

LEONARD J. SAVAGE

Difficultés de la théorie de la probabilité personnelle

Mathématiques et sciences humaines, tome 21 (1968), p. 5-9.

http://www.numdam.org/item?id=MSH_1968__21__5_0

© Centre d'analyse et de mathématiques sociales de l'EHESS, 1968, tous droits réservés.

L'accès aux archives de la revue « Mathématiques et sciences humaines » (<http://msh.revues.org/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

DIFFICULTÉS DE LA THÉORIE DE LA PROBABILITÉ PERSONNELLE *

par

Leonard J. SAVAGE, Yale University¹

(traduit par G. Morlat)

Nous, statisticiens, qui nous occupons tout spécialement de l'incertitude, sommes plus que d'autres susceptibles de rencontrer la philosophie, que cela nous plaise ou non.

Pour ma part, je m'en accommode assez bien. C'est pourquoi l'honneur m'est échu d'ouvrir cette séance de discussion, bien que mes préoccupations habituelles rendent mes connaissances et mon langage quelque peu différents des vôtres.

La théorie de la probabilité personnelle formalise une façon de voir la nature de l'incertitude, qui a été découverte indépendamment par Frank Ramsey et par Bruno de Finetti. Cette façon de voir ne s'est pas beaucoup répandue ni en philosophie ni en statistique, bien qu'elle ait récemment gagné du terrain dans ces deux disciplines. Les statisticiens, chose assez compréhensible, se montrent aisément passionnés dans la controverse sur la nature de l'incertitude. Nous, statisticiens personnalistes, sommes heureux de voir grossir nos rangs, mais nous sommes impatients vis-à-vis de la majorité, qui n'est pas encore venue à nous. Et la majorité ne peut pas comprendre pourquoi une poignée de statisticiens qui ont fait preuve de compétence dans le passé se consacrent maintenant à la propagande d'une doctrine indéfendable et pernicieuse.

Nous qui défendons le point de vue personnaliste, nous regardons naturellement comme illusoire quelques-unes des difficultés les plus frappantes qui nous sont opposées. Je parlerai ici non pas de celles inhérentes au point de vue personnaliste que je peux aisément lever, au moins pour ma propre satisfaction, mais de celles qui me gênent ou m'embarrassent particulièrement. En conséquence, vous entendrez beaucoup de questions vagues et peu de propositions claires. Presque toute difficulté à laquelle se heurte une opinion semble donner naissance à une objection et permettre de marquer un point à l'une ou l'autre des opinions concurrentes. Comme vous pourriez l'imaginer d'après ma position, les difficultés qui me troublent ne me semblent pas tellement plaider contre l'opinion personnaliste en faveur de ses adversaires, les opinions fréquentiste et logicienne (ou nécessitariste), mais me semblent plutôt suggérer qu'il y a lieu de modifier et d'améliorer le point de vue personnaliste lui-même. En fait, le point de vue personnaliste est en réalité un ensemble de points de vue plus ou moins reliés entre eux, dont l'un se trouve davantage présent à mon esprit que les autres, et quelques formes du point de vue personnaliste ont été développées précisément pour échapper à l'une ou l'autre des difficultés en cause.

J'aborderai mon affaire avec une esquisse — n'ayant pas le temps pour une véritable description, voire un développement — de la théorie de la probabilité personnelle dans la forme que je préfère pour des discussions critiques, sinon pour un exposé scolaire ou un usage pratique.

* Texte d'une conférence pour la Section occidentale de l'Association américaine de Philosophie, Chicago, 4-6 mai 1967. L'article original a été publié dans « Philosophy of Science », vol. 34, n° 4, décembre 1967.

Nous remercions la rédaction de cette revue, ainsi que l'auteur lui-même, pour nous avoir autorisé à en publier la traduction.

1. « This research was supported in part by the Army, Navy, Airforce, and the National Air and Space Administration of the United States under a contract administered by the office of Naval Research. » Je suis très reconnaissant de la patience et de la sensibilité dont G. Morlat a fait preuve dans la traduction de cet ouvrage. (Note de l'auteur.)

La théorie vise à distinguer entre le comportement cohérent et la bévue, ou l'incohérence qu'on peut faire apparaître en présence d'incertitude. Pour cela, elle impose des conditions aux préférences de quelqu'un entre les actes. La relation de préférence, « n'est pas préféré à », est supposée constituer un préordre complet sur les actes. Un acte est défini comme une fonction ou un tableau de correspondance, associant une conséquence pour la personne qui choisit, à chaque état possible du monde.

D'une manière non formelle, les conséquences représentent tout ce que la personne éprouve, ou ce pour quoi elle a des préférences, même quand il n'y a pas d'incertitude. Cette idée de l'expérience pure du bon ou du mauvais, semble philosophiquement suspecte, et n'est certainement pas opératoire. Dans les applications, le rôle de conséquence est tenu par des choses comme un paiement comptant ou une journée de pêche, dont les « conséquences réelles » peuvent être très incertaines, mais qui sont néanmoins bien adaptées au rôle de conséquences sûres dans le contexte de certaines applications. Un des mérites de la théorie personnaliste de Richard Jeffrey est d'échapper à ce concept suspect.

Un état du monde sera, toujours de façon non formelle, une liste possible de réponses à toutes les questions qui pourraient être pertinentes dans le problème de décision envisagé. Les états n'ont pas besoin d'être conçus comme absolument atomiques, mais seront considérés comme atomiques pour tel contexte ; encore cet emploi d'atomes est-il didactique plutôt que nécessaire. Donc l'énigme philosophique de propositions véritablement atomiques n'entre pas essentiellement dans la théorie de la préférence, comme cela semble être le cas dans la théorie de Carnap.

Certains postulats, qu'on ne peut malheureusement pas répéter ici, sont attribués à la relation de préférence entre les actes. Quelque chose du caractère de ces postulats peut s'exprimer par cette hypothèse : si deux actes ont les mêmes conséquences pour certains états, la préférence entre ces deux actes ne sera pas changée si on leur donne de nouvelles conséquences communes pour les états où ils avaient déjà les mêmes conséquences, chacun de ces actes restant inchangé par ailleurs.

Portons l'attention pour un moment sur deux conséquences quelconques, dont l'une est effectivement préférée à l'autre ; nous appellerons la conséquence préférée « la récompense » et l'autre « le statu quo ». Quand un acte ne peut procurer que ces deux conséquences, il en résulte qu'on obtient la récompense lorsque se produit un certain ensemble d'états (nous dirons : un événement) et qu'on garde le statu quo sinon. La préférence d'un tel acte à un autre peut bien s'exprimer en langage ordinaire, en disant que la personne tient l'événement entraînant la récompense, avec le premier acte, pour plus probable que l'événement entraînant la récompense avec le second acte. Selon une proposition déduite des postulats, il existe une fonction unique qui associe un nombre non négatif à chaque événement, et qui satisfait aux trois conditions suivantes.

- Primo, si et seulement si la personne tient un événement pour plus probable qu'un autre, le nombre associé au premier sera plus grand que le nombre associé au second.
- Ensuite, le nombre associé à la réunion de deux événements disjoints est la somme des nombres associés à chacun de ces événements.
- Enfin, et accessoirement, le nombre associé à l'événement universel, ensemble de tous les états, est égal à l'unité.

Les mathématiciens appellent toute fonction d'événements (ou d'ensembles) dans l'ensemble des nombres non-négatifs, ayant les deux dernières propriétés (additivité finie) une mesure de probabilité, sans se référer nécessairement à quelque autre aspect ou interprétation de la probabilité, tel que croyance, fréquence ou symétrie.

La mesure de probabilité associée aux événements par l'intermédiaire des récompenses est intimement liée à l'attitude et au comportement potentiel de la personne en présence d'événements incertains. Pour un personnaliste comme moi, elle est sous-jacente à la meilleure analyse qui ait jamais été donnée de la notion ordinaire de probabilité. Convenablement interprétée, elle contient les aspects valables des autres analyses, fréquentiste et logicienne par exemple. Mais le présent exposé porte en réalité davantage sur la théorie de la préférence entre les actes qui conduit à la probabilité personnelle, que sur la probabilité personnelle elle-même. Logiquement, j'aurais presque pu passer sur la probabilité personnelle comme je passerai en fait sur une autre fonction (celle-ci de l'ensemble des conséquences dans l'ensemble des nombres réels) appelée utilité, qui est finalement aussi essentielle à la théorie de la préférence que la probabilité personnelle elle-même.

La théorie de la préférence ne semble pas concerner ou contenir le problème de l'induction dans le sens : si, comment et avec quelle justification, nous en venons à croire aux propositions universelles.

Cependant, elle contient un modèle de la façon dont une opinion est modifiée à la lumière de l'expérience — un phénomène d'une grande importance — qu'on pourrait aussi appeler induction. La théorie de la probabilité personnelle semble être de même dégagée, et pas forcément par accident, de certains autres concepts difficiles qui sont traditionnellement associés à la philosophie de la probabilité, tels que la simplicité et le rasoir d'Occam, les propositions atomiques (déjà mentionnées), le temps, la cause, la permanence de la nature dans l'espace et le temps, et naturellement la notion de limite. La possibilité réelle, ou même apparente, d'échapper à ce qui a été quelquefois regardé comme des questions fondamentales, éclaire le rôle de ces questions, d'une façon dont j'espère parler à une autre occasion ¹.

La théorie de la préférence pourrait être étudiée comme un système axiomatique dépourvu d'interprétation ; mais pour les mathématiciens elle n'a qu'un intérêt limité. Elle peut être interprétée comme une théorie de la façon dont les individus, les collectivités ou autres organismes, se comportent réellement ; mais comme un chapitre de la psychologie, elle a un domaine de validité et d'emploi très limité. L'interprétation pour laquelle elle a été développée est une interprétation normative, selon laquelle une personne peut veiller sur la cohérence de ses propres décisions potentielles.

Pour faire une telle application, il nous faut structurer le problème de décision devant lequel vous vous trouvez en termes d'états du monde (chaque état étant une liste de réponses possibles à toutes les questions pertinentes), de conséquences (qui pourraient vous arriver) et d'actes possibles (les choses que vous pouvez faire, décrites à l'aide des conséquences qui vous affectent dans chaque état du monde). À la fois en pratique et en théorie le choix d'un tel cadre soulève des problèmes. Il ne faut pas attendre un choix unique, ne serait-ce que parce qu'il est en principe indifférent d'ajouter des choses non pertinentes à la liste des questions qui déterminent les états. En pratique, il est souvent — je dirais presque toujours — possible de faire un choix qui marche, c'est-à-dire pour lequel les postulats semblent raisonnablement bien satisfaits. Et on peut, dans une certaine mesure, montrer que différents choix qui marchent ne peuvent pas conduire à des comportements différents. Pourtant, dans la notion même de ce choix d'un cadre, il y a des difficultés impressionnantes. Est-il bon, ou même est-il possible, d'exiger, comme cette théorie de la préférence le fait, que les actes soient sans influence sur les événements et les événements, sans influence sur les satisfactions ? On peut s'arranger pour que de telles conventions marchent mieux qu'on ne pourrait croire à première vue, et elles apportent des avantages considérables à la théorie ; mais elles ont été mises en cause à plusieurs reprises de façon sérieuse, et peut-être avec juste raison ; par exemple, par J. Drèze et R. Jeffrey.

Dans quel sens cette théorie est-elle normative ? Elle vise à ce qu'une personne réfléchie qui se trouve sur le point d'agir en contradiction avec la théorie, reconsidère sa position, et des personnes réfléchies ont effectivement trouvé que la théorie, prise avec un grain de sel, était contraignante. Je sens, mais je ne comprends pas clairement, cette contrainte dont une bonne analyse pourrait être une contribution philosophique intéressante. Jusqu'à quel point une telle analyse serait-elle liée à la philosophie du contrafactuel ? En outre, la philosophie du libre arbitre semble avoir quelque lien avec l'interprétation de toute théorie normative ; dans le cas présent, ce rapport est-il important, académique, ou illusoire ?

L'analyse devrait prendre soin de ne pas trop prouver ; car certains écarts à la théorie sont inévitables, et certains sont louables. Par exemple, une personne à qui on demande de parier de l'argent sur une décimale de rang élevé de π devrait, pour satisfaire pleinement à la théorie, calculer cette décimale, bien que cela dût être en réalité ruineux si le coût du calcul était plus élevé que la récompense promise. Car les postulats de la théorie impliquent que vous devriez vous comporter en accord avec les conséquences logiques de tout ce que vous savez. Est-il possible d'améliorer la théorie à cet égard, en faisant place au coût de la réflexion, ou bien cela entraînerait-il un paradoxe, comme j'ai tendance à le croire tout en restant incapable de le démontrer ? ² Si le remède n'est pas dans une modification de la théorie mais plutôt dans la façon dont nous avons à l'utiliser, il reste encore à rendre les choses plus claires.

L'exemple de π n'exprime pas convenablement l'impossibilité absolue de connaître nos propres idées dans le sens voulu par la théorie. Vous ne pouvez pas être assuré d'avoir composé le télégramme de dix mots qui vous satisfait le mieux, bien que la liste des possibilités soit finie et qu'un grand nombre de possibilités puissent être éliminées immédiatement ; après tous vos efforts, quelqu'un peut suggérer une amélioration évidente qui ne vous était pas apparue.

1. « Implications of personal probability for induction », *Journal of philosophy*, 64 (1967), 593-606. Une traduction de cet autre essai de L. J. Savage paraîtra dans un prochain numéro de « Mathématiques et Sciences Humaines ».

2. À cette question, Ian Hacking répond dans un essai publié dans le même numéro de *Philosophy of Science* (décembre 1967).

Un aspect particulièrement évident de notre incapacité à connaître nos propres idées (et donc à pouvoir satisfaire à la théorie) tient à ce que nous ressentons de façon incertaine, ou vague, nos préférences pour des choix relativement simples tels qu'entre dix dollars ou deux billets de théâtre. Certains ont essayé de représenter ce caractère vague dans la théorie, cependant que d'autres pensent que, bien qu'il faille reconnaître de quelque façon ce caractère vague, sa nature défie toute formalisation.

Voici une importante illustration tirée de la statistique. L'emploi d'échantillons au hasard, si caractéristique de la statistique moderne, n'est pas compatible avec la théorie de la préférence, du moins appliquée au comportement d'un organisme simple. Même si la « randomisation » n'est pas aussi efficace que les statisticiens l'ont quelquefois pensé, elle n'en paraît pas moins utile et importante. La valeur de la randomisation, qui échappe à la théorie strictement interprétée, peut être rattachée dans une large mesure au caractère vague de nos idées et à d'autres incapacités intellectuelles à satisfaire parfaitement à la théorie de la préférence — et peut-être aussi à notre interaction avec d'autres personnes.

Supposez que vous acceptiez l'idée que votre comportement devrait se conformer raisonnablement bien à la théorie ; est-ce que la théorie a quelque valeur pour vous ? et pourquoi ? Un avantage de la théorie, à mon avis, est que toute application de l'idée de probabilité peut de façon assez claire être regardée comme une application de la théorie de la préférence, bien que naturellement il soit souvent possible d'obtenir les bénéfices pratiques de ces applications de la probabilité sans les regarder comme tels. Employer la théorie de la préférence, c'est rechercher l'incohérence parmi les décisions possibles, dont vous, l'utilisateur de la théorie, devez alors revoir une ou plusieurs. La théorie elle-même ne dit pas de quelle façon retrouver la cohérence, et il ne faut pas, je suppose, s'attendre à cela. Néanmoins, certaines préférences sont manifestement pour vous plus faciles à laisser tomber que d'autres ; dans certains cas, l'analyse philosophique pourrait tirer au clair les raisons.

D'après ce que j'ai dit, la théorie de la préférence, bien que non exactement valable, est approximativement valable. Comme tous les philosophes et mathématiciens le savent maintenant, toute contradiction dans une théorie déductive rend la théorie sans valeur. Est-il réellement possible de donner un sens rigoureux à la proposition qu'une théorie empirique ou normative ne soit pas valable, mais soit presque valable ? Avec ou sans rigueur, il nous faut vivre avec des théories approximativement valables. La philosophie peut-elle nous dire comment ?

L'idée de faits connus est implicite dans l'emploi de la théorie de la préférence. D'un côté, la personne doit savoir quels actes lui sont offerts. Si par exemple je vous demande quelles chances vous donnerez que le quatrième jet de cette pièce de monnaie donne pile si les trois premiers l'ont fait, il est normalement implicite non seulement que vous savez que je mettrai ma part de l'enjeu si nous parions, mais aussi que vous saurez qu'il y a eu trois fois pile si vous les voyez. Le statisticien parle toujours de la réaction qui serait appropriée à tel ou tel ensemble de données, ou d'observations. Cependant, les données ne sont jamais tout à fait données, parce qu'il y a toujours quelque doute sur ce que nous avons réellement vu. Naturellement, dans toute application, le doute peut être repoussé plus loin. Nous pouvons remplacer l'événement de trois fois pile par l'événement moins immédiat de trois relevés qui nous sont transmis, et prendre en compte dans notre analyse le fait que tous les relevés ne sont pas forcément corrects. Néanmoins, non seulement les propositions universelles, mais les propositions les plus concrètes et les plus personnelles ne sont jamais réellement hors de doute. En vérité, comme vous le savez mieux que moi, de telles apparentes déclarations de faits doivent réellement être regardées comme des propositions universelles. Peut-on dès lors éviter le manque de clarté et de rigueur, dans notre allusion à des faits connus ? On a avancé que puisqu'il n'y a pas de certitude absolue, nous devrions entendre par « certitude », une certitude relativement forte.

Cette recommandation est provocante mais semble davantage souligner, que résoudre la question présente.

Vous avez écouté un long catalogue de mes douleurs et de mes peines philosophiques. J'espère apprendre au cours de la discussion que certaines sont seulement dans ma tête et que d'autres iront mieux d'elles-mêmes. Mais je crains que nous ne tombions d'accord que même une éthique de l'incertitude aussi neutre et ouverte que la théorie de la préférence est inextricablement liée à bien d'autres parties de la philosophie.

BRÈVE BIBLIOGRAPHIE SUR LA PHILOSOPHIE DE LA PROBABILITÉ PERSONNELLE

Les références qui suivent portent sur le sujet dans des sens très divers, et les références contenues dans les articles cités conduisent à une littérature plus complète.

- DE FINETTI Bruno. — « La prévision : ses lois logiques, ses sources subjectives » ; *Annales de l'Institut Henri-Poincaré*, 1937.
- DE FINETTI Bruno. — « Probability : Philosophy and interpretation », *International Encyclopedia of the Social Sciences*, 1968.
- DE FINETTI Bruno et SAVAGE Léonard. — « Sul modo di scegliere le probabilità iniziali », *Sui fondamenti della statistica, Biblioteca del Metron, Series C. 1*, (1962), p. 81-47 (English summary, p. 148-151).
- DRÈZE Jacques. — « Fondements logiques de la probabilité subjective et de l'utilité » p. 73-87, in *La décision*, Paris, Centre national de la recherche scientifique, 1961.
- EDWARDS Ward, LINDMAN Harold and SAVAGE Léonard J. — « Bayesian statistical inference for psychological research », *Psychological Review* 70 (1963), p. 193-242.
- FISHBURN Peter C. — *Decision and Value Theory*, New York, Wiley, 1964.
- HACKING Ian. — *Logic of Statistical Inference*, Cambridge, Basic Books, 1965.
- JEFFREY Richard D. — *The Logic of Decision*, New York, Mc Graw-Hill, 1965.
- KYBURG Henry E. Jr. and SMOKLER Howard E. — *Studies in Subjective Probability*, New York, Wiley, 1964. Contient les mémoires fondamentaux de Frank P. Ramsey et Bruno de Finetti. Ce dernier est la traduction en anglais de l'article des *Annales de l'Institut Henri-Poincaré* (1937) cité plus haut.
- MORLAT Georges. — « Des poids et des choix », *Mathématiques et Sciences Humaines*, n° 3 (1963), p. 41-46.
- MORLAT Georges. — « Statistique et Théorie de la décision », *Mathématiques et Sciences Humaines*, n° 8 (1964), p. 1-8.
- SALMON Wesley C. — « The foundations of scientific inference », *Mind and Cosmos : Essays in contemporary science and philosophy*, vol. III, University of Pittsburgh Series in the Philosophy of Science, Univ. of Pittsburgh Press, 1966.
- SAVAGE Leonard J., et al. — *The Foundations of Statistical Inference : A Symposium*, New York, Wiley, 1962.