Comment attribuer une probabilité à un évènement ?

Hugo LAVENANT

12 février 2016

> « *Il faut réfléchir pour mesurer et non pas mesurer pour réfléchir.* »
>
> Gaston Bachelard, *La formation de l'esprit scientifique*.

Résumé

Ce document vise à discuter les différentes interprétations qui peuvent être faites du calcul des probabilités à travers la question de l'attribution d'une probabilité à un évènement : sur quels éléments, objectifs ou subjectifs, se fonde-t-on pour attribuer des probabilités à des évènements concernant des domaines aussi divers que les jeux de hasard, les phénomènes sociaux ou encore les rencontres sportives ? La loi des grands nombres, qui semble ramener toute évaluation de probabilité à une limite de fréquences, est abordée.

Introduction

Aussi surprenant que cela puisse paraître, le calcul des probabilités, c'est-à-dire la théorie mathématique traitant des probabilités, ne dit pas ce qu'*est* la probabilité d'un évènement. On ne trouvera nulle part dans un ouvrage de mathématique une preuve que la probabilité de tomber sur pile lorsque l'on joue à pile ou face est $1/2$, et pire encore : même si c'était le cas, l'ouvrage ne vous apprendrait même pas ce que cela *signifie* que la probabilité d'obtenir pile est $1/2$. Le but du calcul du probabilité est, la probabilité de certains évènements étant donnée, de calculer la probabilité d'autres évènements dépendant des premiers. Si vous dites à un mathématicien que la probabilité de tomber sur pile en jouant à pile ou face est $1/2$, il sera capable de vous dire, mieux, de vous *prouver*, que celle d'obtenir exactement deux « pile » lors de trois lancers indépendants successifs est $3/8$. Mais cela ne permet toujours pas d'apprendre ce que ce $3/8$ signifie. Qu'il n'y ait ici aucun malentendu : c'est grâce à cette limitation que le calcul des probabilités a pris son envol et a réalisé des constructions intellectuelles d'une grande beauté qui font honneur à l'esprit humain. Mais lorsqu'il est question d'utiliser les résultats du calcul des probabilités dans des situations concrètes, pour prendre des décisions ou comme argument dans un débat, la signification de la probabilité intervient inévitablement, que ce soit implicitement ou explicitement.

Ce document tentera d'apporter des éléments de réponse à la question suivante : comment peut-on attribuer une probabilité à un évènement ? C'est-à-dire, sur quelles raisons peut-on se fonder pour affirmer que la probabilité d'obtenir pile vaut $1/2$? Une définition des termes s'impose ici, au moins pour brosser un tableau grossier qui sera précisé par la suite. Un *évènement* est dit aléatoire, hasardeux, ou encore fortuit s'il y a une incertitude quant à son arrivée. Cette définition est bien sûr insuffisante et demanderait à être amplement discutée mais ce n'est pas l'objet du présent document, on supposera que chacun a une idée suffisamment claire de ce qu'est un évènement fortuit. Pour donner des exemples, c'est le cas de la pièce qui peut tomber sur pile (ou sur face), mais aussi sur le fait qu'un individu meurt dans l'année qui suit ou qu'une équipe sportive gagne sa prochaine rencontre. Une *expérience aléatoire* décrit les conditions dans lesquelles l'évènement aléatoire peut se produire ou non. Les résultats possibles de l'expérience sont les *issues*, un évènement aléatoire est donc une des issues possibles d'une expérience aléatoire. Pour reprendre les trois exemples ci-dessus, le lancer de la pièce, la donné d'un individu et la rencontre sportive sont les expériences aléatoires respectivement associées aux évènements précédents. Une question centrale est de savoir si une expérience aléatoire est reproductible, c'est-à-dire si les conditions donnant lieu à l'arrivée éventuelle d'un évènement peuvent être reproduites *toutes choses égales par ailleurs*. Enfin, la *probabilité* d'un évènement est une mesure de quelque chose comme son degré de certitude, l'objet de ce qui suit est justement d'éclairer ce que la probabilité mesure. La seule contrainte est que la probabilité d'un évènement est un nombre compris entre 0 et 1, la valeur 0 étant attribuée à un évènement impossible et la valeur 1 à un évènement certain.

Pour tenter d'éclaircir le projet par une métaphore, comparons le calcul des probabilités à la géométrie : un des buts de la géométrie est d'exprimer des relations entre des longueurs, comme le fait si bien le fameux théorème de Pythagore, mais la géométrie ne parle pas de ce qu'est une longueur et surtout de comment une longueur est mesurée. Elle ne dit pas qu'il faut utiliser une règle si la longueur à mesurer est rectiligne, une corde s'il s'agit de mesurer le périmètre d'une figure courbe comme un cercle. Elle ne dit pas les précautions qu'il faut prendre pour que la mesure de la longueur soit la plus précise possible. Par contre la géométrie peut prouver des théorèmes qui expliquent comment mesurer de façon indirecte une longueur : par exemple, pour mesurer la distance entre la Terre et la Lune, nul besoin d'essayer de construire une règle gigantesque, il suffit de mesurer différents angles sous lesquels la Lune est vue depuis la Terre et différentes longueurs terrestres pour en déduire la longueur recherchée. C'est exactement la même chose pour le calcul des probabilités, il suffit de remplacer le terme « longueur » par celui de « probabilité ». En particulier, le calcul des probabilités peut nous apprendre comment mesurer des probabilités de façon indirecte, c'est-à-dire en ramenant la mesure de probabilité d'un évènement à celle d'un autre plus facile à obtenir. La différence est que, s'il y a un consensus sur la manière de mesurer des longueurs, ce n'est pas le cas pour les probabilités.

On passera en revue trois façons d'attribuer des probabilités sans qu'il n'y ait de volonté d'exhaustivité ni de contextualisation historique ; puis l'on discutera du résultat qui semble prétendre unifier ces conceptions : la loi des grands nombres.

1 Trois façons d'attribuer des probabilités

1.1 La conception « classique »

S'il ne fallait donner qu'un seul nom pour symboliser la conception classique des probabilités, ce serait celui de Pierre Simon de Laplace (1749-1827). Pour évaluer la probabilité d'un évènement E dans une expérience aléatoire, il faut commencer par déterminer les cas (les issues pour reprendre la terminologie de l'introduction) qui sont également probables, puis définir

$$\text{Probabilité}(E) = \frac{\text{nombre de cas favorables à } E}{\text{nombre de cas possibles}}.$$

Un exemple paradigmatique est celui du lancer de dé (à six faces) : si le dé est équilibré alors chacune des faces constitue un « cas également probable », de sorte que pour calculer la probabilité d'obtenir n'importe quelle face, par exemple « cinq », il n'y a qu'un seul cas favorable parmi les six possibles et donc

$$\text{Probabilité (« cinq »)} = \frac{1}{6}.$$

Si l'évènement en question se compose de plusieurs cas élémentaires, il suffit de les dénombrer : par exemple il y a trois façons d'obtenir un chiffre pair (en obtenant « deux », « quatre » ou « six ») de telle sorte que

$$\text{Probabilité (« obtenir un chiffre pair »)} = \frac{3}{6} = \frac{1}{2}.$$

Un autre exemple abondamment utilisé est celui du tirage d'une boule dans une urne : si une urne contient p boules blanches et q boules noires toutes identiques sauf pour ce qui concerne la couleur, alors il y a $p+q$ boules donc $p+q$ cas également probables, et p cas sont favorables au tirage d'une boule blanche de sorte que

$$\text{Probabilité (« tirer une boule blanche »)} = \frac{\text{nombre de boules blanches}}{\text{nombre total de boules}} = \frac{p}{p+q}.$$

C'est-à-dire que la probabilité de tirer une boule blanche est égale à la proportion de boules blanches dans l'urne. De manière générale, le problème d'attribution des probabilités est réduit à un problème de *dénombrement*, puisqu'il suffit de compter, de dénombrer, les cas favorables et les cas possibles. Ce type de problème n'est pas toujours simple mais il a l'avantage de posséder une réponse bien déterminée et surtout déterminable grâce à des outils purement mathématiques dans lesquels la notion de hasard ou de chance est éliminée.

Mais en réalité cette définition mérite d'être éclaircie : comment peut-on déterminer les cas « également probables » ? Car si ce terme n'est pas précisé la définition est circulaire, en ce qu'elle présuppose un concept d'équiprobabilité afin d'obtenir celui de probabilité. Dans les exemples ci-dessus, et de manière générale dans les jeux de hasard, ce sont des considérations de symétrie qui permettent de conclure : si le dé est parfaitement symétrique, on ne voit pas pourquoi une des faces serait plus probable que les autres, si les boules sont identiques, on ne voit pas pourquoi l'une d'entre elles aurait plus de chance d'être choisie qu'une autre, etc. C'est parfaitement naturel, presque instinctif, et c'est pourquoi les jeux de hasard sont utilisés pour faire comprendre la signification de la probabilité d'un évènement. Ainsi, pour faire comprendre ce que veut dire, par exemple, que le taux de mortalité est de 3 pour mille, une urne contenant 3 boules noires et 997 boules blanches est introduite, et la mort d'un individu est comparée au tirage d'une boule noire dans cette urne, sa survie à celui d'une boule blanche. Il faut garder à l'esprit qu'il ne s'agit que d'une analogie, elle n'épuise pas entièrement la signification du taux de mortalité.

Mais comment faire pour aller au delà des jeux de hasard ? On pourrait dire que les évènements également probables sont ceux pour lesquels l'on dispose d'une même absence d'information, ceux sur lesquels l'on est également indécis. Cela permet de reconnaître explicitement que les probabilités sont un moyen permettant de mesurer notre incertitude, qu'elles sont un outil qui nous sert à exprimer notre manque de connaissance : les probabilités sont *épistémiques*, c'est-à-dire relatives à notre connaissance ou plus précisément à son imperfection. Le démon de Laplace [1], qui a devant ses yeux l'ensemble de ce qui s'est produit et de ce qui va arriver dans le monde, n'a pas besoin des probabilités, nous, être humains, si. Mais en réalité la question s'est déplacée à la quantification de notre manque de connaissance : comment juge-t-on que nous sommes « également indécis » ? Soit aucune information n'est à notre disposition, mais dans ce cas l'on se demande de quel droit nous prétendons quantifier ce sur quoi nous ne savons absolument rien, soit nous avons des informations qui sont de poids égal, mais il faut alors expliquer comment l'on pèse une information. C'est sans doute pour cela que les probabilités classiques ont du mal à être attribuées lorsque des considérations de symétrie ne s'appliquent pas : le dé pipé est déjà hors d'atteinte [2], alors que dire des probabilités de décès ?

Une autre critique sévère peut être faite contre les probabilités classiques, elle intervient lorsque le nombre de cas également probables est infini et donne lieu au paradoxe de Bertrand. En effet, si le nombre de cas possibles et celui de cas favorables sont infinis, le quotient définissant la probabilité n'a pas un sens univoque. Le paradoxe peut être formulé de la manière suivante : supposons que vous vous trouvez en face d'un verre contenant un mélange de 20 cl d'eau et d'une quantité inconnue d'alcool, que vous savez néanmoins comprise entre 0 et 20 cl. Quelle est la probabilité que la quantité d'alcool soit comprise entre 10 et 20 cl ? Intuitivement vous avez tendance à répondre $1/2$, puisqu'en l'absence d'information les cas « entre 0 et 10 cl » et « entre 10 et 20 cl » sont également probables. Mais le problème peut aussi être formulé de la manière suivante : la proportion (le degré) d'alcool dans ce verre est comprise entre $0°$ (s'il y a 0 cl) et $50°$ (s'il y a 20 cl). Et une quantité d'alcool entre 10 et 20 cl correspond à une proportion entre $33°$ (plus précisément $1/3$ d'alcool s'il y a 10 cl) et $50°$. Or la probabilité pour que la proportion d'alcool soit comprise entre $33°$ et $50°$ sachant qu'elle varie entre 0 et $50°$ est, suivant la méthode classique, de $1/3$ (le calcul est $(50-33)/50 \simeq 1/3$). À la même question, à savoir « quelle est la probabilité pour que la quantité d'alcool soit comprise entre 10 cl et 20 cl ? », on trouve deux réponses différentes : $1/2$ (si l'on raisonne en terme de quantité d'alcool) et $1/3$ (si l'on raisonne en terme de proportion d'alcool). Deux ré-

1. Pour Laplace, le monde est parfaitement déterministe, c'est-à-dire que l'état du futur de l'univers est entièrement déterminé par le présent. Le « démon de Laplace », avec une intelligence absolue et une connaissance parfaite uniquement de l'état présent de l'univers, est capable par le calcul de connaître l'avenir et le passé aussi clairement que le présent.

2. Il faut nuancer ce propos : historiquement, les probabilités ont été introduites au XVIIe siècle pour penser les contrats équitables en situation d'incertitude. Elles étaient donc fortement reliées à l'idée d'*équité* et permettaient de trouver le juste prix. L'idée d'un dé pipé ou plus généralement d'un jeu de hasard inéquitable était impensable du point de vue des probabilités pour l'époque.

ponses pour une même question : cela montre bien que cette dernière est mal posée, c'est-à-dire que la méthode classique est insuffisante.

N'y-a-t-il vraiment rien à sauver de la conception classique, doit-elle être jetée aux oubliettes de l'histoire ? Pas entièrement. Tout d'abord parce que, comme signalé plus haut, ce n'est pas un hasard si le lancer de dé ou le tirage d'une boule dans une urne restent les exemples les plus utilisés pour introduire la notion de probabilité, notre esprit percevant intuitivement dans ces situations le concept d'équiprobabilité puisqu'il se ramène à celui de symétrie. Mais aussi parce que des méthode modernes permettent de pousser plus loin la méthode classique en incorporant les cas où les informations à notre disposition ne nous mettent pas dans une situation d'égale indécision : l'incorporation se fait en prenant la distribution de probabilité qui encode les informations à disposition mais qui ne rajoute pas d'information supplémentaire, c'est-à-dire qui maximise le « désordre » parmi toutes celles contenant les informations à disposition[3], de telle sorte que la probabilité reflète notre connaissance et notre manque de connaissance à un instant donné.

1.2 La conception fréquentiste

Cette conception est souvent présentée comme la seule qui vaille. Si le principe sous-jacent est plutôt simple, une définition rigoureuse et une réelle exploration des conséquences de celle-ci peut être attribuée à Richard von Mises (1883-1953). L'idée de base est qu'un évènement est vu comme l'issue d'une expérience aléatoire, expérience qui définit les conditions de possibilité d'arrivée de l'évènement et qui doit pouvoir être reproduite. Lorsqu'elle est reproduite un certain nombre de fois, la fréquence d'un évènement E est définie comme le rapport du nombre d'expériences où E s'est réalisé par le nombre d'expériences total, c'est-à-dire par

$$\text{Fréquence}(E) = \frac{\text{nombre d'expériences où } E \text{ est réalisé}}{\text{nombre d'expériences total}}.$$

La définition est bien différente de celle de la conception classique : dans cette dernière, le dénombrement se faisait par rapport aux issues possibles de l'expérience, ici il se fait par rapport aux résultats d'expériences effectivement réalisées. Cette définition dépend du nombre d'expériences réalisées et il n'est pas possible d'identifier fréquence et probabilité sans tomber dans des situations paradoxales. Par exemple, si un dé est lancé et s'arrête sur la face « cinq », en ne prenant en compte que cette expérience, la conclusion est que

$$\text{Fréquence (« cinq »)} = \frac{1}{1} = 1.$$

3. La référence est faite ici au concept d'*entropie* de la théorie de l'information. Si une expérience aléatoire possède n issues deux à deux disjointes $E_1, E_2, \ldots, E_n$, pour attribuer une distribution de probabilité $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ aux issues respectives $(E_1, E_2, \ldots, E_n)$, ce que l'on fait est que l'on choisit la distribution $\mathbf{p}$ qui maximise l'entropie

$$S(\mathbf{p}) = -\sum_{k=1}^{n} p_k \ln(p_k)$$

parmi toutes les distributions de probabilité satisfaisant à certaines contraintes correspondant à l'information dont l'on dispose sur l'expérience aléatoire. Par exemple on peut imposer que l'espérance d'une variable aléatoire, calculée à partir de la distribution $\mathbf{p}$, soit une valeur fixée. Maximiser l'entropie garantie que la distribution de probabilité n'encode pas plus d'information que celle déjà contenue dans les contraintes portant sur $\mathbf{p}$.

Or une probabilité de 1 correspond à un évènement certain et ce n'est pas en faisant une seule expérience que l'on a envie de conclure que l'évènement est certain : il n'est pas possible d'identifier probabilité et fréquence. Pour pallier ce problème, il faut s'intéresser à un nombre d'expériences suffisamment grand. Si par exemple 100 lancers de dé sont réalisés et que le chiffre « cinq » sort 18 fois, alors

$$\text{Fréquence (« cinq »)} = \frac{18}{100} \simeq \frac{1}{6}.$$

Plus le nombre d'expériences réalisées est important, plus la fréquence de réalisation d'un évènement E s'approche d'une valeur limite, et c'est cette valeur qui est définie comme la probabilité de E : cette dernière est mesurée de façon de plus en plus précise en calculant les fréquences de réalisation de E pour un nombre d'expériences de plus en plus grand.

Cette contrainte d'un « grand » nombre d'expériences demande une clarification. Tout d'abord, le fait que, en faisant un grand nombre d'expériences, la fréquence de réalisation d'un évènement s'approche de plus en plus d'une valeur limite est une propriété *empirique*, qui n'est pas postulée *a priori* mais qui est vérifiée expérimentalement : en ce sens les probabilités sont *objectives*, elles décrivent les régularités du monde qui nous entoure. C'est seulement en lançant réellement un dé un grand nombre de fois que l'on peut s'assurer que la fréquence de l'occurrence de la face « cinq » s'approche de plus en plus de $1/6$, et en ce sens les casinos nous fournissent la plus belle preuve de ce fait car s'il y a bien un endroit où de nombreux dés sont jetés, c'est celui-ci. Un autre exemple de régularité empiriquement constatée est celle des phénomènes sociaux comme le mariage. Ici l'expérience s'identifie à un individu, et le résultat de l'expérience est positif si l'individu se marie au cours de l'année, négatif sinon. Ce fut une surprise, au XIXe siècle, de découvrir que la fréquence des mariages était d'une régularité à toute épreuve, et que plus le nombre d'individus (c'est-à-dire d'expériences) augmentait, plus la fréquence des individus mariés au cours de l'année se rapprochait d'une valeur limite. Ce fut une surprise parce que le mariage était vu comme l'expression d'une volonté, d'un libre arbitre, trouver une régularité dans un tel phénomène était inattendu. Une fois cette régularité constatée, il est possible de donner un sens à la probabilité qu'un individu se marie au cours de l'année, c'est la limite de la proportion de personnes mariées au cours de l'année dans une population de plus en plus grande.

Une objection possible est que justement il n'est pas possible de réaliser un nombre infini d'expériences, et donc que la « vraie » probabilité d'un évènement ne pourra jamais être connue. En réalité, grâce au calcul des probabilités, il est possible d'estimer l'écart entre la fréquence observée et la probabilité « réelle », et ainsi pour un nombre d'expériences grand mais fini, la fréquence donne, à une incertitude *connue* près, la probabilité [4]. Et à y réfléchir, c'est la même chose pour toute grandeur physique : une grandeur comme la longueur d'une table ne peut être définie avec une précision infinie, toute mesure de la longueur d'une table est entachée d'une incertitude dépendant de l'appareil de mesure (avec un double décimètre l'incertitude est de l'ordre du millimètre), et la « vraie » valeur de la longueur pourrait être définie comme la limite des longueurs mesurées lorsque que des appareils de plus en plus précis sont utilisés. Mais, dira-t-on, pour certains cas comme les phénomènes sociaux le nombre d'expériences ne peut dépasser le nombre d'êtres vivant sur Terre, la probabilité ne pourra donc jamais

4. C'est le concept d'intervalle de confiance : un tel intervalle, centré autour de la fréquence observée, indique où se trouve très probablement la vraie valeur de la probabilité.

être connue avec une précision aussi petite que voulue. Là encore, une comparaison avec le concept de longueur s'impose : la longueur d'une table n'est jamais définie avec une précision arbitraire, à l'échelle microscopique le bord de la table n'est pas droit mais irrégulier, la température peut fluctuer et modifier légèrement la longueur, etc., de sorte qu'une imprécision irréductible se cache dans toute mesure de longueur. Ainsi le concept de probabilité est en ce sens aussi défini que celui de longueur [5] puisque l'on connaît, à l'aide du calcul des probabilités, les conditions dans lesquelles la fréquence donne une bonne approximation de la probabilité.

Une autre objection est celle de la reproductibilité des expériences : rigoureusement parlant aucune expérience n'est semblable à une autre, et cela se voit d'autant plus lorsque les probabilités sont définies pour les phénomènes sociaux. Aucun mariage n'est identique à un autre, l'étude statistique des faits sociaux passe par une mise en équivalence de phénomènes qui ne le sont pas. Cette critique n'est pas injustifiée mais peut être affinée. La probabilité d'un évènement dépend de l'expérience aléatoire dont il est vu comme une issue, c'est une propriété de l'ensemble des expériences et pas d'un cas particulier : dans le cas des jeux de dé, la probabilité de tomber sur le chiffre « cinq » n'est pas une propriété du prochain lancer de dé, mais bien d'un très grand ensemble de lancers ; dans le cas du mariage, la probabilité de se marier au cours de l'année ne concerne pas un individu en particulier mais tout un groupe social, et peut donc varier pour un même individu en fonction du groupe auquel il est rattaché. Ainsi, la probabilité de se marier d'un individu de 25 ans encore étudiant n'est pas la même s'il est considéré comme appartenant au groupe des personnes de 25 ans ou celui des étudiants [6]. À trop préciser les caractéristiques du groupe social dont fait partie l'individu (25 ans, étudiant, habitant à Paris, diabétique, sportif, amateur de brocoli, etc.), le groupe social finit par n'être restreint qu'à un seul individu et la probabilité n'a plus de sens puisque le nombre d'expériences est trop petit. Le choix de l'expérience aléatoire, c'est-à-dire d'une classe d'équivalence de différents phénomènes, est dicté par deux exigences contraires entre lesquelles un compromis doit être trouvé : il faut que la classe d'équivalence comporte assez d'éléments pour que la fréquence approche suffisamment bien la probabilité, mais en même temps les différents éléments de la classe d'équivalence doivent être suffisamment semblables pour que la notion de « reproductibilité » puisse avoir un sens.

Un problème subsiste : il faut pouvoir distinguer, parmi les suites de résultats d'expériences répétées dont les fréquences de réalisation d'évènements convergent, celles qui possèdent des régularités supplémentaires de celles correspondant à de véritables expériences aléatoires. Que l'on pense aux jours de la semaine : la fréquence d'apparition du « jeudi » est $1/7$, pourtant dire que la probabilité qu'un jour quelconque soit un jeudi est $1/7$ semble être vide de sens : la suite des jours de la semaine n'est pas une suite aléatoire, elle exhibe au contraire une régularité à toute épreuve. Pour pallier ce défaut, il est nécessaire d'avoir un critère pour définir ce qu'est une suite aléatoire. Si l'on considère une expérience répétée un grand nombre de fois et que l'on s'intéresse à la venue d'un évènement E dont il est supposé qu'il arrive avec une probabilité p, on

5. Ce qui ne signifie pas pour autant que ces concepts sont bien définis : penser qu'il y a une valeur « vraie », réelle, objective de la longueur d'une table ou de la probabilité d'obtenir « cinq » avec un dé et que nos mesures tentent de la révéler ou du moins de s'en approcher n'est pas une conception acceptée par tous, mais il est argumenté ici que le problème n'est pas de nature différente pour la probabilité ou la longueur.

6. Même si cela demande à être vérifié, elle est sûrement plus élevée dans le premier cas que dans le second.

dit que la suite des résultats des expériences est *aléatoire* s'il n'existe pas de moyen de gagner de l'argent en pariant sur la survenue de E. C'est-à-dire que si avant chaque expérience, un joueur peut parier la somme S de son choix sur la survenue de E (auquel cas il gagne S/p si E survient effectivement) ou sur la non survenue de E (auquel cas il gagne $S/(1-p)$ si E ne survient pas), et que ce joueur n'a accès qu'aux résultats des expériences précédentes, alors il n'a pas d'algorithme, de procédure qui lui permette (*via* le choix des sommes misées S) de s'assurer un gain d'argent : sur le long terme il ne gagnera pas d'argent ni n'en perdra. Pour reprendre l'exemple des jours de la semaine, un joueur comprendrait vite qu'il a intérêt à parier sur « jeudi » sept jours après que le précédent « jeudi » soit arrivé et il gagnerait alors de l'argent à coup sûr. Au contraire, pour les jeux de hasard il n'existe pas de procédure miraculeuse qui permette de gagner de l'argent à coup sûr, et là encore ce jugement ne découle pas de considérations *a priori* mais bien d'une observation empirique, observation notamment faite dans les casinos [7]. Ici encore, cette hypothèse d'aléatoire d'une série d'expériences est une propriété *globale*, qui dépend de l'ensemble des expériences et pas d'un cas particulier. Cela ne fait donc aucun sens de dire qu'un évènement singulier est aléatoire, il l'est toujours relativement à une série d'expériences dont il n'est qu'une des instances.

La conception fréquentiste est-elle satisfaisante ? Pas entièrement. Tout d'abord parce que le calcul des probabilités tel qu'il est aujourd'hui, bien qu'inspiré fortement par cette conception, présuppose en réalité des hypothèses plus fortes [8]. Mais aussi parce que si les expériences ne sont pas suffisamment reproductibles et présentes en assez grand nombre, il faut selon la conception fréquentiste tout simplement renoncer à utiliser le calcul des probabilités. Or le concept de probabilité est parfois utilisé dans de tels cas, notamment comme on le verra par la suite lorsqu'il concerne la loi unique du hasard.

1.3 La conception subjectiviste

C'est peut-être la plus inhabituelle des trois conceptions, celle présentée ci-dessous prend la forme qui lui est donnée dans l'œuvre du philosophe et mathématicien italien Bruno de Finetti (1906-1985), qui arrive de manière élégante à appuyer ses prétentions philosophiques sur des résultats mathématiques. Pour lui, la probabilité d'un évènement est une *mesure du degré de croyance* quant à l'arrivée de cet évènement. La probabilité d'un évènement est le résultat du jugement d'un sujet humain, elle peut donc varier, pour un même évènement, d'un sujet à un autre : c'est ce sens qu'elle est *subjective*. Elle diffère de la probabilité épistémique car cette dernière est relative à la connaissance et est donc identique pour différents sujets rationnels connaissant les mêmes choses, alors que la probabilité subjective pourrait quand même différer pour de tels sujets. L'exemple classique, en dehors des habituels jeux de hasard, est le sport. Avant qu'un match de football ait lieu, son issue est incertaine et chacun a une certaine croyance quant à l'équipe qui va remporter la victoire, la probabilité permet justement

7. Richard von Mises compare cette observation à celle de l'impossibilité du mouvement perpétuel : de même que la physique ne prouve pas que le mouvement perpétuel n'existe pas mais au contraire le constate et élève cette impossibilité en postulat (le second principe de la thermodynamique), le calcul des probabilités ne prouve pas qu'il est impossible de gagner à coup sûr de l'argent au casino mais au contraire élève cette impossibilité au rang d'axiome, de critère permettant de distinguer les suites aléatoires des autres.

8. Pour les connaisseurs, l'additivité dénombrable (la σ-additivité) de probabilité peut être mise en défaut par la conception fréquentiste alors que c'est un des axiomes de Kolmogorov.

de quantifier cette croyance. En fonction des informations à disposition, de la sensibilité et de l'expérience de chacun, les croyances et donc les probabilités diffèrent. Un autre exemple est celui de la prospection pétrolière : un géologue est mandaté par une compagnie pétrolière pour indiquer si, dans un zone donnée, se trouve du pétrole ou non. Le géologue enquête, rassemble des informations et acquiert un certain degré de croyance, appuyé sur les informations récoltées et son propre savoir, quant à la présence de pétrole ou non dans la zone en question. Ce degré de croyance doit être intégré dans un processus de décision complexe de la compagnie : la décision de forer à cet endroit dépend de la présence potentielle ou non de pétrole, mais aussi du coût du forage, peut-être des forages potentiels dans d'autres zones, des objectifs économiques de l'entreprise, etc. Or le géologue n'est pas au courant de tous ces détails, c'est pour cela qu'il exprime son degré de croyance à l'aide de la probabilité : cette dernière est un moyen opérationnel d'exprimer dans un langage compréhensible par tous son niveau de certitude. À l'aide de ce langage, la compagnie peut intégrer l'avis de l'expert géologue dans son processus de décision, sans qu'elle n'ait à acquérir les compétences du géologue et sans que celui-ci n'ait à connaître les autres enjeux de la décision.

Il est nécessaire d'avoir un moyen d'opérationnaliser le concept de probabilité, c'est-à-dire d'attribuer au degré de croyance une valeur chiffrée. L'outil clé est le *pari*. Un sujet attribue une probabilité p à un évènement E s'il est prêt à miser une somme pS en échange du gain de la somme S si E arrive effectivement. Autrement dit, on mesure la probabilité attribuée par un sujet en regardant à quelle cote il est prêt à parier[9]. Plusieurs remarques s'imposent pour dissiper les malentendus. Tout d'abord le pari est censé être accepté pour n'importe quelle somme S, du moment qu'elle n'est pas trop grande (car pour des sommes trop grandes les gens sont susceptibles d'aversion au risque). En particulier la somme peut être négative, ce qui revient à parier sur l'évènement contraire avec probabilité $1 - p$. Ainsi le sujet a intérêt à parier à la cote qui lui paraît équitable, celle pour laquelle il n'est ni perdant ni gagnant. Deuxièmement, et justement pour contrer l'aversion au risque, le sujet est forcé de prendre le pari et de proposer une probabilité p. Enfin, cette faculté d'exprimer son degré de croyance à l'aide de la méthode des paris s'acquière avec l'expérience. Bruno de Finetti avait entraîné ses étudiants en les faisant parier de manière régulière sur les résultats des matchs de football de la ligue italienne, et au fur et à mesure ils ont acquis une meilleure intuition de ce que mesure la probabilité. La probabilité n'est pas un nombre donné de manière arbitraire, elle permet de *traduire* une conviction quant à l'arrivée ou non d'un évènement, et cette traduction par la méthode des paris demande de la pratique, de l'expérience pour être correcte.

La probabilité d'un évènement n'a alors de sens qu'avant que l'évènement soit arrivé, puisqu'elle est là pour quantifier l'incertitude. Une fois l'évènement arrivé, seul reste de la certitude et la probabilité n'a plus lieu d'être. Une probabilité est donc associée à un évènement singulier, au contraire de la conception précédente où elle n'avait de sens que pour la répétition d'une expérience aléatoire et dépendait donc de l'ensemble des résultats des différentes expériences. Si la probabilité n'a pas de sens après la survenue éventuelle de l'évènement, elle peut en revanche être modifiée avant au cas où des nouvelles informations sont à disposition. Pour reprendre l'exemple du match de football, si une des deux équipes marque, alors la probabilité subjective de la victoire de l'équipe qui vient de marquer augmentera sûrement : cela ne signifie pas que la nouvelle probabilité est plus précise ou plus « juste » : la probabilité change tout simple-

9. La probabilité est alors définie comme l'inverse de la cote.

ment parce que les informations accessibles au sujet (ici le fait qu'un but a été marqué) changent.

Une des beautés de cette méthode est qu'il est possible de retrouver les règles usuelles du calcul des probabilités. À titre d'exemple, supposons qu'un sujet, appelons-la Alice, soit amené à parier sur le résultat d'un match de football opposant Bourges à Caen. Elle doit choisir parmi les trois options : (A) victoire de Bourges, (B) victoire de Caen et (C) match nul. Elle attribue à ces trois options des probabilités respectives p_A, p_B et p_C. Mais supposons qu'elle les choisisse telles que $p_A + p_B + p_C \neq 1$: c'est contraire aux règles du calcul des probabilités, puisque une et une seule de ces trois options se réalisera. Alors il existe un système de pari, c'est-à-dire un choix de trois sommes S_A, S_B et S_C tel que si Alice parie sur (A), (B) et (C) les sommes respectives S_A, S_B et S_C aux cotes dictées par p_A, p_B et p_C, alors elle perd de l'argent à coup sûr. Pour être sûr de bien comprendre ce résultat, une application numérique peut être utile, supposons donc que $p_A = 0.5,\ p_B = 0.4$ et $p_C = 0.3$, de sorte que $p_A + p_B + p_C = 1.2 > 1$. Alors, si Alice parie 5€ sur (A) (et gagne donc 10€ si (A) se réalise), 4€ sur (B) (et gagne donc 10€ si (B) se réalise) et 3€ sur (C) (et gagne donc 3€ si (C) se réalise) alors elle parie au total $5 + 4 + 3 = 12$€ et gagne 10€ quelque soit le résultat : elle perd 2€ à coup sûr. De manière plus générale, avec la seule contrainte que le comportement d'un sujet est cohérent, c'est-à-dire qu'il n'accepte pas un système de pari dans lequel il perd à coup sûr, alors il est possible d'en déduire les règles du calcul des probabilités.

La conception subjectiviste a pour elle l'avantage d'être d'une grande cohérence (sûrement plus que les deux autres) et d'une grande beauté conceptuelle. Elle est parfois accusée d'autoriser des attributions de probabilité « folles » car arbitraires, mais en réalité il faut voir la méthode du pari comme un moyen de mesurer le jugement d'un sujet sur un phénomène incertain, le degré de croyance préexiste à la mesure et cette dernière demande de l'entraînement, de l'expérience afin d'avoir une réelle signification. Une autre critique, ou du moins un autre point à éclaircir, est de savoir si les probabilités objectives peuvent être amenées à contraindre les probabilités subjectives : dans le cas où la fréquence des résultats d'une expérience aléatoire répétée un grand nombre de fois s'approche d'une valeur limite, l'évaluation de probabilité subjective du résultat de l'expérience aléatoire doit-elle s'accorder avec la limite de la fréquence ? Ce sera discuté dans la partie qui suit. Mais une critique majeure subsiste : ce concept de probabilité est foncièrement *antiréaliste* [10], ce qui pose problème lorsqu'il est utilisé dans les théories physiques. En physique statistique, grande utilisatrice du concept de probabilité, on veut que les probabilités aient un sens objectif, ou du moins intersubjectif (c'est-à-dire que l'attribution d'une probabilité soit la même pour tous les sujets), car on veut que les probabilités parlent d'une réalité qui soit dans un sens ou dans un autre indépendante de nous. Et les probabilités subjectives ne saisissent justement pas cet aspect.

2 La loi des grands nombres

Il serait malhonnête de discuter ces différentes conceptions de la probabilité sans aborder le résultat qui prétend les unifier en faisant triompher l'interprétation fréquentiste, à savoir la loi des grands nombres. Schématiquement, elle énonce que *si une*

10. Un des traités de probabilité de de Finetti commence par cette phrase volontairement provocatrice : « Probability does not exist ».

expérience aléatoire est répétée un grand nombre de fois de telle sorte que les différentes expériences soient indépendantes les unes des autres, alors la fréquence des réalisations d'un évènement est proche de la probabilité de cet évènement. Mais dans cette formulation il n'est pas dit ce à quoi renvoie la probabilité, et surtout le statut de la loi des grands nombres reste à éclaircir : s'agit-il d'un théorème prouvé mathématiquement ou d'une loi empirique ? Pour répondre à ces questions, il est nécessaire d'expliciter la signification de cette loi dans chacune des conceptions présentées ci-dessus.

La conception classique Pour simplifier la situation, on s'intéresse à une expérience possédant uniquement deux issues également probables, « succès » et « échec », de sorte que la probabilité de « succès » est égale à $1/2$. Si n expériences indépendantes sont réalisées, le nombre de cas également probables pour l'ensemble des n expériences est 2^n : par exemple si 3 expériences sont effectuées, il y a 8 issues possibles également probables [11] :

« succès, succès, succès »
« échec, succès, succès »
« succès, échec, succès »
« échec, échec, succès »
« succès, succès, échec »
« échec, succès, échec »
« succès, échec, échec »
« échec, échec, échec ».

La loi des grands nombres énonce que si n est grand, parmi les 2^n cas également possibles pour la série de n expériences, une majorité correspond à des situations où la fréquence de succès parmi les n expériences est proche de $1/2$. Autrement dit, si l'on se fixe un écart $\varepsilon > 0$ aussi petit soit-il, alors pour n assez grand la probabilité (au sens de la proportion de cas favorables parmi les 2^n cas équiprobables possibles) que la fréquence de succès parmi les n expériences soit comprise entre $1/2 - \varepsilon$ et $1/2 + \varepsilon$ est proche de 1. Par exemple, si l'on choisit $\varepsilon = 0.1$, alors en effectuant $n = 500$ expériences, la probabilité pour que la fréquence de succès parmi ces 500 expériences soit comprise entre 0.4 et 0.6 est plus grande que 0.95.

Éclaircissons ce résultat : les deux probabilités apparaissant dans cette loi des grands nombres sont des probabilités au sens classique, c'est-à-dire des rapports du nombre de cas favorables sur le nombre de cas total. La loi des grands nombres est donc un *théorème*, c'est un résultat qu'il est possible de prouver mathématiquement et sa preuve repose sur du dénombrement, elle ne fait intervenir aucun « hasard ». La messe est-elle pour autant dite ? Non, car cette loi des grands nombres ne dit pas que dans un très grand nombre d'expériences, la fréquence des réalisations d'un évènement est proche de la probabilité de l'évènement, elle dit qu'*il est très probable* que la fréquence des réalisations de l'évènement soit proche de la probabilité de l'évènement [12]. Pour boucler le

11. L'hypothèse d'indépendance se traduit précisément en disant que ces issues sont également probables.

12. Le connaisseur aura remarqué que seule la loi « faible » des grands nombres, c'est-à-dire de convergence en probabilité et pas presque sûre, est abordée. La première raison est que la convergence en probabilité est celle qui sert en pratique car elle donne des informations utiles pour un nombre fini d'expériences ; la deuxième est qu'elle est plus simple à exposer. Mais surtout la loi forte des grands nombres ne résout pas le problème de la loi unique du hasard (voir ci-dessous), car il faut toujours expliquer pourquoi un évènement presque sûr (qui n'est pas un évènement certain) se produit toujours en pratique.

raisonnement, il faudrait pouvoir dire que *si un évènement possède une probabilité très proche de* 1*, alors en pratique il se produira* (ou de manière équivalente, si un évènement possède une probabilité très proche de 0, alors en pratique il ne se produira pas), assertion à laquelle peut être donnée le nom de *loi unique du hasard* [13]. Or la définition « classique » de la probabilité ne dit pas pourquoi, si la probabilité d'un évènement est extrêmement proche de 1, il faut se comporter en pratique comme s'il arrive de façon certaine. La conception classique n'explique pas de façon logique pourquoi nous ne voyons pratiquement jamais les évènements dont la probabilité est quasi nulle et pratiquement toujours ceux dont la probabilité est très proche de 1. C'est une hypothèse qui dépasse le cadre de la conception classique, de sorte qu'à l'intérieur de cette dernière il n'est pas possible de *prouver* la convergence des fréquences.

La conception fréquentiste La conception fréquentiste donne l'impression au contraire de partir de la loi des grands nombres puisqu'elle définit les probabilités comme des limites de fréquence. En particulier, et comme cela a déjà été dit, la convergence des fréquences est une constatation *empirique*, elle n'est pas prouvable. Pourtant il existe bel et bien une loi des grands nombres qui ne correspond pas à cette hypothèse de convergence des fréquences mais qui spécifie la « vitesse » de convergence. Elle possède le même énoncé que celui de la conception classique, à savoir que quelque soit l'écart $\varepsilon > 0$ que l'on se donne, il est possible de choisir un nombre n d'expériences assez grand tel qu'avec une probabilité proche de 1, la fréquence de réalisation d'un évènement de probabilité p soit comprise entre $p - \varepsilon$ et $p + \varepsilon$, et il est même possible de quantifier à quel point n doit être grand [14]. Cette quantification est utile car c'est grâce à elle qu'il est possible d'estimer l'écart entre la probabilité « réelle » et la fréquence observée, et ainsi savoir à quel point la mesure de la probabilité effectuée à l'aide d'une observation de la fréquence est précise.

Là aussi éclaircissons ce résultat : tout d'abord il se *déduit* de deux prémisses, l'hypothèse de convergence des fréquences et l'hypothèse d'aléatoire de la suite des expériences (c'est-à-dire qu'il n'existe pas de système de pari permettant de gagner de l'argent à coup sûr à partir de la série des résultats des expériences). Dans cette loi des grands nombres, les probabilités sont à prendre en un sens fréquentiste : la probabilité de l'évènement est égale à la limite de la fréquence de ses réalisations si le nombre d'expériences est grand, et de dire que la probabilité qu'en n expériences la fréquence de réalisation d'un évènement est comprise entre $p - \varepsilon$ et $p + \varepsilon$ (où p est la probabilité de l'évènement) est proche de 1 signifie qu'en effectuant un grand nombre de fois des séries de n expériences, dans la majeure partie de ces séries la fréquence observée

13. Ce nom est du à Émile Borel. Ce dernier donne même une estimation de la probabilité à partir de laquelle un évènement ne se produit en pratique jamais : pour des évènements qui concernent un unique individu, elle est fixée à 10^{-6}. C'est-à-dire que si un évènement possède une probabilité inférieure à un millionième, alors il faut se comporter en pratique comme s'il n'arrive pas. Ce chiffre n'est pas arbitraire, il découle de la constatation que si un individu devait prendre en compte l'ensemble des évènements dont la probabilité est inférieure à 10^{-6}, il serait retrouverait bien vite paralysé et n'agirait plus. Si par contre un évènement peut toucher un individu parmi l'ensemble de l'humanité, sa probabilité doit être inférieure à 10^{-15} pour que l'on agisse comme s'il ne se produit pas. Par exemple, à l'échelle individuelle, un sujet doit se comporter comme s'il allait perdre au loto, mais à l'échelle d'un pays, il n'est plus possible de se comporter comme si tout le monde allait perdre au loto.

14. Dans sa forme mathématique, elle énonce que si n expériences sont réalisées alors avec une probabilité supérieure à $1 - \dfrac{1}{4n\varepsilon^2}$ l'écart entre la fréquence d'apparition d'un évènement et sa probabilité ne dépasse pas ε.

de l'occurrence de l'évènement est comprise entre $p - \varepsilon$ et $p + \varepsilon$. De même que dans le cas classique, la loi des grands nombres fréquentiste dit qu'*il est très probable* que la fréquence de réalisation de l'évènement soit proche de sa probabilité, mais l'expression « il est très probable » prend un autre sens. Dans la conception fréquentiste, un évènement est très probable si, lorsque l'expérience aléatoire est répétée un grand nombre de fois, il se réalise dans la majeure partie des cas. Pour reprendre l'exemple numérique donné plus haut (celui dans lequel l'évènement « succès » a une probabilité de $1/2$), si l'on répète un grand nombre de fois des séries de 500 expériences, alors parmi la majeure partie de ces séries, la fréquence de succès sera comprise entre 0.4 et 0.6. Mais cela n'explique pas pourquoi la loi unique du hasard doit s'appliquer : dans la plupart des cas, nous sommes confrontés à une unique occurrence d'un évènement extrêmement probable, pas plusieurs, or la probabilité d'un évènement isolé n'a de pas de sens dans la conception fréquentiste. En pratique nous ne faisons jamais plusieurs séries de 500 expériences afin de vérifier que la fréquence de succès varie peu d'une série à l'autre. De même que la conception classique mais pour des raisons différentes, la conception fréquentiste n'arrive pas à justifier de façon interne la loi unique du hasard : comme la probabilité est toujours définie relativement à une série d'expériences, il n'est pas possible de dire pourquoi elle doit dicter le comportement en face d'un évènement singulier même si la probabilité de cet évènement est très proche de 1. Or la loi unique du hasard recommande précisément, en face d'un évènement singulier dont la probabilité est proche de 1, de se comporter en pratique comme si celui-ci allait arriver. Si la conception fréquentiste peut « prouver » la loi des grands nombres, elle ne permet pas de justifier l'usage pragmatique qui en est fait.

La conception subjective La conception subjectiviste étant précisément ancrée dans la prise de décision, la loi unique du hasard est naturelle : presque par définition, si j'attribue à un évènement une probabilité proche de 1 alors je me comporte en pratique comme s'il était certain. La nécessité qui me pousse à agir ainsi n'est pas dans le monde, elle est en moi, dans le sens où la probabilité que j'attribue correspond à un jugement, pas à un état du monde extérieur. C'est pourquoi le problème de la justification de la loi unique du hasard disparaît dans une conception subjectiviste, alors que dans les autres conceptions il tenait à ce qu'un état du monde extérieur devait dicter la conduite d'un individu rationnel.

Mais un autre problème se pose alors pour la conception subjectiviste : en quoi les régularités du monde doivent influer les probabilités que j'attribue ? Si par exemple j'observe 100 lancers de pile ou face et que « pile » sort 52 fois, pourquoi est-ce que je devrais attribuer à la probabilité de tomber sur pile une valeur proche de $1/2$? Il serait possible d'essayer de penser en même temps l'existence de probabilités objectives et subjectives, les secondes étant égales aux premières lorsque les premières sont disponibles, ainsi les évènements répétables à l'infini seraient décrits par les limites des fréquences de leur réalisation et les probabilités subjectives pourraient s'appliquer aux évènements singuliers. Mais ce n'est pas ce qu'accepte de Finetti, il propose une autre vision s'appuyant sur un élégant théorème mathématique. Supposons qu'une série d'expériences aléatoires ayant deux issues (« succès » et « échec ») prenne part devant moi, par exemple un jeu de hasard. Supposons que je fasse l'hypothèse que le nombre de succès parmi n lancers ne dépende pas de l'ordre des lancers, c'est-à-dire par exemple que j'attribue la même probabilité à ce qu'il y a 1 « succès » suivi de 5 « échecs » que 3 « échecs » suivis d'un « succès » puis de 2 « échecs ». Cette supposition est un jugement subjectif que je suis enclin à faire si les expériences semblent être reproduites à l'identique et de façon indépendante. Alors simplement avec cette hypothèse d'indifférence de l'ordre et les règles du calcul des probabilités (qui, rappelons-le, découlent d'une exigence de cohérence), je dois modifier mes estimations de probabilité pour les résultats suivants du jeu, formellement de la même manière que si la fréquence de succès allait converger lorsque le nombre d'expérience augmente de plus en plus [15]. La « loi des grands nombres » montre comment, en rajoutant ce jugement subjectif d'indifférence de l'ordre des succès, le raisonnement probabiliste *contraint* à attribuer à la probabilité une valeur de plus en plus proche de la fréquence observée au fur et à mesure que le nombre d'expériences augmente.

Que retenir de cette discussion ? Tout d'abord qu'au delà d'une idée vague de convergence des fréquences vers la probabilité, les formulations de la loi des grands nombres se révèlent plus subtiles. En particulier, elles permettent d'estimer le nombre nécessaire d'expériences à réaliser pour qu'il soit très probable que la fréquence observée de réalisation d'un évènement soit proche de sa probabilité, information extrêmement utile pour savoir avec quelle précision les fréquences s'approchent des probabilités. Mais à un moment ou à un autre l'argument devient circulaire, car il faut expliquer pourquoi en pratique, il faut se comporter comme si un évènement très probable arrive à coup sûr, chose que ni les conceptions classiques et subjectives n'arrivent à expliquer de par leur structure interne. La conception subjectiviste est en un sens construite pour répondre à cette question, et elle peut même se payer le luxe d'expliquer pourquoi et dans quelles conditions il est rationnel de se comporter comme si les fréquences allaient converger au fur et à mesure de l'augmentation du nombre de réalisations d'expériences semblables. Mais le problème qui demeure est qu'en faisant ceci, elle rejette tout idée de régularité statistique des phénomènes qui pourrait être *objectivement* décrite puisque les probabilités n'existent que dans notre esprit.

Conclusion

Comment concilier ces différentes visions ? Plutôt que de tenter de compter les points afin de dégager un « vainqueur » qui serait la bonne manière d'attribuer des probabilités, il est sans doute plus utile de voir l'étude de ces points de vue inconciliables comme une exploration des possibles [16]. Chacune de ces conceptions explore un point de vue sur les probabilités jusqu'au bout, montre les questions auxquelles ce point de vue peut

15. Pour les connaisseurs, la formulation est imprécise. Le résultat de de Finetti (pour une présentation moderne on peut consulter l'article de J. F. C. Kingman, « Uses of exchangeability » dans *The annals of Probability*, 1978, Vol. 6, No. 2, 183-197) peut s'énoncer de la manière suivante : si $X_1, X_2, \dots$ sont des variables aléatoires prenant les valeurs 0 (« échec ») ou 1 (« succès ») telles pour toute permutation σ de $\{1, 2, \dots, n\}$ la loi de $(X_1, X_2, \dots, X_n)$ est la même que celle de $(X_{\sigma(1)}, X_{\sigma(2)}, \dots, X_{\sigma(n)})$ (indifférence de l'ordre), alors la loi du processus $(X_n)_{n\geqslant 1}$ est une combinaison convexe de lois de processus de Bernoulli i.i.d. C'est-à-dire qu'il existe une mesure μ de probabilité sur $[0, 1]$, telle que la probabilité qu'il y ait r succès et $n-r$ échecs parmi $(X_1, X_2, \dots, X_n)$ vaille $\displaystyle\int_0^1 \binom{n}{r} p^r (1-p)^{n-r} \mathrm{d}\mu(p)$. Ainsi l'hypothèse d'indifférence de l'ordre n'implique pas que le processus soit indépendant, mais que ce soit un « mélange » de processus indépendants. La fréquence empirique $\displaystyle\frac{1}{n}\sum_{k=1}^{n} X_k$ converge alors vers une variable aléatoire suivant la loi μ.

16. Cette idée a été entendue dans une conférence de Soazig le Bihan.

répondre, les principaux avantages mais aussi les problèmes qui se posent, les aspects qui sont insaisissables. Par exemple, si les probabilités objectives sont des attributs bien définis des phénomènes, au même titre que les longueurs, elles pèchent quand il s'agit de capturer les phénomènes singuliers et ne peuvent notamment saisir directement la loi unique du hasard qui sert à régler nos décisions. Seule une analyse rigoureuse des concepts mis en jeu, que ce document espère avoir esquissé, permet d'arriver à de telles conclusions et en particulier de différencier la véritable insuffisance de celle qui ne l'est qu'en apparence.

Il ne s'agit pas non plus d'une analyse statique, d'un campement dans des positions fixes : tout l'intérêt d'avoir plusieurs interprétations est de pouvoir naviguer de l'une à l'autre. En effet, il ne faut pas oublier que les probabilités peuvent servir un dispositif argumentatif et s'inscrire à l'intérieur d'un réseau rhétorique. C'est d'ailleurs leur caractère herméneutique qui leur a permis d'être reprises par des acteurs aussi variés que les adeptes du casino, les physiciens, les sociologues, les parieurs, etc. Pour exemplifier leur variété d'utilisation, il est possible de se tourner vers les sciences climatiques. Les climatologues doivent gérer trois sortes d'incertitude : celle liée à la variabilité naturelle du climat, dont on peut dire qu'elle se décrit objectivement par l'analyse des fréquences des phénomènes ; celle liée à l'imperfection de notre connaissance du système climatique ; et celle liée au comportement humain, puisqu'il est difficile de prévoir l'état du futur de notre civilisation et donc de nos émissions de gaz à effet de serre. La magie des probabilités permet, dans les rapports à l'intention du grand public ou des décideurs, de combiner ces incertitudes fondamentalement différentes en un unique chiffre qui se retient et qui peut s'utiliser hors de la sphère scientifique.

Une meilleure connaissance des interprétations des probabilités permet de clarifier leur rôle dans une argumentation, notamment en repérant les glissements de sens : combien de fois une probabilité objective, basée sur des observations statistiques, est identifiée à la propension individuelle d'un individu ? Combien de fois les jeux de hasard sont utilisés pour donner une signification à une probabilité alors que le contexte ne s'y prête pas ? Un outil aussi puissant et polysémique que le calcul des probabilités mérite plus que d'être utilisé sans interrogation sur la signification des réponses qu'il apporte.

Bibliographie

- Émile Borel, *Valeur pratique et philosophie des probabilités*, 1939, éditions Jacques Gabay, Paris.
- Bruno de Finetti, « La prévision : ses lois logiques, ses sources subjectives », dans *Annales de l'I.H.P.*, tome 7, no. 1, 1937, p. 1-68.
- Bruno de Finetti, *Probability, induction and statistics – The art of guessing*, 1972, John Wiley & Sons.
- Alan Hájek, « Interpretations of Probability », *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.).
- Richard von Mises, *Probability, statistics and truth*, 1957 (2e édition anglaise), Dover Publications, New York.
- Jan von Plato, *Creating modern probability*, 1994, Cambridge University press, Cambridge.