'être révisé et par suite que le «certain» est de la même nature qu'une hypothèse de travail. Si l'on adopte ce point de vue, en théorie de la connaissance, la convention terminale de Cournot n'a plus rien qui puisse surprendre.

II.

Le principe de Bernoulli

1° Le principe

Au chapitre précédent a été mis en évidence le rôle fondamental de la «convention terminale de Cournot». On n'a pas prétendu toutefois que cette convention, à elle seule, suffisait à justifier toutes les applications des probabilités, même en se limitant aux problèmes du type C dans lesquels il est admis que les probabilités sont parfaitement connues.

Le théorème de Bernoulli que l'on appelle encore «loi des grands nombres» permet d'affirmer que la répétition, un grand nombre de fois, d'une expérience dont le résultat est aléatoire, c'est-à-dire à laquelle on associe une distribution, donne un résultat moyen qui a une petite probabilité de s'écarter sensiblement du premier moment de la dite distribution ¹⁾.

Cet énoncé prépare l'application de la convention de Cournot ²⁾ en mettant en évidence l'ensemble formé par les valeurs non voisines du premier moment auquel n'est affecté au total qu'une petite probabilité. Il suffit de distinguer cet ensemble et de lui appliquer la convention de Cournot pour conclure qu'il est vide et par conséquent pour aboutir à de nouveaux énoncés :

La moyenne arithmétique des résultats obtenus en répétant l'expérience est voisine du premier moment de la distribution dont on part. On peut encore dire que cette moyenne arithmétique sera d'autant plus voisine du premier moment que le nombre des répétitions sera grand ou encore, en d'autres termes, elle pourra d'autant mieux être confondue avec le premier moment que le nombre de répétitions sera plus grand.

Ces énoncés résultent de la loi des grands nombres à laquelle on superpose la convention de Cournot.

Par «principe de Bernoulli» nous désignerons un autre énoncé :

la valeur d'une opération aléatoire est égale au premier moment de la distribution qui lui est associée,

ou en d'autres termes :

la valeur d'une opération aléatoire est égale à son espérance mathématique ³⁾.

¹⁾ Nous pouvons ici nous contenter de cet énoncé; on trouve dans tous les traités de probabilités des énoncés plus précis, plus complets et plus généraux, sous le nom — lorsqu'il y a lieu de préciser — de «loi faible des grands nombres».

²⁾ On ne pourrait se dispenser de recourir à cette convention qu'en attribuant l'expression «ensemble de résultats ayant une petite probabilité» non pas le sens que nous lui avons attaché «ensemble de résultats ayant été affecté d'une petite probabilité» mais le sens «ensemble de résultats qui ne se produiront que si rarement qu'on peut admettre qu'ils ne se produiront pas». Ce *glissement de sens* est souvent considéré comme intuitif mais en l'analysant on constate qu'il équivaut à l'application de la convention de Cournot.

³⁾ Le recours systématique à l'espérance mathématique apparaît dans le mémoire fondamental de Jacob Bernoulli: *Ars Conjectandi* 1713.

Bien que dans les applications on tire de l'un ou l'autre de ces énoncés les mêmes conclusions, il peut être intéressant de s'arrêter quelque peu à la notion de valeur d'une opération aléatoire.

Nous désignerons par «opération aléatoire» une expérience dont le résultat est un gain positif ou négatif x_i auquel est associée une probabilité p_i (on pourrait encore dire une loterie). Soit O une opération aléatoire et $D = \begin{Bmatrix} x_i \\ p_i \end{Bmatrix} i = 1, 2, \dots, n$ la distribution qui lui est associée de premier moment μ et de variance V .

La répétition n fois de l'opération aléatoire O sera dite une «répétition aléatoire» et notée $O^{(n)}$: on conviendra, de plus, de dire, pour abrégé, que le résultat de cette répétition est la moyenne arithmétique des résultats obtenus dans les n expériences. A cette moyenne arithmétique s'associe une distribution $D^{(n)}$ de premier moment μ et de variance $\frac{V}{n}$.

Ainsi, en vertu de la loi des grands nombres et de la convention de Cournot on peut dire lorsque n est grand que le résultat de $O^{(n)}$ est voisin de μ .

Soient deux opérations aléatoires O_1 et O_2 auxquelles sont associées des distributions D_1 et D_2 de premiers moments μ_1 et μ_2 et de variances V_1 et V_2 . Par définition la somme de ces deux opérations sera l'opération aléatoire qui a pour résultat la somme des résultats de chacune des deux opérations, à laquelle s'associe la distribution D_s , dite de la somme des deux variables aléatoires, de premier moment $\mu_s = \mu_1 + \mu_2$ et de variance $V = V_1 + V_2$ si nous supposons que les deux opérations O_1 et O_2 sont indépendantes. L'opération aléatoire ainsi définie comme la somme de O_1 et de O_2 sera notée O_s ou $O_1 + O_2$.

A la répétition de O_s c'est-à-dire à la répétition aléatoire $O_s^{(n)}$ sera associée une distribution $D_s^{(n)}$ qui aura pour premier moment μ_s et pour variance $\frac{V_1 + V_2}{n}$. En vertu d'abord de la loi des grands nombres et ensuite de la convention de Cournot, pour n grand, le résultat de $O_s^{(n)}$ sera voisin de μ_s .

D'une manière analogue la différence des deux opérations O_1 et O_2 sera une opération aléatoire qui aura pour résultats la différence des résultats et à laquelle sera associée une distribution D_d de premier moment $\mu_d = \mu_1 - \mu_2$ et de variance $V_d = V_1 + V_2$. L'opération aléatoire ainsi définie comme la différence de O_1 et de O_2 sera notée O_d ou

$O_1—O_2$. A la répétition de O_d , c'est-à-dire à la répétition aléatoire $O_d^{(n)}$, sera associée une distribution $D_d^{(n)}$ de premier moment μ_d et de variance $\frac{V_1 + V_2}{n}$. En vertu, toujours, de la loi des grands nombres et de la convention de Cournot, pour n grand, le résultat de $O_d^{(n)}$ sera voisin de μ_d .

Nous nous proposons maintenant d'attacher à chaque opération aléatoire O un nombre qui représentera sa valeur soit $v(O)$. A quelles conditions doit satisfaire $v(O)$?

a) Il est naturel de poser par définition que $O_1 \geq O_2$ suivant que le résultat de la répétition de O_d , c'est-à-dire le résultat de $O_d^{(n)}$, sera voisin d'un nombre $\alpha \geq O$.

On devra ensuite définir $v(O)$ de telle sorte que $O_1 \geq O_2$ entraîne $v(O)_1 \geq v(O)_2$. Il résulte de ce qui précède qu'il en est bien ainsi si l'on prend $v(O) = \mu$ pour toute opération O .

b) Il est encore naturel d'exiger que l'opération O_s , qui a été définie comme la somme des deux opérations O_1 et O_2 , ait pour valeur la somme des deux valeurs de O_1 et de O_2 , c'est-à-dire

$$v(O_s) = v(O_1) + v(O_2).$$

Il en est encore ainsi si l'on prend $v(O) = \mu$ pour toute opération O .

Ayant trouvé une solution $v(O) = \mu$ satisfaisant aux conditions posées nous allons montrer qu'il n'en existe pas d'autre.

Soit F une autre solution. En vertu de la condition a) on aurait lorsque $O_1 = O_2$ à la fois $\mu_1 = \mu_2$ et $F_1 = F_2$ ce qui nécessite que F soit une fonction de μ seulement: $F(\mu)$. En vertu de a) toujours, lorsque $O_1 < O_2$ on doit avoir à la fois $\mu_1 < \mu_2$ et $F_1 < F_2$, c'est-à-dire $F(\mu_1) < F(\mu_2)$, ce qui nécessite que F soit une fonction croissante de μ . En vertu de b) si l'on considère $O_s = O_1 + O_2$ on doit avoir $\mu_s = \mu_1 + \mu_2$ et $F(\mu_s) = F(\mu_1) + F(\mu_2)$ c'est-à-dire $F(\mu_1 + \mu_2) = F(\mu_1) + F(\mu_2)$ et par suite $F(\mu) = m\mu$ ou m doit être positif puisque F est une fonction croissante de μ .

Nous assujettissons maintenant la valeur attachée à l'opération aléatoire à une condition supplémentaire:

c) l'opération aléatoire à laquelle est associée la distribution

$$\begin{vmatrix} x' & x \\ \alpha & 1 - \alpha \end{vmatrix}$$

de premier moment $\mu = \alpha x' + (1 - \alpha) x$ a pour valeur, d'après ce qui précède,

$$m\mu = m[\alpha x' + (1 - \alpha) x].$$

On s'impose de retrouver, pour $\alpha = 0$, la valeur de l'opération certaine à laquelle se réduit l'opération aléatoire d'où $m x = x$ et $m = 1$.

On a donc l'unique solution $v(0) = \mu$.

Tout ceci aboutit à un résultat bien connu que se trouve, sous une forme ou sous une autre, dans tous les traités. Il nous reste, ainsi qu'il a été fait au chapitre précédent à propos de la convention de Cournot, à passer en revue les critiques qui ont été adressées au principe de Bernoulli ainsi que les arguments en sa faveur.

2^o Première critique. Notion d'utilité

Les difficultés de l'application du principe ont été aperçues par Daniel Bernoulli ¹⁾ qui les formule à peu près en ces termes dans le «mémoire de St-Pétersbourg» ²⁾:

«Un pauvre hère est en possession — on ne sait trop comment — d'un billet de loterie qui peut gagner 20 000 ducats ou rien du tout avec des probabilités égales pour les deux éventualités. Cet homme évaluera-t-il ses chances de gagner à 10 000 ducats? Serait-il mal avisé de vendre son billet pour 9000 ducats? Il me semble que la réponse est négative. Par contre je croirais volontiers qu'un homme riche serait mal avisé s'il refusait d'acheter le billet pour 9000 ducats. Aussi, si je ne me trompe pas, il semble clair que la même règle ne peut être suivie par tout le monde pour estimer la valeur du billet dont il est question.»

Le célèbre paradoxe de St-Pétersbourg, qui figure dans le même mémoire, conduit aussi à l'impossibilité d'appliquer le principe de Bernoulli. De nombreux autres exemples pourraient être invoqués: ce principe proscrirait tous les jeux inéquitables, et par suite toutes les assurances.

Daniel Bernoulli répond à l'objection en substituant, dans la définition de l'élément aléatoire, à la quantité, par exemple d'unités moné-

¹⁾ Un neveu de Jacob, cité au paragraphe précédent.

²⁾ Ainsi nommé parce qu'il parut dans les Commentaires de l'Académie impériale des sciences de St-Pétersbourg 1738. Le texte original est en latin, une traduction anglaise par D^r Louise Sommer a paru dans *Econometrica*, January 1954.

taires, une notion nouvelle l'«utilité»: dans sa terminologie, par cette substitution l'espérance mathématique devient l'«espérance morale»¹⁾.

L'introduction de la notion d'utilité a une telle importance en économétrie qu'elle a été l'objet de multiples exposés, commentaires et aussi qu'elle a soulevé de nombreuses discussions. On s'est accordé assez facilement pour admettre une utilité ordinale, c'est-à-dire susceptible d'indiquer un ordre de préférence. La possibilité de définir une utilité cardinale, c'est-à-dire à laquelle il est possible d'affecter une mesure, s'est heurtée à une résistance bien plus forte. Il est bien évident que la justification du principe de Bernoulli par le recours à la notion d'utilité exige que celle-ci soit cardinale.

Ce n'est qu'assez récemment que la question a été résolue, d'une manière satisfaisante²⁾, dans l'ouvrage si remarquable de von Neumann et Morgenstern³⁾. Les auteurs ont montré qu'il était possible de définir une mesure de l'utilité à une transformation linéaire près et en même temps, ont formulé un système d'axiomes tels que tout système de préférences qui y satisfait résulte de l'application du principe de Bernoulli à partir des utilités numériquement définies. D'autres systèmes d'axiomes ont été proposés par divers auteurs⁴⁾. D'autres auteurs mettent en doute la possibilité d'appliquer l'utilité ainsi définie au comportement réel⁵⁾. Il n'y a pas lieu ici de passer en revue les travaux⁶⁾ sur cette question ou d'entreprendre sa discussion; nous admettrons

¹⁾ Cette expression est tombée en désuétude: dans tous les cas on dit «espérance mathématique» étant entendu qu'elle est calculée avec un système d'utilités qui peuvent, en particulier, coïncider avec les quantités d'unités monétaires.

²⁾ Tous les auteurs ne s'accordent pas à la reconnaître: voir en particulier, H.Wold, *Demand Analysis*, New York Wiley 1953, p. 127. Voir aussi un exposé remarquable des différents points de vue, accompagné de références bibliographiques, dans L.-J. Savage, *The foundations of statistics*. New York Wiley, 1954, pp. 21-104.

³⁾ *Theory of games and economic behaviour*. Third edition 1953. Princeton University Press. Chap. 1, § 3 et Appendix.

⁴⁾ Voir en particulier L.J. Savage, loc. cit.

⁵⁾ Voir M. Allais, *Fondements d'une théorie positive du choix comportant un risque et critique des postulats et axiomes de l'école américaine*. Voir *Colloques internationaux du Centre National de la Recherche Scientifique*. XL. *Econométrie* 1952. Dans ce volume on trouve encore de nombreuses autres communications portant du moins partiellement sur la notion d'utilité. Se reporter également à *Symposium on cardinal utilities* *Econometrica* 20.

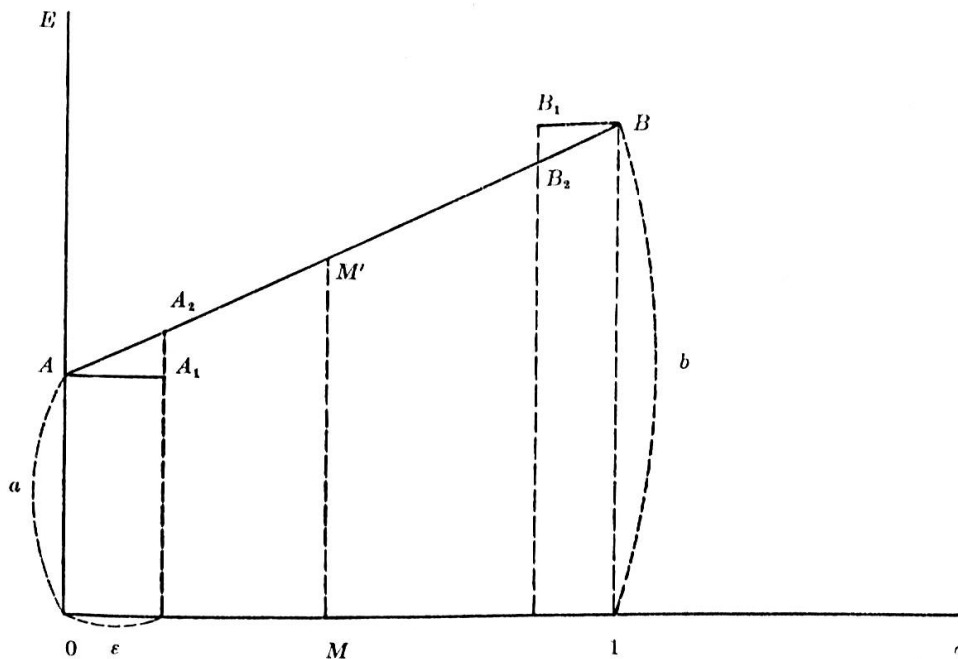
⁶⁾ Voir K.J. Arrow, *Utilities, Attitudes, Choices: a Review Note*. *Econometrica*, January 1958.

donc qu'une utilité numérique peut être définie de telle sorte que le principe de Bernoulli continue à être applicable. Ainsi est sauvegardée la validité du principe ¹⁾).

3° Deuxième critique. Attitudes

L'application du principe de Bernoulli et celle de la convention terminale de Cournot pourraient paraître contradictoire.

Soit, par exemple, une expérience qui ne comporte que deux résultats possibles a et b auxquels sont associées respectivement les probabilités $1-\alpha$ et α . L'espérance mathématique du résultat $E = a(1-\alpha) + b\alpha$ est représentée par la droite AB dans la figure ci-dessous.



Le principe de Bernoulli conduit à prendre, pour toute probabilité α , le segment $MM' = E$ comme valeur du résultat aléatoire.

Toutefois, en appliquant la convention terminale de Cournot on peut, lorsque $\alpha < \epsilon$, considérer que la probabilité de b est suffisamment

¹⁾ Le principe de Bernoulli complété par l'introduction de la notion d'utilité, c'est-à-dire fondé sur l'espérance morale, au lieu de l'espérance mathématique peut être appelé le principe de Bernoulli-Cramer en hommage au mathématicien genevois que cite Daniel Bernoulli dans une note adjointe au «mémoire de St-Petersbourg» où il reconnaît que «cet auteur a formulé une théorie sur le même sujet plusieurs années avant son travail». Cf. L.-J. Savage, loc. cit. p. 92.

petite pour que l'on puisse faire de b le sous-ensemble que l'on distingue et conclure en faveur de a , ce qui donnerait a comme valeur du résultat aléatoire.

Ainsi à l'extrémité du segment AB (le même raisonnement pourrait être fait pour l'autre extrémité) on obtient pour valeur du résultat aléatoire :

AA_2 par le principe de Bernoulli,

AA_1 par la convention terminale de Cournot.

La réponse à cette objection se trouve dans la première partie du présent travail : on applique dans les deux cas la convention de Cournot mais les sous-ensembles que l'on distingue sont différents et nous avons vu que le choix du sous-ensemble distingué était entièrement libre.

Il est intéressant de remarquer que, dans les applications pratiques, le choix du sous-ensemble distingué définit une *attitude*. Dans l'exemple ci-dessus

le choix qui conduit à AA_2 peut être dit rationnel, et

le choix qui conduit à AA_1 peut être dit prudent ou pessimiste

et à l'autre extrémité,

le choix qui conduit à BB_2 peut être dit rationnel ¹⁾ et

le choix qui conduit à BB_1 peut être dit téméraire ou optimiste.

On arrive ainsi à établir une correspondance entre les différents sous-ensembles que l'on peut distinguer et les différentes attitudes que l'on peut adopter.

Un autre exemple est fourni par les différentes attitudes que peut adopter chaque joueur dans un jeu à deux personnes, notamment

- a) éviter le pire par une stratégie pure,
- b) chercher à atteindre le meilleur par une stratégie pure,
- c) atteindre le *maximin* de son espérance mathématique par une stratégie mixte.

¹⁾ Par définition on dit rationnel le choix qui conduit à la conclusion fournie par le principe de Bernoulli. Nous reviendrons sur le caractère rationnel de ce choix page 94).

4° *Troisième critique. Justification à la longue et intérêt immédiat*

La troisième critique est peut-être la plus évidente.

La justification du principe de Bernoulli par la convention de Cournot a pour point de départ la répétition un grand nombre de fois de l'événement dont il s'agit et la conclusion n'est valable que pour une longue suite de répétitions: on peut dire qu'elle est valable «à la longue». Le principe de Bernoulli ne serait donc pas applicable dans les cas où l'on sait que l'expérience ne sera pas répétée un grand nombre de fois, par exemple, à un jeu où l'on ne joue qu'une fois. Or, même dans ces cas on rencontre de nombreux exemples de son application.

L'exemple des opérations viagères est particulièrement instructif:

- l'assureur, en prenant pour prime unique pure l'espérance mathématique de ses engagements, applique le principe de Bernoulli et peut se contenter d'une conclusion valable à la longue puisqu'il conclut un grand nombre de contrats,
- l'assuré au contraire n'a que faire du principe de Bernoulli puisqu'il ne joue qu'une seule fois – cependant en fait il accepte la prime pure proposée par l'assureur,
- deux particuliers qui ne sont assureurs ni l'un ni l'autre concluent entre eux une opération viagère, par exemple cession d'un immeuble contre une rente viagère, qui ne peut avoir lieu entre eux qu'une fois: le principe de Bernoulli est absolument sans justification et cependant l'opération se fait en prenant pour estimer la valeur de la rente viagère l'espérance mathématique tout comme si les deux parties étaient des assureurs.

Ainsi on constate que bien que le recours à l'espérance mathématique ne soit justifié qu'à la longue il tend à s'étendre aux cas où il n'y a aucune répétition ou un petit nombre seulement. Lorsqu'il en est ainsi on se comporte devant une seule réalisation de l'expérience de la manière qui serait justifiée si l'on se trouvait devant une longue répétition; on peut encore dire, en revenant aux termes introduits au paragraphe précédent, que l'on adopte devant une seule réalisation de l'expérience l'attitude qui serait rationnelle «à la longue». Bien entendu ceci n'est pas une règle générale ¹⁾ mais une constatation que l'on peut observer

¹⁾ Il ne faudrait pas tomber dans une sorte de fétichisme du principe de Bernoulli: par exemple dans une seule partie de pile ou face avec mise unité la valeur du jeu est 0 si je gagne, 2 si je perds et l'espérance mathématique 1 n'est

fréquemment. Elle n'admet évidemment aucune justification logique et son explication ne peut être que psychologique. Je proposerai la suivante :

A côté du but principal poursuivi il existe un ou plusieurs buts secondaires — qui restent plus ou moins dans l'inconscient ; ces buts secondaires se rapportent à de longues répétitions et la conclusion qui vise l'ensemble de tous les buts est encore justifiée par le principe de Bernoulli. Par exemple, dans le cas de la rente viagère, ci-dessus considérée, l'une et l'autre partie bien que n'ayant en vue qu'une seule transaction préfèrent que celle-ci s'effectue selon une norme plus généralement admise. Dans d'autres cas, on tient à dégager sa responsabilité en prenant l'attitude qui serait rationnelle à la longue. Ou encore on poursuit en outre du résultat immédiat un but plus lointain par exemple pédagogique : éduquer sa main, ses réflexes, prendre de bonnes habitudes, etc. En résumé, le comportement ne dépend pas de plus souvent d'un but unique mais de plusieurs buts qu'une analyse un peu poussée permet de distinguer ; l'analyse des buts se relie d'ailleurs étroitement à l'analyse des attitudes, le choix d'une attitude ne pouvant être arrêté, le plus souvent, qu'en prenant en considération les buts secondaires ¹⁾.

5° Conclusion

Ayant fait leur part aux considérations psychologiques, à tout ce que peut apporter l'analyse des buts et des attitudes, il semble possible d'admettre que les applications du principe de Bernoulli ²⁾ sont admises parce qu'elles peuvent être justifiées « à la longue ». De la sorte elles relèvent encore de la convention terminale de Cournot et l'on peut conclure que celle-ci, soit directement comme au chapitre I soit par l'intermédiaire du principe de Bernoulli, est la seule justification que l'on puisse retenir pour servir de fondement à toute solution de tout problème probabilitaire du type C.

qu'une indication théorique dont l'intérêt est bien moindre que le couple (0,2) dont se préoccupe exclusivement celui qui ne connaît pas l'espérance mathématique.

¹⁾ On pourrait adjoindre ces problèmes psychologiques à ceux qu'ont posés Fréchet, Un problème psychologique sur les probabilités subjectives irrationnelles, *Journal de psychologie pure et appliquée*, oct.-déc. 1954, p. 431, et de Finetti, Les problèmes psychologiques sur les probabilités subjectives, *Journal de psychologie pure et appliquée*, avril-juin 1955, p. 253. Il serait intéressant d'étudier les circonstances dans lesquelles est adoptée une attitude téméraire, rationnelle ou prudente, et les buts secondaires qui interviennent.

²⁾ Et en particulier toutes les stratégies mixtes dans la théorie des jeux.

III.

Le recours aux problèmes du type C

Nous étant efforcé dans les chapitres précédents d'analyser les conclusions qui permettent d'énoncer les problèmes du type C, nous allons maintenant envisager les circonstances dans lesquelles se posent ces problèmes et sont intéressantes leurs conclusions. Nous remarquerons tout d'abord que les décisions les plus courantes de la vie et par conséquent de toute expérimentation reposent sur des règles que nous admettons parce que nous les avons apprises ou acquises par l'expérience ou encore par analogie sans qu'intervienne consciemment aucun raisonnement probabilitaire. Nous admettons par exemple qu'une maladie est grave sans calculer la probabilité des cas mortels ou encore qu'un acte est dangereux sans recourir à l'estimation du risque d'accident. Alors même que nous connaîtrions, par la statistique, la proportion de cas mortels dans les observations antérieures de la maladie considérée nous adopterions cette proportion — qui serait une espérance mathématique — comme la base de nos raisonnements — ce qui reviendrait à appliquer, sans s'en rendre compte, le principe de Bernoulli ¹⁾ (en admettant que la statistique qui est à notre disposition constitue un échantillon suffisamment représentatif mais, pour le moment, nous laissons cette difficulté dans l'ombre car elle est du ressort des problèmes du type A). Il est bien certain que les règles empiriques pratiquées, dont nous disposons et qui ne sont comme l'a fortement mis en évidence David Hume que le souvenir d'expériences antérieures — pas nécessairement personnelles — suffisent à nous guider dans la presque totalité des cas où nous avons une décision à prendre. Il n'en est pas moins vrai

- a) qu'au point de vue épistémologique la justification d'une de ces règles ne peut résulter que d'une analyse probabilitaire qui serait fondée — s'il s'agit exclusivement d'un problème du type C — soit sur le principe de Bernoulli soit sur la convention de Cournot;
- b) que lorsqu'il s'agit de préciser une de ces règles c'est encore une analyse probabilitaire qui permettra de le faire.

¹⁾ Autres exemples: nous prenons comme valeur «équitable» d'un jeu son espérance mathématique, comme valeur d'une assurance sa prime unique, et plus généralement nous traitons l'espérance mathématique comme une grandeur sans penser qu'il s'agit d'une conclusion probabilitaire.

Des règles empiriques qui viennent d'être mentionnées on peut encore tirer autre chose de tout à fait important pour les applications des probabilités :

Si l'on a *toujours* constaté qu'un phénomène x est accompagné d'un phénomène y on cherchera une explication dans une loi non probabilitaire, certaine.

Si l'on a constaté que x est *souvent* accompagné du phénomène y on cherchera l'explication dans une loi probabilitaire. Mais nous ajoutons ceci : dans la plupart des cas l'observation a donné quelque précision sur la valeur du mot «souvent» qui se manifeste en général non pas par une proportion, par une probabilité, mais par une ou plusieurs règles empiriques qui nous font dans telles circonstances attendre y lorsque x s'est produit et dans d'autres ne pas l'attendre. Si nous essayions de soumettre l'une de ces règles à une analyse probabilitaire nous constaterions qu'elle équivaut à l'application de la convention de Cournot avec un degré de probabilité α et un sous-ensemble V déterminés. Nous ne vivons pas au milieu de coefficients α et d'ensemble V dont nous avons conscience, mais nous vivons dans un monde de règles empiriques dont la justification épistémologique équivaudrait à ces coefficients et à ce choix d'ensembles. Ainsi que nous l'avons déjà dit ces règles guident nos décisions dans la vie courante et encore davantage dans la technique. Toute technique comprend d'une part certaines lois que l'on peut dire scientifiques et d'autre part certaines règles empiriques pas toujours formulées avec précision mais qui constituent «l'expérience du technicien».

Toute technique, qu'elle soit mécanique, industrielle, médicale, commerciale, financière, actuarielle, administrative, comporte des règles empiriques — ce sont parfois des secrets de fabrication. Il ne s'agit pas de chercher à exprimer ces règles en termes de probabilités — ce qui le plus souvent ne servirait à rien — mais il importe de remarquer que l'on pourra tirer de ces règles, dans chaque cas particulier, au moins une indication sur les deux arbitraires α et V qui interviennent dans la convention de Cournot. Il est clair aussi qu'en interprétant les règles empiriques comme des applications soit du principe de Bernoulli soit de la convention de Cournot on est à même de les comparer entre elles : l'exemple élémentaire le plus connu est le recours au schéma des urnes qui permet de parler «d'une chance sur cent, sur mille» ainsi qu'on le fait couramment sans prétendre appliquer la théorie des probabilités.

On pourrait encore maintenir un doute sur le choix de la valeur de α qui sera considérée suffisamment petite pour l'application de la Convention de Cournot. On pourrait vouloir prendre α de plus en plus petit. Il en serait bien ainsi si en diminuant α on ne faisait surgir d'autres inconvénients. Dans la presque totalité des applications la recherche d'une certitude absolue ou même celle d'une probabilité de succès trop grande aurait pour conséquence la mise en œuvre de moyens si importants qu'ils conduiraient à différer toute action ou alors qu'ils enlèveraient tout intérêt à cette action (par exemple si l'on exigeait une probabilité de risque trop petite on devrait renoncer à tous les moyens de transport). On se trouve donc en pratique devant une certaine marge dans laquelle il faut prendre α suffisamment petit pour la sécurité et pas trop petit pour ne pas être contraint à l'immobilisme. Cette marge est connue empiriquement d'une manière plus ou moins précise, parce qu'elle est implicitement contenue dans une des règles empiriques qui guident nos actions, qui forment une partie importante de toute technique et qui, répétons-le, proviennent d'expériences antérieures.

Toute analyse, même partielle, des règles empiriques, toute possibilité de les ramener, même incomplètement, à un schéma rationnel ouvre la voie à une méthode permettant de les transmettre. On considère le plus souvent qu'il y a, d'un côté, la théorie que l'on apprend par l'enseignement oral ou écrit et, de l'autre, un domaine mystérieux de règles empiriques qui s'accumulent dans chaque individu à la suite de nombreuses expériences et qui ne sont transmissibles par aucune autre méthode qu'un contact prolongé avec les problèmes pratiques et avec les personnes qui les possèdent déjà. Il est évident que tout pas en avant qui pourrait être fait dans la direction que nous indiquons permettrait de transférer une des règles empiriques au domaine rationnel de la théorie, par suite réduirait le domaine considéré comme mystérieux en donnant la possibilité de transmettre par une méthode rationnelle, analytique, ce qui a été tiré des expériences antérieures. N'est-ce pas, en fin de compte, tout ce que l'on peut demander à la science?

Certaines et peut être la plupart de nos explications, justifications, précautions pourront paraître superflues : elles ont été conservées parce qu'il a paru essentiel dans notre mode de présentation d'éclaircir tout ce qui se rapporte aux problèmes du type C avant d'entreprendre l'étude, beaucoup plus délicate, des problèmes du type A, que nous nous proposons d'envisager dans un travail ultérieur.

Zusammenfassung

Die Begründung der Anwendungen der Wahrscheinlichkeitstheorie ist seit langem Gegenstand von Untersuchungen, ja von Kontroversen gewesen, an welchen sich Philosophen, Mathematiker und Techniker verschiedener Spezialgebiete beteiligt haben. Verfasser knüpft an die Darlegungen zweier früherer Arbeiten an und führt diese weiter aus. Wenn man von experimentellen Erfahrungen und den aus den bezüglichen Resultaten sich ergebenden Ereignisverteilungen ausgeht, so kann daraus keine für das Verhalten der Experimentierreihe aufschlussreiche Folgerung gezogen werden, ohne dass man – explizit oder implizit – auf die Regel von Cournot abstellt. Letztere muss daher mit der nötigen Vorsicht umschrieben und mit «Anwendungsvorschrift» versehen werden.

Aus der Cournotschen Regel kann das Bernoullische Prinzip und weiter die Definition des wahrscheinlichkeitstheoretischen Erwartungswertes hergeleitet werden. Verfasser unterzieht Regel und Prinzip einer kritischen Betrachtung, namentlich im Hinblick auf die gegen sie erhobenen Einwände.

Schliesslich wird nochmals auf die Unbestimmtheit eingetreten, welche der Regel von Cournot inhärent ist und gezeigt, dass das Wissen um dieselbe durchaus der bewussten Anwendung empirischer Regeln gleichkommt, auf welche jegliche Technik angewiesen ist. So erscheint die Möglichkeit, die in Betracht gezogenen empirischen Regeln durch eine rationelle Methode auf gleich gelagerte Probleme zu übertragen, oft wie das Ergebnis einer individuellen Ansammlung früherer Erfahrungen, welche einer Analyse unzugänglich sind und folglich dem Wesen nach nicht übertragbar wären.

Beispiele zeigen, dass diese erkenntnis-theoretischen Überlegungen auch dem Gebiete der Versicherungsmathematik nicht fremd sind, dass man vielmehr unmittelbar auf sie stösst, sowie man sich die Frage nach Wert und Tragweite der klassischen aktuariellen Methoden stellt.

Summary

The justification for applying the theory of probability has for a long time been the object of discussions and even controversy among philosophers, mathematicians and specialists in different fields. The author refers to the ideas which he has set forth in two previous papers and develops them further. Taking an experiment and the distribution of its results as a starting point, it is impossible to draw any conclusions about the conduct of the series of experiments without applying – implicitly or explicitly – the Cournot principle which therefore needs careful definition and a strict «rule of application».

From the Cournot principle it is possible to derive the Bernoulli principle and the definition of the value of a random operation. The author examines these principles thoroughly and mainly discusses the objections which have been raised against them.

Finally he considers the arbitrary elements intervening in the Cournot principle and shows that knowledge of these elements is equivalent to application of the empirical rules needed by every technique. Thus the possibility arises of applying efficiently

these very empirical principles – often considered as the result of accumulated previous individual experiences beyond analysis and hence essentially incommunicable – to similar problems.

Examples show, that these epistemological considerations are also applicable to the domain of actuarial science and that one encounters them directly as soon as one asks the question about the extent and value of the classical actuarial methods.

Riassunto

Le basi di applicazione delle probabilità sono sempre state l'oggetto di studio, anzi di controversie alle quali hanno partecipato filosofi, matematici e tecnici di specializzazione diversa. L'autore riprende anzitutto l'esposizione introdotta nei due articoli anteriori sviluppandola ulteriormente. Partendo da un'esperienza e dalla distribuzione associata ai suoi risultati, non si può giungere ad una conclusione avente un interesse per la condotta della sperimentazione, se non con l'aiuto – in modo esplicito od implicito – della convenzione terminale di Cournot. Questa dovrà essere espressa con cautela ed accompagnata dalla «regola d'applicazione».

Dalla convenzione di Cournot si può ricavare il principio di Bernoulli e, in seguito, la definizione del valore di una operazione aleatoria.

La convenzione ed il principio sono l'oggetto di un esame critico dove vien tenuto conto delle obiezioni sollevate contro di essi.

Infine sono nuovamente considerati gli arbitri che intervengono nella convenzione di Cournot ed è dimostrato che la conoscenza di questi equivale alla conoscenza delle regole empiriche dalle quali dipendono tutte le tecniche. Così appare la possibilità di trasmettere con un metodo razionale le regole empiriche considerate sovente come il prodotto di un'accumulazione di esperienze anteriori, le quali sfuggono a ogni analisi e sarebbero per conseguenza essenzialmente non trasmissibili.

Esempi dimostrano che queste considerazioni epistemologiche non sono sconosciute nel campo dell'attuariato, ma che ci si urta immediatamente appena interrogati sul valore e sulla portata dei metodi classici.

