

Ce que mesurent les probabilités

Mikaël Cozic¹ et Bernard Walliser^{2,3}

Prisme N°24

Septembre 2012

¹ Mikaël Cozic est maître de conférences en philosophie des sciences et de la connaissance à l'Université de Paris-Est Créteil et responsable de l'équipe « Décision, rationalité et interaction » de l'Institut d'histoire et de philosophie des sciences et des techniques (Paris I-ENS Ulm-CNRS). Ses recherches portent sur la philosophie de l'économie et sur les théories formelles de la rationalité.

² Bernard Walliser est professeur à l'École nationale des ponts et chaussées, directeur d'études à l'École des hautes études en sciences sociales et chercheur à l'École d'économie de Paris. Auteur de nombreux ouvrages, il concentre ses recherches sur les liens entre la théorie économique et les sciences cognitives, ainsi que sur l'épistémologie de l'économie.

³ Nous remercions Anouk Barberousse, Philippe Mongin, Denis et Hervé Zwirn pour les discussions que nous avons eues avec eux à propos de cet article ainsi que Josselin Garnier et Laurent Mazziak pour leurs relectures. Nous sommes particulièrement redevables envers Isabelle Drouet pour des critiques et des suggestions essentielles.

Résumé

Les probabilités sont l'un des outils fondamentaux sur lesquels un modélisateur peut s'appuyer pour décrire et expliquer les phénomènes qu'il étudie. Il s'en sert pour représenter à la fois les propriétés d'événements de tous ordres (sociaux, psychologiques ou naturels) et les degrés de croyances des acteurs. La compréhension des probabilités soulève de redoutables difficultés conceptuelles, qui sont l'objet propre de la philosophie des probabilités. La définition des probabilités repose sur une ontologie souvent implicite et leur évaluation soulève des problèmes épistémologiques spécifiques. L'objectif de cet article est d'esquisser un cadre conceptuel dans lequel les catégories fondamentales des philosophes des probabilités comme des probabilistes puissent communiquer.

Sommaire

Introduction	7
1. Caractéristiques d'une probabilité.....	9
2. Cadre d'analyse de la situation	11
3. Typologies des probabilités	13
3.1. Typologie de base	13
3.2. Probabilités ontologique et épistémique	16
3.3. Probabilités objective et subjective.....	17
4. Comparaison des typologies	18
4.1. Croisement des typologies.....	18
4.2. Attitudes face aux interprétations multiples	21
5. Mesure des probabilités	22
6. Interprétations classiques.....	25
7. Combinaison de probabilités.....	28
8. Révision des probabilités	30
9. Problèmes ontologiques et méthodologiques	34
10. Extensions des probabilités	38
11. Application à l'inférence statistique	41
12. Application aux sciences empiriques.....	43
12.1. Considérations générales	43
12.2. Cas de la théorie des jeux	44
13. Application à l'épistémologie.....	46
14. Conclusion	47
Références	49

Introduction

Il existe une variété de modèles formels qui visent à représenter les croyances des acteurs sur leur environnement. En se restreignant à ceux de la logique épistémique, les croyances sont exprimées dans deux cadres distincts. Dans un cadre syntaxique, elles sont définies sur des propositions⁴ par le biais d'opérateurs de croyance individuels qui expriment qu'un acteur croit que telle ou telle proposition est vraie. On note par exemple $B_i p$ le fait que l'individu i croit que la proposition p est vraie. Dans un cadre sémantique, elles sont définies sur des mondes possibles par le biais de relations d'accessibilité individuelles qui indiquent les mondes qu'un acteur considère comme envisageables à partir d'un monde donné. Dans les deux cas, les croyances des acteurs portent sur leur environnement commun, mais aussi sur les croyances d'autrui sur ces croyances primaires. Sous des conditions peu contraignantes, les deux cadres sont équivalents moyennant la définition de principes de correspondance.

Les croyances des acteurs sont également distinguées selon le degré de finesse avec lequel elles appréhendent leur objet. Dans une approche ensembliste, la croyance est affaire de tout ou rien, au sens où un acteur croit ou non que telle ou telle proposition est vraie. Dans une approche probabiliste, la croyance est graduée au sens où l'acteur exprime un *degré de croyance* sur telle ou telle proposition. Les croyances probabilistes peuvent être mises en relation avec les croyances ensemblistes en syntaxe ou en sémantique. En syntaxe, on peut considérer que l'opérateur de croyance s'applique à une proposition p si le degré de croyance de l'individu en p excède un certain seuil (thèse de Locke). En sémantique, on peut considérer que la relation d'accessibilité représente le support d'une distribution de probabilités sur les mondes possibles à partir d'un monde donné. En s'appuyant sur cette correspondance, les propriétés définies dans un cadre ensembliste sont transférées (d'une manière qui n'est pas toujours immédiate) à une approche probabiliste.

⁴ Par souci de simplicité, aucune distinction n'est faite entre énoncés ou formules et propositions, comme il est d'usage en philosophie.

Par la suite, on s'en tient à un acteur unique qui se forge des croyances probabilistes sur son environnement physique ou social. C'est dire que l'on n'envisage pas de croyances croisées, en l'occurrence des croyances d'un acteur sur celles d'un autre. Les croyances directes sont tout naturellement exprimées dans le cadre sémantique des mondes possibles, chaque monde traduisant un état possible de l'environnement. Le monde réel étant fixé (par le modélisateur), l'acteur se contente de définir une distribution de probabilités sur l'ensemble des mondes possibles. En fait, il définit plus généralement une probabilité sur des événements définis comme des sous-ensembles de mondes. Les probabilités sont « extrêmes » si elles prennent la valeur 0 ou 1, et « non extrêmes » dans le cas contraire.

Le présent article précise le cadre conceptuel dans lequel s'insère la définition des probabilités. Il situe ensuite dans ce cadre les différentes écoles qui ont proposé une interprétation particulière des probabilités. Il discute enfin les opérations communément réalisées sur les probabilités en fonction de leurs interprétations. Ce travail tente ainsi de jeter un pont entre l'analyse technique des probabilités et leur analyse philosophique. Pour les probabilistes, il cherche à mettre en évidence les nombreux présupposés implicites qui conditionnent leur travail. Pour les philosophes, il cherche à rendre saillantes les conditions formelles qui sont supposées satisfaites en amont du calcul.

La première section introduit les caractéristiques générales d'une probabilité et la seconde section les caractéristiques de la situation aléatoire analysée. La troisième section présente les principales typologies des probabilités et la quatrième les croise en assortissant les catégories obtenues d'exemples. La cinquième section détaille les protocoles possibles de mesure empirique des probabilités. La sixième section situe les quatre principales écoles qui ont conceptualisé les probabilités au cours de l'histoire. La septième section s'intéresse à la combinaison des probabilités et la huitième à leur révision. La neuvième section soulève des problèmes méthodologiques liés à la comparaison des catégories de probabilités. La dixième section présente quelques généralisations de la notion de probabilité. Les onzième, douzième et treizième sections examinent l'application des probabilités aux statistiques, aux sciences empiriques et à l'épistémologie.

1. Caractéristiques d'une probabilité

Toute probabilité est d'abord caractérisée par son *attributeur*, en l'occurrence l'individu qui exprime la probabilité. L'attributeur endosse traditionnellement deux rôles différents en fonction de ses motivations. Il peut s'agir du *modélisateur* lui-même qui cherche à construire un modèle probabiliste d'un système qu'il observe. Par exemple, le modélisateur va attribuer une probabilité donnée à l'émission de particules par un corps radioactif. Il peut s'agir d'un *acteur* ordinaire qui cherche à représenter le contexte dans lequel il se meut à des fins de connaissance ou d'action. Par exemple, un parieur va attribuer une probabilité au fait que tel ou tel cheval va gagner une course. Cependant, même si la probabilité est exprimée par un acteur, c'est en dernier ressort le modélisateur qui considère son évaluation par l'acteur.

Toute probabilité est ensuite caractérisée par son *objet*: un système qui manifeste un certain phénomène, isolable de son contexte. Le phénomène se réalise selon diverses modalités ou instances. En pratique, il peut s'actualiser sous deux formes, statique ou dynamique. Il s'agit d'un phénomène *sériel* s'il peut se produire de façon répétée pour un même système. Par exemple, le modélisateur va attribuer une probabilité donnée au résultat du lancer d'un dé bien défini. Il s'agit d'un phénomène *populationnel* s'il s'incarne dans des individus distincts d'une même population. Par exemple, le modélisateur va attribuer une probabilité donnée au fait qu'un individu d'une certaine population soit atteint de telle ou telle maladie.

Toute probabilité est enfin caractérisée par la date à laquelle elle est censée s'appliquer au phénomène. De fait, cette date est définie en prenant comme référence la date d'occurrence du phénomène. Une probabilité *ex ante* est définie avant l'occurrence du phénomène. Un électeur affiche ainsi une probabilité qu'un candidat à une élection soit élu avant l'élection elle-même. Une probabilité *ex post* est définie après l'occurrence du phénomène. Ainsi, un juré énonce la probabilité qu'un accusé soit coupable après que le crime a été commis. De plus, la date à laquelle le phénomène est évalué peut différer de la date à laquelle cette évaluation est énoncée. Un chercheur peut penser aujourd'hui que la probabilité de découverte d'un vaccin contre le sida a été de tant hier et sera de tant demain. Ces deux dates

coïncident usuellement, la probabilité énoncée par l'attributeur s'appliquant instantanément.

En revenant à l'objet de la probabilité, il est toujours possible de considérer que les réalisations successives du phénomène sériel forment une population, et donc de se rabattre sur le cas populationnel. En sens inverse, les individus d'une population sont choisis séquentiellement (dans un ordre aléatoire à préciser) et définissent ainsi des réalisations successives du phénomène. Un phénomène peut être simultanément sériel et populationnel : il en est ainsi d'une course de chevaux dans laquelle les mêmes chevaux effectuent plusieurs courses successives. Si l'attributeur de la probabilité est le parieur, l'objet de la probabilité est le résultat de l'une des courses, en l'occurrence la place d'un cheval donné.

Quel que soit le phénomène considéré, l'attributeur peut lui affecter une probabilité selon deux points de vue différents. Ce point de vue dépend de la façon de considérer les expériences (ou épreuves) séquentielles ou parallèles de réalisation du phénomène. Le phénomène est *spécifique* si l'émetteur s'intéresse à une instance unique (*token* en anglais) de sa réalisation. Par exemple, le modélisateur s'intéresse au prochain tirage du dé et le médecin à la maladie du patient qu'il a en face de lui. Le phénomène est *générique* si l'émetteur s'intéresse à une classe (*type* en anglais) de réalisations. Par exemple, le modélisateur s'intéresse à un lancer non défini du dé ou le médecin à la maladie d'un patient quelconque.

Par ailleurs, le phénomène peut être défini par rapport à un système plus ou moins finement différencié. Les modalités du phénomène sont supposées être les mêmes dans chaque classe qui est distinguée. Le phénomène est *univoque* si le système incriminé est considéré comme identique dans toutes les instances de sa réalisation. C'est ainsi que l'on peut considérer les lancers successifs d'un dé bien défini ou les maladies d'un patient appartenant à un groupe homogène. Le phénomène est *multivoque* si l'on possède une typologie précise des systèmes incriminés qui manifestent le même phénomène, une probabilité étant définie pour chacune des catégories considérées. C'est ainsi que l'on considère la probabilité relative à des dés de différentes formes ou couleurs ou à des populations de patients de différents sexes et âges.

2. Cadre d'analyse de la situation

Telle qu'elle vient d'être introduite, la description par le modélisateur de l'objet de la probabilité met en jeu trois types de critères permettant de décrire le phénomène incriminé. Les *critères phénoménaux* sont ceux qui permettent de différencier les modalités possibles du phénomène et qui sont identiques dans toute expérience. Les *critères factuels* sont ceux qui permettent de caractériser les conditions de chaque expérience particulière de réalisation du phénomène (s'il est générique) et qui peuvent varier de l'une à l'autre. Les *critères structurels* sont ceux qui permettent de définir et de classer les systèmes qui sont à la base du phénomène (s'il est multivoque) et qui sont invariants à travers les expériences.

En pratique, le modélisateur se laisse guider par les relations causales qu'il sait exister entre ces critères. Il fait l'hypothèse que les critères tant factuels que structurels font partie des facteurs explicatifs des modalités du phénomène. En sens inverse, les critères phénoménaux rendent compte des effets engendrés par ces facteurs. Cependant, les probabilités affichées sont fonction des seuls critères structurels qui exercent une influence déterministe. Elles ne sont pas modulées par les critères factuels, dont l'influence - déterministe ou non - est précisément intégrée dans l'évaluation de la probabilité. Ainsi, le modélisateur ne cherche pas à construire un modèle décrivant complètement l'influence entre les critères, mais se contente de mesurer la dépendance des probabilités (et donc des critères phénoménaux) aux critères structurels.

Dans le cas d'une course hippique, le critère phénoménal se limite à la place occupée par un certain cheval dans une course. Les critères factuels sont l'état météorologique, la lourdeur de la piste ou l'habileté du jockey, mais aussi la nature de la selle ou l'heure de la course. Les critères structurels sont le sexe du cheval, l'âge du cheval ou son écurie, mais aussi l'hérédité du cheval ou l'identité de son propriétaire. L'attributeur va alors définir par exemple une probabilité pour un cheval de gagner une course en fonction de son âge et de son sexe. S'il dispose d'une information supplémentaire sur l'écurie du cheval étudié, il est amené à affiner sa catégorie d'appartenance et à ajuster sa probabilité en conséquence. En revanche,

l'état de la piste varie d'une course à l'autre et fait partie des causes intégrées dans la probabilité.

Toute probabilité se trouve ainsi définie dans un *cadre de référence* qui fixe précisément les caractéristiques du système et les modalités du phénomène. Des conventions d'équivalence affirment que c'est bien le même phénomène qui se reproduit dans ses instances parallèles ou séquentielles (caractéristiques structurelles fixées, caractéristiques factuelles variables). Une première condition impose ainsi que les situations comparées soient semblables. Cela doit être réalisé dans chacune des classes qui résultent de la discrimination éventuelle du système. Une seconde condition impose de plus que les expériences considérées soient indépendantes entre elles. Dans un cadre temporel, elle implique que les expériences ne soient répétées à des termes ni trop rapprochés (pour éviter des corrélations), ni trop lointains (pour éviter des évolutions trop fortes). Dans un cadre spatial, elle implique que les individus de la population interagissent de façon stable quant au phénomène considéré.

Dans le cas du dé, chaque dé est censé ne pas se modifier au cours des lancers successifs et ces lancers doivent être effectués selon le même protocole. Le dé est aussi supposé ne pas s'user et ne pas avoir de mémoire concernant ses réalisations passées. Dans le cas de la maladie, les patients sont considérés comme interchangeable et soumis aux mêmes conditions environnementales. Ils doivent aussi entretenir des influences réciproques fixées à court terme (contagions). Dans le cas d'une course, les chevaux sont supposés ne pas vieillir et les conditions de course (météorologie, état du terrain) ne pas varier ou ne pas avoir d'importance. En particulier, les résultats passés des chevaux ne doivent pas influencer ceux des courses futures, ce qui n'est souvent pas le cas (handicaps).

Bien entendu, le cadre de référence est une construction délibérée du modélisateur. Ainsi, bien que l'attribution d'une probabilité soit une opération épistémique, sa définition repose sur des présupposés ontologiques. Les catégories qui caractérisent le phénomène comme les similitudes entre les phénomènes sont définies par le modélisateur. Lorsque l'attributeur de la probabilité devient un acteur ordinaire, on suppose de plus qu'il adopte les mêmes typologies que le modélisateur.

Par ailleurs, le cadre de référence est préexistant à l'évaluation de la probabilité et ne saurait être affecté en première instance par les résultats des épreuves. Ainsi, si le modélisateur fixe *a priori* les critères structurels qui discriminent les systèmes, ils ne sauraient être remis en cause par l'observation des résultats des expériences (problème de la résilience). Si une pièce de monnaie est jugée *a priori* équitable (voir §5), un déséquilibre observé entre tirages « pile » et « face » ne saurait faire conclure que la pièce n'est pas équilibrée.

3. Typologies des probabilités

3.1. Typologie de base

Dans le cadre explicatif précédent, on peut poursuivre l'analyse des critères factuels (appelés désormais facteurs) du point de vue de l'attributeur. Un facteur est *explicite* si l'attributeur est conscient qu'il exerce une certaine influence sur le phénomène. Un facteur est *implicite* si l'attributeur ignore l'action qu'il exerce d'une façon ou d'une autre sur le phénomène. Dans certains cas, aucun facteur explicite permettant de distinguer entre deux expériences n'est repéré par l'attributeur et le phénomène est dit *primitif*. Un exemple-type est celui d'un corps radioactif qui émet aléatoirement des rayonnements de divers types. Un tel phénomène ne peut être rapporté à des causes antécédentes repérées et doit être étudié de façon autonome. Dans tous les autres cas, un facteur explicite au moins est repéré et le phénomène est dit *émergent*. Par itération, un tel phénomène peut être rapporté à des facteurs préalables qui eux-mêmes peuvent être engendrés par des facteurs encore antérieurs.

Par ailleurs, un facteur explicite est *connu* si l'attributeur peut observer la valeur qu'il prend dans chaque instance du phénomène. Un facteur explicite est *inconnu* si l'attributeur est incapable de l'observer avec précision, pour des raisons aussi bien physiques que morales. Cependant, la connaissance d'un phénomène varie selon qu'il est considéré comme spécifique ou générique. Dans un lancer de dé spécifique, l'impulsion initiale est déterministe, mais impossible à mesurer en pratique. Dans un lancer de dé générique, les impulsions varient et le phénomène prend précisément une allure aléatoire pour un observateur extérieur. Le lancer du

dé générique correspond à un phénomène aléatoire délibérément construit par l'attributeur. Il transforme en quelque sorte des facteurs explicites en facteurs implicites.

Enfin, un facteur implicite est *régulier* si l'attributeur peut considérer que son influence suit une loi aléatoire, quelle que soit cette loi. Un critère implicite est *irrégulier* si l'attributeur considère son influence comme erratique, c'est-à-dire non représentable par une loi de probabilité. Cependant, cette classification dépend de la finesse d'analyse du phénomène en jeu. En effet, il est très difficile de distinguer un phénomène stochastique d'un phénomène produit sur d'autres bases, en particulier artificiellement sans recours à des tirages au sort. Un exemple-type d'un phénomène irrégulier est un phénomène d'innovation scientifique ou technique auquel on attribue une probabilité malgré son côté erratique.

Par croisement des deux derniers critères, on peut dériver une première typologie des probabilités en quatre grandes catégories (dans ce tableau, par convention, un facteur absent est considéré comme régulier et connu) :

Facteurs implicites Facteurs explicites	tous réguliers	certains irréguliers
tous connus	probabilité irréductible : <i>corps radioactif</i>	probabilité approchée : <i>innovation technique</i>
certains inconnus	probabilité maîtrisée : <i>lancer de dé</i>	probabilité radicale : <i>course hippique</i>

Une *probabilité irréductible* est relative à un phénomène considéré comme primitif (ou à facteurs tous réguliers et connus). La probabilité traduit alors une propriété intrinsèque du phénomène, indépendante de tout facteur antécédent. C'est ainsi que le corps radioactif émet des particules selon des lois aléatoires classiques, non dérivées de probabilités liées à des facteurs plus profonds. La probabilité est irréductible, car on suppose que le phénomène est intrinsèquement aléatoire et ne peut se ramener à un déterminisme avec des variables cachées. La mécanique quantique est souvent considérée comme le seul phénomène non déterministe, du

moins dans les sciences de la nature. Le caractère aléatoire disparaît à un niveau plus macroscopique par un phénomène d'agrégation des lois microscopiques.

Une *probabilité maîtrisée* est relative à un phénomène dont l'explication repose sur des facteurs qui sont tous explicites, mais dont certains sont non connus. La probabilité résume alors la méconnaissance de la valeur effectivement prise par ces facteurs. Le lancer d'un dé par un opérateur humain ou mécanique, dans son interprétation spécifique, en est l'illustration la plus simple. La probabilité est maîtrisée si on admet que le processus de base est parfaitement déterministe et que l'attributeur ne peut observer les conditions initiales du lancer, ce qui l'empêche de prévoir son résultat (parmi plusieurs valeurs possibles) par les lois de la mécanique (voir §4.1). Dans ce cas, on a affaire à un « aléa au sens de Cournot », dû au croisement de deux chaînes causales indépendantes. Dans le jet du dé, le mouvement du dé vient rencontrer la surface sur laquelle il tombe.

Une *probabilité approchée* est relative à un phénomène qui fait l'objet de critères explicites connus, et dont les éventuels critères implicites s'avèrent irréguliers. La probabilité reflète alors la meilleure estimation que l'on puisse faire de l'occurrence du phénomène. Une innovation technologique semble obéir à ce schéma si on fait abstraction des conditions socio-économiques de son occurrence. La probabilité est approchée si on admet qu'elle résume les facteurs implicites irréguliers en leur conférant une certaine probabilité. Là encore, en sciences sociales, les phénomènes d'innovation sont souvent considérés comme les seuls phénomènes d'allure aléatoire, hormis les éventuels éléments de libre arbitre associés au comportement.

Une *probabilité radicale* est relative à un phénomène dont l'explication mobilise à la fois des facteurs implicites irréguliers et des facteurs explicites mal connus. La probabilité reflète alors l'influence conjointe de ces deux types de facteurs (qui peuvent se confondre). Il en est ainsi d'une course hippique ou de toute compétition entre plusieurs entités (élection). Les facteurs qui interviennent sont très nombreux et agissent de façon enchevêtrée. La probabilité est radicale au sens où elle ne fait qu'introduire une certaine régularité là où elle ne s'impose pas *a priori*.

Ces distinctions sont faites par le modélisateur en fonction de ses connaissances du moment et peuvent varier au cours du temps. Des facteurs explicites peuvent être découverts par le modélisateur et transformer une probabilité irréductible en une probabilité maîtrisée. C'est ainsi que la radioactivité peut se voir expliquée par des facteurs plus profonds agissant comme des variables cachées. En sens inverse, des facteurs implicites peuvent être découverts et transformer une probabilité radicale en une probabilité maîtrisée. C'est ainsi que, dans une élection en milieu restreint, tous les facteurs pertinents peuvent être repérés sinon connus. Enfin, des facteurs non connus peuvent être de mieux en mieux observables et transformer un phénomène aléatoire en un phénomène déterministe. Il en serait ainsi si l'impulsion du dé pouvait être appréhendée, ce qui permettrait de parfaitement prévoir le résultat du lancer.

3.2. Probabilités ontologique et épistémique

Au regard de la philosophie traditionnelle, une deuxième typologie des probabilités est fondée sur la distinction entre une propriété attribuée au phénomène lui-même et une propriété « localisée » dans l'esprit du modélisateur :

- une *probabilité ontologique* (*chance* en anglais) reflète le degré d'occurrence d'un phénomène, indépendamment des états mentaux de son attributeur ; elle prend la valeur α pour un événement E si cet événement a une probabilité intrinsèque $\text{Ch}(E) = \alpha$ de se produire ;
- une *probabilité épistémique* (*credence* en anglais) reflète le degré de croyance assigné par l'attributeur à la survenance du phénomène, compte tenu de son incertitude ; elle prend la valeur α pour un événement E si l'attributeur i évalue sa survenance au degré $\text{Cr}_i(E) = \alpha$

La probabilité ontologique d'un événement est l'*objet* du jugement d'un attributeur, comme celui-ci peut se donner pour objet toute propriété physique d'une entité quelconque. De la même façon qu'il peut bien ou mal évaluer la longueur ou la masse d'un corps, il peut bien ou mal évaluer la probabilité accordée à un événement, mesurée selon une échelle supposée cardinale (voir §10). Par contraste,

la probabilité épistémique attachée à un événement est un *mode* du jugement que l'attributeur porte sur cet événement. Le spectre des valeurs de probabilités admet alors un grain plus ou moins fin, qui raffine la tripartition habituelle (croire que l'événement survient, croire que l'événement ne survient pas, croire que l'événement peut survenir ou non).

Bien entendu, la probabilité ontologique (postulée) $Ch(E)$ qui fait l'objet du jugement de l'attributeur est à découvrir. L'attributeur peut afficher son estimation de cette probabilité que l'on note $Ch^*(E)$. Cette dernière possède une nature épistémique et peut être interprétée comme une croyance certaine de l'attributeur sur $Ch(E)$. Un pluraliste (voir §3.4) peut même recourir à des probabilités hiérarchiques (voir §10) et analyser l'estimation $Ch^*(E)$ comme une probabilité épistémique définie sur une probabilité ontologique : $Cr_i (Ch(E) = \alpha) = 1$. Cette situation ne diffère pas de celle relative à la vitesse de la lumière qui est inconnue, mais fait néanmoins l'objet d'une estimation quasi consensuelle.

3.3. Probabilités objective et subjective

Dans la même veine, une typologie différente des probabilités est fondée sur le fait que le phénomène est appréhendé ou non de la même façon par tous les attributeurs :

- une *probabilité objective* caractérise une situation dans laquelle chaque attributeur dispose d'informations suffisantes sur le phénomène et raisonne idéalement sur elles ; une telle probabilité est *universelle* en tant qu'elle est la même pour tous les attributeurs ; pour un événement E , elle est notée $Po(E)$;
- une *probabilité subjective* caractérise une situation dans laquelle les informations d'un attributeur sont incomplètes et/ou ses capacités de raisonnement limitées ; une telle probabilité est *personnelle* au sens où elle diffère d'un attributeur à l'autre ; pour un événement E et un attributeur i , elle est notée $Ps_i(E)$.

Contrairement à la précédente distinction, cette dernière est plus affaire de degré que d'opposition binaire dans la mesure où une probabilité apparaît comme plus ou moins objective. La distinction demeure relativement floue en tant qu'elle renvoie à une notion d'information parfaite et de rationalité parfaite (« objectivité épistémique »). En théorie, l'information est parfaite quand tous les facteurs pertinents ont été explicités et observés, ce qui reste un vœu pieux. En pratique, l'information est jugée parfaite à partir du moment où une information supplémentaire ne remet plus en cause la probabilité retenue, ce qui introduit une manière d'autoréférence. Quant à la rationalité parfaite, elle renvoie au respect des axiomes des probabilités (voir §7).

L'univocité ou non du lien entre les informations et les probabilités est mise en évidence, dans le cas des probabilités épistémiques, par les deux attitudes extrêmes suivantes :

- pour l'objectiviste radical, tout ensemble d'informations détermine univoquement les degrés de croyances adoptés par un attributeur rationnel ; un corollaire affirme que la probabilité ne peut se modifier que si l'attributeur reçoit une information supplémentaire⁵,
- pour le subjectiviste radical, pour un ensemble d'informations donné, tous les degrés de croyances sont envisageables pour chaque attributeur, pourvu qu'ils obéissent aux axiomes des probabilités.

4. Comparaison des typologies

4.1. Croisement des typologies

Si on compare les deux premières typologies, on peut considérer que les probabilités irréductibles et approchées (première ligne du premier tableau) sont de nature ontologique. Elles correspondent à des facteurs tous connus et ne diffèrent que par leurs modes d'influence. En revanche, les probabilités maîtrisées et radicales

⁵ Pour une discussion récente de l'idée selon laquelle un ensemble d'informations détermine univoquement les probabilités, voir Levi (2010).

(deuxième ligne du premier tableau) sont de nature épistémique. Elles font en effet intervenir des facteurs mal connus de l'attributeur. Cela revient à dire que le seul élément d'incertitude de l'attributeur est le fait de ne pas pouvoir appréhender certains facteurs explicites. Cela est manifeste pour le jet de dé comme pour la course hippique. Cependant, un autre élément d'incertitude concerne la liste des facteurs qui influencent le phénomène, quelle que soit la possibilité de les mesurer.

Si on compare la première et la troisième typologie, on peut considérer que les probabilités irréductibles et maîtrisées (première colonne du tableau) sont objectives. En revanche, les probabilités approchées et radicales (deuxième colonne du tableau) sont subjectives. Cela revient à dire que la raison principale de l'objectivité est le fait de ne pas avoir de critères implicites ou de considérer ceux-ci comme parfaitement réguliers. La probabilité d'un corps radioactif est objective, car sa loi d'engendrement est parfaitement connue. La probabilité d'un lancer de dé est objective, car elle résulte d'un calcul fondé sur des données objectives relatives au lancement. Comme la loi de mouvement (et d'atterrissage) du dé est donnée par la mécanique, il est possible de définir, pour chaque condition initiale, le résultat du lancer. Dans l'espace des conditions initiales (direction et vitesse du lancer), il en résulte des zones d'attraction pour chacune des réalisations possibles du dé. Si l'on observe de plus que ces zones sont distribuées de façon homogène dans l'espace (compte tenu de l'indiscernabilité qui les frappe), on peut en déduire que les probabilités des différentes réalisations sont uniformes⁶.

Si on compare enfin les deux dernières typologies, on retrouve le tableau initial avec d'autres intitulés :

⁶ L'idée que des probabilités objectives peuvent émerger des propriétés d'un système physique déterministe, et en particulier de la manière dont il associe les conditions initiales aux types de conditions finales, est à rapprocher de la « méthode des fonctions arbitraires » que l'on fait souvent remonter à Poincaré (1896/1912). Elle a donné lieu à une série de résultats dont von Plato (1983) donne une présentation d'ensemble. Ces résultats montrent que, pour certains systèmes physiques, une très large classe de distributions sur les conditions initiales aboutissent (approximativement) à la même distribution sur les conditions finales. De tels résultats jouent un rôle important dans les discussions récentes sur la possibilité de probabilités objectives au sein de systèmes déterministes ; voir Strevens (2011).

	Probabilité objective	Probabilité subjective
Probabilité ontologique	probabilité irréductible <i>radioactivité</i>	probabilité approchée <i>innovation technique</i>
Probabilité épistémique	probabilité maîtrisée <i>lancer de dé</i>	probabilité radicale <i>course hippique</i>

Les quatre familles de probabilités obtenues par croisement des deux typologies apparaissent pertinentes en première instance. Cependant, on peut arguer que les deux critères qui sont croisés sont fortement congruents. D'une part, une probabilité ontologique apparaît comme plutôt objective dans la mesure où l'attributeur peut effectivement accéder à l'information nécessaire. D'autre part, une probabilité épistémique apparaît comme plutôt subjective dans la mesure où un degré de croyance n'est fondé *a priori* que sur des informations partielles.

Cependant, les catégories mixtes peuvent intuitivement avoir leurs raisons d'être. D'une part, il est possible d'émettre un jugement de probabilité ontologique et subjectif si une propriété aléatoire est attribuée à un phénomène, sans que son évaluation ne s'impose à tous. Ainsi, même si la radioactivité apparaît comme un phénomène primitif objectif, sa mesure peut faire l'objet de débats entre scientifiques. D'autre part, il est possible d'émettre un jugement de probabilité épistémique et objectif si le degré de croyance exprimé est fondé sur des informations suffisantes pour que tout attributeur soit contraint de l'endosser. Ainsi, tous les parieurs peuvent se focaliser sur une même estimation du résultat d'une course. Le périmètre de ces deux situations d'attribution dépend des normes de formation des degrés de croyances que l'attributeur accepte.

Surtout, le problème est de savoir si les deux formes de probabilité mixtes (approchée, maîtrisée) peuvent se rabattre sur les deux formes pures (irréductible, radicale). D'une part, une probabilité maîtrisée est réduite à une probabilité irréductible par l'artifice consistant à passer d'une situation spécifique à une situation générique. Les conditions du lancer de dé sont rendues opaques et l'attributeur ne considère que les résultats de lancers successifs qui ne diffèrent plus des rayonnements d'un corps radioactif. D'autre part, une probabilité approchée peut soit

basculer en une probabilité irréductible si une régularité est dégagée derrière l'aspect erratique, soit vers une probabilité radicale si certains critères mal connus se font jour.

Une pratique philosophique courante conduit ainsi à confondre les deux couples de notions, bien que la première (épistémique *vs* ontologique) soit conceptuellement mieux définie (Carnap, 1950/1962 ; Gillies, 2000). Les praticiens des probabilités ne font généralement pas non plus la distinction et emploient essentiellement le second vocabulaire (objectif *vs* subjectif), qui paraît plus opératoire. Par la suite, l'une ou l'autre distinction est introduite, selon qu'elle éclaire plus ou moins bien le problème à l'étude.

4.2. Attitudes face aux interprétations multiples

La pluralité d'interprétations proposée est en partie cristallisée dans des écoles (voir §6). L'attitude par rapport à cette pluralité des interprétations est elle-même éminemment variable. Les *monistes* pensent qu'une seule de ces interprétations est légitime alors que les *pluralistes* acceptent d'emblée une pluralité d'interprétations. Les *contextualistes* pensent que toute interprétation préconisée doit être indépendante de la situation analysée alors que les *non contextualistes* estiment que la situation considérée peut moduler les interprétations.

Par croisement des deux typologies, on obtient quatre positions possibles (les exemples correspondent à deux interprétations possibles a et b successivement dans une première et une deuxième situation) :

	Non contextualiste	Contextualiste
Moniste	une seule interprétation valable dans toute situation : (a,a)	une seule interprétation variable selon la situation : (a,b)
Pluraliste	plusieurs interprétations, chacune valable en toutes situations : (ab,ab)	interprétations diverses dans chaque situation : (a,ab)

Cette différence d'attitude est pertinente pour la distinction entre probabilités ontologique et épistémique. Le moniste considère qu'une seule des conceptions est légitime ou nécessaire dans toutes les situations pertinentes d'attribution de probabilités. Le moniste ontologique soutient que $Cr_i(E)$ se laisse en fait réduire à $Ch^*(E)$, tout degré de croyance n'étant en fait qu'une croyance en un degré d'occurrence. Le moniste épistémique défend la réduction de $Ch^*(E)$ à $Cr_i(E)$, toute croyance en un degré d'occurrence n'étant qu'une projection « dans les choses » de degrés de croyance personnels. Inversement, le pluraliste accepte plusieurs conceptions à la fois, $Ch^*(E)$ et $Cr_i(E)$ ayant simultanément un sens, mais différent. Il peut alors estimer que, dans un contexte d'attribution donné, une seule des deux conceptions est nécessaire ou bien que, dans certains contextes, les deux conceptions sont mobilisées.

5. Mesure des probabilités

On peut définir quatre modes de mesure des probabilités, fondés successivement sur leurs occurrences passées, sur leurs caractéristiques instantanées et sur leurs conséquences futures. Plus précisément, à chaque catégorie de probabilité préalablement définie correspond un mode naturel de mesure par le modélisateur :

	Probabilité objective	Probabilité subjective
Probabilité ontologique	probabilité irréductible : <i>fréquence</i>	probabilité approchée : <i>déclaration soumise à des contraintes subjectives</i>
Probabilité épistémique	probabilité maîtrisée : <i>estimation fondée sur des contraintes objectives</i>	probabilité radicale : <i>révélation</i>

Les probabilités *irréductibles* sont typiquement inférées d'une *fréquence* dans une suite S d'expériences. Bien entendu, cette évaluation ne s'applique que si le phénomène est répétitif dans des conditions analogues. La fréquence d'une modalité du phénomène n'est autre que le nombre d'expériences ayant conduit à cette

modalité rapporté au nombre total d'expériences. La fréquence « actuelle » considère une suite S finie qui ne permet que d'approximer la probabilité. La fréquence « hypothétique » (Reichenbach, 1949/1971, von Mises, 1928/1957) considère une suite virtuelle infinie, qui converge vers la probabilité (si elle existe) en vertu de la loi des grands nombres⁷. Ainsi, on estime les probabilités d'émission de particules radioactives en observant les fréquences correspondantes.

Les probabilités *maîtrisées* sont estimées à partir de *contraintes objectives* qui viennent encadrer les valeurs des probabilités et même les déterminer. Ces contraintes sont relatives aux caractéristiques factuelles, explicites et non connues, qui déterminent le phénomène examiné. Elles sont engendrées par les caractéristiques structurelles du système étudié. Il en est ainsi des conditions de symétrie ou d'invariance qui s'imposent à certains phénomènes à travers les lois qui les gouvernent. Ainsi, le lancement d'un dé de forme régulière avec une impulsion quelconque conduit à prévoir une équiprobabilité dans l'occurrence des valeurs faciales du dé.

Les probabilités *approchées* sont mesurées par simple *déclaration verbale* des acteurs, encadrée par des contraintes subjectives. Elles résultent d'une interrogation pure et simple de l'agent, qui doit toutefois accepter d'y répondre. Elles sont néanmoins souvent jugées peu fiables par rapport à des probabilités profondes dont l'agent est censé être pourvu. Elles peuvent être volontairement biaisées par des considérations stratégiques concernant la fourniture d'information à autrui. Elles sont aussi involontairement biaisées par une prise de conscience déformée par l'agent de ses propres croyances. Ainsi, tout individu peut afficher la probabilité qu'il attribue à une innovation technique en fonction de sa connaissance subjective des processus d'innovation.

Les probabilités *radicales* sont obtenues par *révélation* à partir des choix (en incertitude) effectués par les acteurs qui les portent. Elles résultent d'un processus d'abduction du modélisateur, qui révèle les probabilités sous-jacentes à partir des

⁷ La loi des grands nombres s'écrit, $\Pr (| \text{Ch} (E) - \text{Fr}_n (E) | \leq \varepsilon) \rightarrow 1$ si $n \rightarrow \infty$, où $\text{Ch}(E)$ est la probabilité ontologique supposée vraie du phénomène, $\text{Fr}_n (E)$ la fréquence du phénomène après n preuves et \Pr la probabilité calculée d'une hypothèse pour le modélisateur.

actions observées. Par exemple, il peut faire l'hypothèse qu'un individu croît au degré α qu'un cheval va gagner si α est la somme maximale qu'il est prêt à payer pour entrer dans un pari qui rapporte un euro si la proposition est vraie et zéro sinon. Les probabilités ainsi révélées ne sont pas forcément conscientes pour l'acteur. Elles sont souvent multivoques dans la mesure où elles résultent d'un raisonnement non déductif. Elles reposent sur un modèle de décision postulé, le plus souvent la maximisation de l'espérance d'utilité. Elles nécessitent la connaissance des préférences de l'acteur sur les conséquences des choix, qui déterminent ses choix conjointement aux croyances. Ainsi, les paris qu'un agent est amené à faire sur une course hippique révèlent ses croyances sous-jacentes.

Une même méthode de mesure peut fort bien s'appliquer à des catégories différentes de probabilités. En pratique, les méthodes se différencient essentiellement selon que les probabilités sont objectives ou subjectives. Pour les premières, les fréquences s'appliquent dès lors que le phénomène est véritablement répétitif (dans le temps ou au sein d'une population). Elles concernent aussi bien la radioactivité que le lancement de dé. Les estimations conformes à des contraintes objectives s'appliquent à tout phénomène pour lequel on dispose d'indications suffisantes sur ses causes, ce qui est plus restrictif. Elles valent pour le dé, mais non pour la radioactivité. Pour les secondes, les paris sur les issues possibles du phénomène s'appliquent dès lors qu'un acteur accepte de s'y soumettre. Ils concernent aussi bien une élection qu'une innovation. Les déclarations sont parfois plus difficiles à obtenir, du moins sous forme quantitative. Elles sont pertinentes dès lors que l'attributeur dispose d'un modèle suffisamment structuré du phénomène. Cela vaut pour une innovation, mais peut aussi s'appliquer à une élection.

En sens inverse, un même phénomène peut être appréhendé par des probabilités mesurées selon des méthodes différentes et pour des acteurs différents. Ainsi, la probabilité qu'une liste de symptômes soit associée à telle ou telle maladie peut être évaluée par les quatre méthodes. La fréquence avec laquelle les symptômes conduisent à la maladie peut être mesurée sur un échantillon de malades. Des contraintes peuvent aussi donner des ordres de grandeur des rapports de ces probabilités en fonction de la proximité des maladies. Par ailleurs, un médecin peut

fournir une estimation de cette probabilité en s'appuyant sur son expérience qualitative passée. La probabilité peut enfin être inférée à partir d'un pari que ferait ce médecin sur l'état du malade à travers une opération qu'il décide. Bien entendu, ces mesures peuvent diverger les unes des autres dans la mesure où elles obéissent à des logiques différentes et mobilisent des informations différentes.

6. Interprétations classiques

Dans le prolongement des typologies générales présentées au §3, des interprétations plus spécifiques des probabilités⁸ ont été proposées au cours de l'histoire. Ces interprétations sont censées fournir une analyse conceptuelle et parfois méthodologique d'un concept de probabilité mobilisé dans la pratique scientifique ou dans la vie ordinaire. Elles cherchent à répondre à des questions particulières qui se sont posées dans l'utilisation des probabilités.

L'approche fréquentiste, initiée par Jacques Bernoulli (1713), puis développée par Fisher (1925) et von Mises (1928/1957), repose sur l'*identification* (asymptotique) des probabilités à des fréquences d'apparition d'événements génériques dans des expériences successives (ou à des proportions, dans des expériences parallèles). Elle proscriit donc l'attribution de probabilités à des événements spécifiques ; selon von Mises, « c'est l'une des conséquences les plus importantes de notre définition des probabilités »⁹. Le fréquentisme est essentiellement justifié par la loi des grands nombres qui affirme que la fréquence d'un événement converge vers sa probabilité. Cette convergence fonctionne soit en probabilités (loi faible), soit presque sûrement (loi forte). L'interprétation fréquentiste est destinée à fournir un concept ontologique de probabilité : les probabilités sont censées traduire des propriétés empiriques de l'environnement de l'acteur ou du modélisateur, que ceux-ci peuvent découvrir.

L'approche propensionniste, introduite par Popper (1959), se range également sous le concept ontologique de probabilité. La « propension » qualifie le

⁸ voir Hajek (2010) pour une discussion générale.

⁹ Cité par Gillies (2000a), p.115.

pouvoir causal exercé par certaines conditions matérielles dans la production du phénomène étudié. Elle a été introduite pour rendre compte de la probabilité d'événements spécifiques, mais admet en fait deux versions. Selon la première (*long-run*, Popper, Gillies), la propension d'un ensemble de conditions répétables est sa tendance à produire des fréquences quand on les instancie de manière répétée. Selon la seconde (*single case*, Popper, Miller, Fetzer), la propension traduit le pouvoir exercé par un ensemble complet de conditions pertinentes pour produire un événement spécifique. Au regard du fréquentisme, le propensionnisme est ontologiquement moins parcimonieux dans la mesure où il postule des capacités du système à agir (causalement) sur le phénomène considéré. Il est méthodologiquement moins opérationnel dans la mesure où il ne fournit pas de protocole propre pour mesurer les probabilités.

L'approche logiciste a été introduite par Keynes (1921) et Carnap (1950/1962)¹⁰. Les « probabilités logiques » (ou inductives) sont parfois conçues comme qualitatives (voir §9), mais peuvent aussi être quantitatives. Surtout, elles s'appliquent d'emblée à des propositions plutôt qu'à des phénomènes matériels. Plus précisément, si p et q sont deux énoncés, la probabilité logique $Ind(p,q)$ est une probabilité conditionnelle qui traduit le degré auquel p implique q (maximal si p implique q et minimal si p implique non q). Elles permettent ainsi d'étendre les relations de conséquence logique à des relations probabilisées et endossent le même statut que celle-ci. Cela signifie, pour certains du moins, qu'elles ne sont ni des propriétés psychologiques des agents, ni des propriétés empiriques de notre univers. Elles sont néanmoins plutôt rapprochées de l'interprétation épistémique des probabilités, sous l'hypothèse suivante : si toutes les informations dont dispose un agent tiennent dans la proposition e , alors le degré auquel il doit croire une hypothèse h est donné par $Ind(e,h)$. Par ailleurs, les probabilités logiques sont présentées comme objectives. Elles se présentent même comme la meilleure illustration d'un objectivisme radical (voir §3.3). Cependant, cette approche n'est pas directement opérationnelle, car elle ne suggère aucun protocole précis de mesure des probabilités.

¹⁰ Pour une défense récente, voir Maher (2006). Pour une présentation succincte, voir Hajek (2010, § 3.2).

L'approche subjective a été développée par Ramsey (1931), de Finetti (1937) ou Savage (1954/1972). Elle développe une interprétation épistémique et subjective en termes de degré de croyance, mais qui se veut néanmoins opérationnelle. En effet, elle s'appuie sur l'idée que les croyances ne sont pas directement observables, mais se révèlent dans des actions virtuelles ou réelles qui sont fondées sur elles, en particulier les paris. Elle peut dès lors s'appliquer à une expérience spécifique pour autant que l'acteur concerné accepte d'exprimer un choix fondé sur le phénomène correspondant. Ce sont donc les propriétés « causales » des croyances dans l'action qui permettent de les mesurer. Cette approche est parfois renforcée par des théorèmes de convergence qui montrent que la probabilité subjectiviste converge vers une valeur asymptotique (indépendante de la probabilité *a priori*) si elle est révisée de façon bayésienne (voir § 7) au regard d'informations suffisamment riches.

Enfin, plus récemment, une approche bayésienne objective (Jaynes, 2003 ; Williamson, 2009) a été introduite. Elle est bayésienne parce qu'elle entend fournir une interprétation épistémique des probabilités. Elle est objective parce qu'elle prétend, à l'encontre des subjectivistes radicaux (de Finetti), que les contraintes qui pèsent sur les degrés de croyance d'un acteur rationnel sont bien plus fortes que le simple respect des axiomes de probabilités. Deux contraintes supplémentaires sont ainsi ajoutées au respect des axiomes : l'alignement par rapport aux fréquences empiriques partielles (« calibration »), le respect de contraintes théoriques (généralement la maximisation de l'entropie).

Il n'y a pas de correspondance univoque entre les quatre interprétations classiques et les catégories esquissées au §4. Cependant, des rapprochements s'imposent de façon évidente. Les approches fréquentistes et propensionniste sont clairement ontologiques, la première se voulant de plus objective. L'approche subjectiviste trouve dans les situations de probabilité radicale, épistémique et subjective, un terrain d'application privilégié. L'approche logique est bien plus ambiguë, car elle comporte une dimension à la fois épistémique (par l'évaluation des probabilités) et objective (par l'utilisation calculatoire des probabilités). Par ailleurs, certains auteurs se rangent sous la bannière des *monistes* non contextuels (Keynes, 1921 ; von Mises, 1928/1957 ; Popper ; 1959 ; de Finetti, 1974) alors que d'autres

sont *pluralistes* (Ramsey, 1931 ; Carnap, 1950/1962 ; Mellor, 1971 ; Lewis, 1980/1986 ; Gillies, 2000).

7. Combinaison de probabilités

Pendant longtemps, la définition la plus influente des probabilités, attribuée à Laplace (1812), a été la suivante : la probabilité d'un événement est le rapport entre le nombre de cas où cet événement se réalise et le nombre de cas possibles. La définition mathématique contemporaine des probabilités, qui prend en compte la possibilité d'un nombre infini de cas, remonte à l'axiomatisation de Kolmogorov (1933/1950). Elle conçoit la probabilité comme une fonction numérique qui prend pour domaine une algèbre \mathcal{A} sur un ensemble non vide W de mondes. La fonction P de \mathcal{A} dans \mathbb{R} est une fonction de probabilité si et seulement si les trois axiomes suivants sont vérifiés :

(P1) non négativité : $P(E) \geq 0$, pour tout $E \in \mathcal{A}$

(P2) normalisation : $P(W) = 1$

(P3) additivité finie : si $E \cap E' = \emptyset$, $P(E \cup E') = P(E) + P(E')$, pour tout $E, E' \in \mathcal{A}$

L'ensemble W comme l'ensemble \mathcal{A} résultent de conventions de l'attributeur. Le premier n'est jamais exhaustif quant aux caractéristiques du phénomène à prendre en compte. Le second ne considère pas tous les événements matériellement différents les uns des autres, mais un sous-ensemble. Les conditions P1 et P2 sont essentiellement techniques. Seule la condition d'additivité finie P3 est une véritable contrainte s'imposant aux probabilités. Elle s'étend en une condition d'additivité dénombrable P'3 en considérant un nombre dénombrable d'éléments de \mathcal{A} . Elle a pour conséquence la propriété suivante :

(P'3) croissance : si $E \subseteq E'$, alors $P(E) \leq P(E')$

Les probabilités ontologiques satisfont aux axiomes de Kolmogorov (du moins avec l'additivité dénombrable). Tout d'abord, on peut vérifier que les fréquences relatives à des suites finies d'expériences satisfont à ces axiomes. Ce résultat est maintenu dans le cadre proposé par von Mises¹¹. Il est moins évident pour d'autres interprétations ontologiques. A la suite de Lewis (1980/1986), certains auteurs ont recours à une stratégie de contournement. Ils postulent que les probabilités épistémiques s'alignent (dans certaines circonstances) sur les probabilités ontologiques (voir §9) et en déduisent que premières obéissent aux axiomes de Kolmogorov dans la mesure où les secondes le font.

Les probabilités épistémiques satisfont également aux axiomes de Kolmogorov (y compris l'additivité dénombrable). Ainsi, Savage (1954/1972) soumet les préférences d'un acteur entre des actes à une liste d'axiomes, dont certains sont spécifiques aux probabilités (en pratique, ceux des probabilités ordinales, voir §9). Dès lors, ses degrés de croyances obéissent aux axiomes de base de la théorie des probabilités, pour peu que l'on accepte certaines hypothèses sur la façon dont ces degrés de croyance s'expriment dans les choix. Ce résultat généralise l'argument du *Dutch Book* (de Finetti, 1937) qui s'applique aux paris que réalise un agent si son utilité est connue du modélisateur. Il affirme, que si l'agent ne respecte pas les axiomes usuels des probabilités, le modélisateur peut construire des séquences de loteries qui entraînent une ruine certaine de l'acteur.

Les probabilités logiques soulèvent un problème particulier dans la mesure où elles s'appliquent souvent à des propositions, notamment à des hypothèses. Ces hypothèses ne correspondent plus nécessairement à des phénomènes observables, mais à des structures sous-jacentes. En particulier, elles expriment des relations entre des phénomènes et plus précisément les critères phénoménaux et les critères structurels d'un phénomène. Elles peuvent être déterministes ou aléatoires. La probabilité d'une hypothèse est généralement interprétée de façon épistémique et traduit un degré de croyance. L'hypothèse est supposée vraie ou fausse, mais

¹¹ Voir Howson et Urbach (1993), pp. 208-9.

l'attributeur a une incertitude sur sa valeur de vérité. Cependant, la probabilité d'une hypothèse peut aussi être interprétée de façon ontologique en considérant qu'elle est inscrite dans la nature. La nature est alors supposée aléatoire, non seulement dans ses manifestations directes, mais dans ses structures profondes.

Ces hypothèses obéissent cependant au cadre général précédemment décrit sous beaucoup d'aspects. Ainsi, une hypothèse admet d'autres modalités même si l'une au plus est supposée vraie¹². Elle permet de construire des événements qui sont des ensembles d'hypothèses et sont eux-mêmes probabilisés. Elle donne lieu à des expériences successives qui sont autant de tests empiriques de l'hypothèse. Cependant, ce ne sont pas les modalités de l'hypothèse qui sont observées, mais les données relatives aux phénomènes qu'elle met en relation. En pratique, l'hypothèse (obtenue par abduction ou posée *a priori*) est testée face à ces données. Sa probabilité résulte précisément d'un mécanisme d'inversion des causes cherchant à donner une probabilité aux causes, connaissant les conséquences (voir §8). Par exemple, on fait l'hypothèse qu'une pièce de monnaie est équitable et on effectue des séries de tirages pour tester sa validité.

8. Révision des probabilités

La révision des probabilités consiste à transformer une probabilité *a priori* P^- en une probabilité *a posteriori* P^+ en cas de réception d'un message $M \in \mathcal{A}$. La plupart des règles de révision considèrent comme invariante l'algèbre \mathcal{A} sur l'ensemble des mondes possibles. Ce qui revient à exclure la prise en compte de nouvelles possibilités. La règle de révision la plus utilisée est la règle de conditionnalisation de Bayes (qui ne s'applique que si la probabilité *a priori* du message est non nulle) :

$$P^+(E) = P(E / M) = P^-(E \cap M) / P^-(M) \text{ si } P^-(M) > 0$$

¹² A une hypothèse bien formée correspond au moins deux modalités : celle où elle est fausse et celle où elle est vraie. Quand l'hypothèse fixe la valeur θ d'un paramètre dans un espace Θ , alors on peut considérer par extension que Θ est l'ensemble des modalités de l'hypothèse.

Cette règle exprime que les poids des mondes qui sont rendus caducs par le message sont ventilés sur les mondes qui étaient déjà et restent considérés comme possibles, de façon homothétique. D'autres règles sont aussi proposées. La règle d'« imaging » de Lewis (1976) exprime que le poids d'un monde qui est rendu caduc par le message est ventilé sur le monde qui reste possible le plus proche au sens d'une certaine métrique.

Trois contextes de révision sont traditionnellement considérés dans la littérature, selon la teneur du message reçu. Ils portent tant sur les critères phénoménaux que structurels et se présentent sous forme d'événements :

- dans le contexte de mise en cause (*revising*), le message écarte certaines modalités du phénomène. Il contredit ainsi la croyance initiale.
- dans le contexte de mise à jour (*updating*), le message indique que certaines modalités du phénomène ont changé dans un sens donné. Il ne contredit pas la croyance initiale dans la mesure où le système considéré a changé.
- dans le contexte de mise au point ou de focalisation (*focusing*), le message restreint la classe de référence du phénomène (par exemple à une sous-population dans la population entière). Il change ainsi le point de vue de l'attributeur sans lui fournir de véritable information nouvelle.

Dans l'exemple médical, des patients sont atteints de diverses variantes A, B et C d'une certaine maladie dans des proportions modulées en fonction de certaines caractéristiques des malades (sexe, âge). Un message de mise en cause indique que, compte tenu des conditions qui ont cours localement, la population considérée n'est pas touchée par la variante A (ou que les femmes ne sont jamais touchées). Un message de mise à jour indique que, compte tenu d'un médicament nouvellement apparu, tous les malades atteints de la variante A (ou toutes les femmes) sont guéris. Un message de mise au point indique que les variantes de la maladie (en particulier son absence) ne sont plus considérées que pour les malades des variantes B et C (ou pour les seules femmes).

La règle de Bayes peut être justifiée dans divers contextes et pour divers types de probabilités. En pratique, si elle est justifiée pour une probabilité ontologique, elle l'est aussi pour une probabilité épistémique en vertu du principe d'alignement (voir §9). L'inverse n'est évidemment pas vrai. Les justifications sont de plus de deux types, étudiées successivement :

- les justifications cognitives sont seulement fondées sur les probabilités considérées comme des croyances de l'attributeur, en dehors de toute utilisation de ces croyances ;
- les justifications pragmatiques sont fondées sur des décisions de l'attributeur supposées fondées à leur tour sur ses degrés de croyance.

Dans le contexte de *mise au point*, la règle de Bayes s'applique naturellement à toute forme de probabilité. En effet, la distribution de probabilités initiale est transformée homothétiquement en une nouvelle distribution sur une classe restreinte de mondes. À l'intérieur de chaque sous-classe, les rapports de probabilités traduisent les proportions de mondes possibles de chaque type et sont donc conservés. Par exemple, la nouvelle distribution de probabilités (supposée ontologique) des variantes de la maladie s'applique désormais aux seules femmes en appliquant la règle de Bayes.

Dans le contexte de *mise en cause*, la règle de Bayes s'applique de façon plus délicate aux probabilités ontologiques. En effet, ce n'est pas la probabilité vraie qui peut changer, mais seulement son estimation. De plus, le message donne une indication sur la distribution de probabilités, mais aurait pu en donner d'autres tout aussi pertinents, mais conduisant à des distributions finales différentes. Par exemple, si seule la variante A a cours dans la population (comme le sait le modélisateur), le message peut indiquer aussi bien que la variante C est absente ou que la variante B est absente. En revanche, pour des probabilités épistémiques, la règle de Bayes a reçu une justification cognitive directe en termes d'axiomes généraux de révision des croyances, transposés aux probabilités à partir des axiomes AGM (Walliser et Zwirn, 2002).

Dans le contexte de *mise à jour*, la règle de Bayes peut fort bien s'appliquer aux probabilités ontologiques. En effet, comme l'univers a changé, les probabilités ontologiques vraies Ch ont également pu changer. Les probabilités estimées Ch^* peuvent alors être adaptées de façon semblable. Comme condition nécessaire, il faut supposer que « rien d'autre ne s'est passé » dans le système que ce qu'indique le message. Dans l'exemple de la maladie, la règle de Bayes va permettre de calculer la nouvelle proportion de chacune des variantes, à condition que l'occurrence relative des variantes B et C n'ait pas changé. Par ailleurs, pour des probabilités épistémiques, c'est la règle de Lewis (*imaging*) qui découle des axiomes généraux de révision des croyances, mais elle coïncide avec celle de Bayes pour des messages exhaustifs.

Quant aux justifications pragmatiques de la règle de Bayes, elles s'appliquent essentiellement aux probabilités épistémiques. En effet, la règle est justifiée en théorie de la décision par les axiomes de Savage (1954/1972), interprétés en dynamique. Une justification semblable existe en termes de *Dutch Book*, là encore prolongé à des choix dynamiques (Lewis, 1999).

Dans ce qui précède, un message avait toujours une forme ensembliste. Ce message peut prendre une forme probabiliste : afficher une distribution de probabilités sur un ensemble d'événements possibles. Vu à travers un filtre imparfait, un dé est ainsi tombé sur le chiffre un avec une certaine probabilité et sur deux avec la probabilité complémentaire. La règle de Bayes a été généralisée à ce cas (Jeffrey, 1992), la probabilité *a posteriori* d'un monde n'étant autre qu'une pondération avec les probabilités correspondantes des probabilités *a posteriori* obtenues pour chaque événement appris.

Les probabilités conditionnelles permettent enfin d'envisager des règles d'inférence à base probabiliste, particulièrement pertinentes dans l'approche logique des probabilités. Ce raisonnement probabiliste peut être comparé au raisonnement déductif classique (pour le modélisateur) ou au raisonnement décrit par la logique épistémique (pour un acteur). La proposition E est remplacée par la proposition $\Pr(E) \geq \alpha$ ou $\Pr(E) = \alpha$, et il s'agit non plus de préserver la *vérité* de E, mais de préserver sa *probabilité* (Suppes, 1966 ; Adams, 1975). Il est facile de voir que

plusieurs des schémas d'inférence fondamentaux de la logique déductive n'ont pas de contrepartie probabiliste directe. Par exemple, avec le conditionnel matériel noté \supset , la monotonie permet d'inférer $E \wedge G \supset F$ ¹³ de $E \supset F$. Cependant, avec les probabilités conditionnelles, il se peut parfaitement que $\Pr(E/F) > \Pr(E \wedge G/F)$. Des remarques similaires peuvent être faites pour le *modus ponens* qui autorise à inférer F de E et de $E \supset F$. Si l'on conserve le conditionnel matériel, on obtient le résultat suivant : si $\Pr(E \supset F) = \alpha$ et $\Pr(E) = \beta$, alors $\alpha + \beta - 1 \leq \Pr(F) \leq \alpha$ ¹⁴. Si l'on préfère avoir recours à la probabilité conditionnelle, alors le résultat s'énonce ainsi : si $\Pr(E/F) = \alpha$ et $\Pr(E) = \beta$, alors : $\alpha\beta \leq \Pr(F) \leq 1$ ¹⁵ (Wagner, 2004).

Ce type de raisonnements, permet de construire des « réseaux bayésiens », qui se présentent sous forme de graphes reliant des variables selon des probabilités conditionnelles. Les probabilités se propagent dès lors dans ce réseau à partir de sources. Ces réseaux peuvent prendre la forme de « hiérarchies bayésiennes » lorsqu'il s'agit pour un acteur de traiter de problèmes aléatoires emboîtés. Dans tous ces cas, les probabilités conditionnelles apparaissent comme la notion première et les probabilités traditionnelles n'en sont que des cas particuliers. Si $P(E/F)$ est une probabilité conditionnelle de base, $P(E) = p(E/W)$.

9. Problèmes ontologiques et méthodologiques

Dans leur interprétation ontologique, les probabilités sont des propriétés d'événements. Comme en témoigne une bonne partie de la littérature philosophique sur les probabilités, ces propriétés s'avèrent difficiles à cerner. Trois discussions cherchent à les analyser plus avant.

Une première discussion porte sur l'existence ou non de probabilités ontologiques non-extrêmes dans un contexte déterministe. Elle oppose les *compatibilistes* (Levi 2010 ; Maher, 2009 ; Hofer, 2007) aux *incompatibilistes*

¹³ \wedge est le symbole de conjonction.

¹⁴ En particulier, si $\alpha = \beta = 1 - \varepsilon$ pour $\varepsilon \leq 1/2$, alors $\Pr(F) \geq 1 - 2\varepsilon$.

¹⁵ En particulier, si $\alpha = \beta = 1 - \varepsilon$ pour $0 \leq \varepsilon < 1$, alors $\Pr(F) \geq (1 - \varepsilon)^2$.

(Lewis, 1980/1986 ; Loewer, 2001). Dans les termes du §3, les premiers reconnaissent comme d'authentiques probabilités ontologiques les probabilités irréductibles et maîtrisées, tandis que les seconds n'acceptent que les probabilités irréductibles. En fait, la plupart des philosophes admettent l'existence de probabilités non-extrêmes dans le cas du lancement du dé, sous peine de limiter l'usage des probabilités aux situations primitives comme la radioactivité. Les incompatibilistes ont en outre la charge d'expliquer pourquoi nous sommes si enclins à considérer les propriétés probabilistes d'un dé comme des propriétés objectives dont il est doté.

Une deuxième discussion porte sur l'existence ou non de probabilités non-extrêmes d'événements spécifiques, qui peuvent s'exprimer à l'occasion d'une expérience particulière. Elle oppose les *particularistes* et les *non-particularistes*, les fréquentistes étant les représentants les plus revendiqués de la seconde famille. Les non-particularistes ont notamment la charge d'expliquer pourquoi et selon quels principes nos attributions de probabilités passent continuellement d'événements génériques à des événements spécifiques.

Une troisième discussion porte sur l'existence ou non de probabilités ontologiques non-extrêmes pour des phénomènes ayant déjà eu lieu. Elle porte donc sur l'évolution des probabilités au cours du temps, ces probabilités étant *temporelles* pour certains (Lewis, 1980/1986) et *non temporelles* pour d'autres (Hofer, 2007). Pour les premiers, si les probabilités *ex ante* sont quelconques, les probabilités *ex post* sont forcément extrêmes (bien qu'éventuellement inconnues). Si la proposition p affirme d'un événement qu'il aura lieu au temps t , alors, après t , la probabilité objective de x est 1 si l'événement a eu lieu et 0 sinon. Comme le dit Lewis, « *what's past is no longer chancy* ». Quant aux probabilités épistémiques, elles varient naturellement en fonction des informations reçues.

Par ailleurs se pose le problème du passage entre probabilités ontologiques et probabilités épistémiques. Ce passage n'est possible que lorsqu'elles sont toutes deux soigneusement définies sur un même objet. Cette exigence concerne plutôt les probabilités ontologiques, dans la mesure où les probabilités épistémiques sont toujours définies si l'on admet que toute incertitude est susceptible d'un pari. Lorsqu'il est possible, le passage des probabilités ontologiques aux probabilités

épistémiques se fait selon un *principe d'alignement*. Il affirme que les probabilités épistémiques Cr_i d'un acteur s'alignent purement et simplement sur les probabilités ontologiques Ch s'il vient à les connaître. Sous sa forme la plus simple proposé par Miller (1966), on peut l'exprimer ainsi :

$$\text{principe de Miller : } Cr_i(E / Ch(E) = \alpha) = \alpha$$

Une formulation plus sophistiquée est proposée par Lewis (1980/1986) sous la forme suivante (où la probabilité ontologique est datée) :

$$\text{principe principal : } Cr_i(E / [Ch_t(E) = \alpha] \wedge d) = \alpha$$

Dans cette expression, d est une proposition (ou un événement) qui est à la fois (i) compatible avec la proposition selon laquelle $Cr_t(E) = \alpha$ et (ii) admissible au temps t , c'est-à-dire qui n'influence les croyances de i à propos de E qu'en influençant ses croyances sur les chances de E . Ainsi, une proposition signalant que E a eu lieu n'est pas admissible. En revanche, une proposition « historique », rapportant des faits particuliers ayant eu lieu avant t , est admissible. Il en va de même d'une proposition concernant la dépendance de probabilités ontologiques à l'égard de l'histoire, en l'occurrence d'un conditionnel dont l'antécédent porte sur des faits de ce genre et dont le conséquent est une proposition sur la probabilité ontologique. Pour Lewis, le principe principal « semble capturer tout ce que nous savons à propos de la chance ».

Une variante du principe d'alignement intervient dans le contexte de focalisation. Soit E_g le résultat d'une expérience générique et E_s le résultat d'une expérience spécifique. Par exemple, dans le cas de la maladie, E_g concerne la maladie d'un individu d'un certain type et E_s celle d'un patient particulier. Le principe d'alignement amendé indique que la probabilité ontologique correspondant à l'expérience générique est simplement projetée en une probabilité épistémique sur l'expérience spécifique :

principe de focalisation : $Cr_i (Es / Ch(Eg) = \alpha) = \alpha$

Le principe de focalisation effectue une double opération. Il permet d'abord de passer d'une probabilité ontologique à une probabilité épistémique comme le principe d'alignement. Il permet ensuite de passer d'un événement générique à un événement spécifique. De ce point de vue, il permet de rendre compte du fait que l'on peut attribuer une probabilité à un événement spécifique comme l'affirment les particularistes. Ce principe pose cependant le problème de la distinction entre événement générique et spécifique du point de vue formel.

Le statut du principe d'alignement ne fait pas consensus. Lewis semble stipuler que le principe principal est un principe de rationalité, qui s'impose aux degrés de croyance de tout agent rationnel, mais ne le justifie pas. Van Fraassen (1989) conteste qu'on puisse justifier de manière non circulaire un principe d'alignement. Par contraste, Howson et Urbach (1993) défendent une version du principe d'alignement adaptée à la forme d'interprétation des probabilités qu'ils défendent (le fréquentisme hypothétique de von Mises) et proposent une justification pragmatique de leur version.

Le passage inverse des probabilités épistémiques aux probabilités ontologiques est moins conceptualisé. Il est généralement envisagé dans un cadre dynamique, en considérant un agent qui révisé successivement ses probabilités épistémiques. La question s'entend différemment selon que l'on est subjectiviste moniste ou pluraliste. Pour le moniste (de Finetti, 1937), les probabilités ontologiques n'existent pas. Il s'agit alors d'expliquer l'apparence d'objectivité de certaines attributions probabilistes, et en particulier le vaste accord intersubjectif qui existe entre celles-ci : à la faveur de l'apprentissage, c'est un « fait psychologique » (de Finetti) que les différences interindividuelles disparaissent. Pour le pluraliste, il s'agit en outre de montrer à quelles conditions les probabilités épistémiques vont converger vers les probabilités ontologiques, qui sont supposées exister indépendamment des premières. Ces idées s'illustrent dans différents résultats de

convergence¹⁶, qui s'obtiennent typiquement lorsque les informations sont suffisamment riches et nombreuses et lorsque les probabilités épistémiques initiales ne s'écartent excessivement ni du monde réel, ni les unes des autres.

10. Extensions des probabilités

La variante la plus proche de l'axiomatique de Kolmogorov concerne les *probabilités ordinales*, qui correspondent à des attributions du genre : « E est strictement plus probable que E' ». Les probabilités ordinales (dites également qualitatives ou comparatives) sont définies par une relation binaire sur une algèbre \mathcal{A} d'un ensemble non vide W (Kreps, 1988, Fishburn, 1994 ; Fine 1973). La relation binaire notée $>$, asymétrique et négativement transitive, est une probabilité ordinale si les trois conditions suivantes sont satisfaites pour tous $E, E', E'' \in \mathcal{A}$:

(PO1) non ($\emptyset > E$)

(PO2) $W > \emptyset$

(PO3) si $(E \cup E') \cap E'' = \emptyset$, alors $E > E'$ si $(E \cup E'') > (E' \cup E'')$

Une distribution de probabilité P sur \mathcal{A} représente une relation $>$ sur la même algèbre si et seulement si, pour tout couple d'événements E et E' ,

$$[P(E) > P(E')] \text{ si } [E > E']$$

¹⁶ Voir Earman (1992), chap. 6 pour un panorama récent et un examen de la portée philosophique de ces résultats ; et Gillies (2000a, pp. 69-83) pour une exposition et une critique du résultat fondateur de de Finetti (1937).

Si une relation binaire est représentable par une distribution de probabilité, alors elle satisfait (PO1)-(PO3). En revanche, une relation qui satisfait ces conditions n'est pas nécessairement représentable par une distribution de probabilité. Cependant, plusieurs propriétés suffisent pour garantir la représentabilité (parfois unique) d'une probabilité ordinale. Un cas très particulier de probabilités ordinales (ou cardinales) concerne les « probabilités booléennes ». Les probabilités des événements ne peuvent y prendre que les valeurs 0 ou 1.

A priori, les probabilités ordinales sont susceptibles des mêmes interprétations que les probabilités quantitatives usuelles. De fait, elles ont surtout retenu l'attention des tenants d'une interprétation épistémique des probabilités. Pour ceux qui sont attachés aux méthodes introspectives de révélation des degrés de croyance, les jugements comparatifs paraissent plus réalistes. Dans certaines méthodes comportementales de révélation, comme celle de Savage (1954/1972), les probabilités ordinales sont le point de départ de la mesure des probabilités quantitatives. Les jugements révélés par les choix binaires d'un acteur sont en effet du type : « l'acteur i juge E plus probable que E' ».

Une extension des probabilités de types différents concerne les *probabilités hiérarchiques*, qui admettent des jugements probabilistes d'un attributeur sur des probabilités déjà définies (Walliser et Zwirn, 2011). Pour ce faire, on définit des strates successives k formées respectivement de mondes, de méta-mondes et ainsi de suite (k -mondes). Au sein du niveau compris entre deux strates, l'acteur porte un jugement en chaque $(k+1)$ -monde sur les k -mondes inférieurs, jugement qui peut être de deux natures. Un jugement ensembliste associe à chaque $(k+1)$ -monde un sous-ensemble de k -mondes. Un jugement probabiliste associe à chaque $(k+1)$ -monde une distribution de probabilités sur les k -mondes. A partir de règles simples, il est possible de calculer une valeur inférieure et une valeur supérieure pour tout événement défini sur les mondes de base.

Si l'on considère simplement deux niveaux (trois strates), trois structures intéressantes ont été étudiées dans la littérature. Les probabilités bi-niveaux (Skrms, Baron, Kyburg) résultent de jugements probabilistes à chaque niveau (probabilités sur des probabilités). Les fonctions de croyances (Dempster, Shafer)

résultent de jugements ensemblistes au niveau inférieur et de jugements probabilistes au niveau supérieur. Les familles de probabilités résultent de jugements probabilistes au niveau inférieur et de jugements ensemblistes au niveau supérieur. Les deux dernières structures se révèlent interchangeables sous certaines conditions. Elles s'expriment comme restrictions des « capacités de Choquet », des mesures sur des événements qui satisfont aux axiomes P1 et P2 et à l'axiome de croissance. Elles permettent en particulier de prendre en compte les situations d'ignorance totale, ce que les probabilités usuelles ne permettent pas.

Les mondes de chaque niveau peuvent eux-mêmes être interprétés de deux manières. Les mondes physiques correspondent à des entités dotées d'une existence matérielle. A partir d'un monde physique, on définit une distribution de probabilités ontologiques sur les mondes de niveau inférieur. Les mondes psychiques correspondent à des états mentaux dont est doté un acteur. A partir d'un monde psychique, on définit une distribution de probabilités épistémiques sur les mondes de niveau inférieur. Un « principe de subordination » implique que les niveaux supérieurs sont forcément psychiques et les niveaux inférieurs physiques. C'est dire qu'un acteur ne peut qu'avoir des probabilités épistémiques sur des probabilités ontologiques. De plus, un « principe de non-schizophrénie » soutient que l'acteur ne peut former une probabilité épistémique (non extrême) sur ses propres croyances épistémiques de niveau inférieur (il ne peut hésiter entre deux croyances).

Si l'on s'en tient à un ou deux niveaux, deux structures seulement sont possibles (en négligeant le caractère ensembliste ou probabiliste des évaluations). Une croyance élémentaire porte simplement un jugement sur un phénomène. Ainsi, un acteur peut avoir un jugement sur la couleur d'une boule. Une croyance composite porte un jugement sur un méta-phénomène formé d'une collection de phénomènes. Ainsi, un acteur peut avoir un jugement sur la composition d'une urne formée de boules. Un jugement ensembliste traduit les compositions de l'urne que l'acteur considère comme possibles. Un jugement probabiliste traduit l'« ambiguïté » (incertitude de second ordre) de l'acteur sur le contenu d'une urne, c'est-à-dire son degré de confiance dans des compositions alternatives de l'urne. En revanche, un

acteur ne peut croire qu'il a des croyances alternatives sur la couleur d'une boule donnée.

Des principes généraux proches du principe d'alignement permettent de transformer des mondes physiques en mondes psychiques (principe de projection) et réciproquement (principe d'anti-projection). De façon corollaire, les probabilités ontologiques sont transformées en probabilités épistémiques et réciproquement. L'acteur prend à son compte les probabilités ontologiques qui lui sont fournies ou imagine un dispositif qui donne une réalité à ses probabilités épistémiques.

Ces transformations permettent de définir des équivalences entre les trois contextes usuels de changement des croyances (mise en cause, mise à jour, mise au point) en adaptant les structures concernées. Les règles de révision sont aisées à déterminer selon des principes simples pour des structures essentiellement ontologiques. Elles sont alors transposées pour réviser des structures plus épistémiques. On retrouve ainsi les nombreuses règles de révision des structures hiérarchiques qui avaient déjà été proposées sur des bases intuitives (Walliser et Zwirn, 2011). La plupart des règles permettent de réviser les probabilités même quand la probabilité *a priori* du message est nulle.

11. Application à l'inférence statistique

Les méthodes statistiques consistent à confronter un modèle aléatoire aux données selon deux points de vue. D'une part, l'*inférence statistique* permet de déterminer des paramètres inconnus des relations du modèle au regard des observations. D'autre part, des *tests statistiques* permettent de porter un jugement sur la validité des relations du modèle au regard des observations. Dans les deux cas, le paramètre joue le rôle d'une hypothèse que le statisticien accepte de probabiliser.

Dans l'inférence statistique, une variable aléatoire X est supposée avoir une distribution de probabilités $f(\Theta, X)$ dépendant d'un paramètre aléatoire Θ . Cette distribution est donnée *a priori* comme contrainte et n'est pas susceptible d'être modifiée par les observations. On dispose d'un échantillon d'observations de la variable X , soit $x = (x_1, x_2, \dots, x_i, \dots, x_I)$. On cherche à estimer la valeur du paramètre de façon à minimiser la « distance » entre la distribution théorique et les

observations. Pour ce faire, on peut disposer d'une distribution *a priori* sur le paramètre Θ , soit $\pi(\theta)$. On peut alors calculer la distribution de probabilités *a posteriori* de Θ , soit $\pi(\theta/x)$. Par ailleurs, on peut calculer la vraisemblance (*likelihood*) de l'hypothèse Θ , soit $\Pi_i(f(\Theta, x_i))$. Dans tous les cas, le statisticien recherche un intervalle de confiance à α % du paramètre $[\underline{\theta}, \theta]$, dont l'interprétation est variable¹⁷.

À titre d'exemple, l'opération de base du statisticien est la régression linéaire. Elle part d'un nuage de points correspondant à des couples d'observations (x_i, y_i) . Elle cherche les paramètres a et b de la droite d'équation $y = ax + b$ qui passe au plus près des points. La méthode usuelle, celle des moindres carrés, prend comme mesure de cette distance la somme des carrés des écarts (mesurés verticalement) entre la droite et les points. La droite de régression est censée représenter la loi qui gouverne la relation entre la variable explicative x et la variable expliquée y . Les aléas par rapport à la loi, traduits par les écarts des points à la droite, reçoivent deux interprétations. L'interprétation ontologique met l'accent sur des déviations qui se produisent par rapport à la loi. L'interprétation épistémique met l'accent sur la méconnaissance du modélisateur quant à la forme de la loi.

L'*approche classique* consiste à rechercher le paramètre par la méthode du maximum de vraisemblance (ou celle des moments). Cette méthode coïncide avec la méthode des moindres carrés sous des hypothèses classiques. De fait, la probabilité $f(\Theta, X)$ est *a priori* objective et le conventionnalisme du statisticien n'intervient que dans le choix de la règle du maximum de vraisemblance (ou d'une autre) pour calculer ou situer le paramètre. L'intervalle de confiance sur le paramètre signifie qu'il y a une probabilité $1-\alpha$ d'obtenir l'échantillon x tiré, si l'hypothèse est vraie : $\Pr(x / \Theta \in [\underline{\theta}, \theta]) \geq 1-\alpha$. Autrement dit, $(1-\alpha)$ % des intervalles de confiance associés à des échantillons x contiennent la vraie valeur θ , ce qui rend cette interprétation fort peu intuitive.

L'*approche bayésienne* résulte du calcul d'une probabilité *a posteriori* du paramètre à partir d'une probabilité *a priori*. De fait, si la probabilité $f(\Theta, X)$ est *a priori* objective, la probabilité $\pi(\theta)$ est subjective (attachée au statisticien) ainsi que

¹⁷ Pour une analyse précise des rapports entre approches différentes de l'inférence, voir Courgeau (2012).

la probabilité *a posteriori* $\pi(\theta / x)$. L'intervalle de confiance signifie qu'il y a une probabilité α que la valeur du paramètre soit dans l'intervalle : $\Pr (\Theta \in [\underline{\theta}, \theta] / x) \geq 1 - \alpha$. Cette fois, le paramètre se trouve bien dans un intervalle avec une probabilité α . Le problème essentiel est celui de l'origine de la probabilité *a priori*, que l'on peut pour le moins soumettre à des calculs de robustesse.

Par ailleurs, des tests statistiques ont pour but de comparer une hypothèse de référence (appelée hypothèse nulle) h_0 et l'hypothèse testée (appelée hypothèse alternative) h . Pour n'importe quelle hypothèse h , deux types d'erreurs sont considérés. Une erreur de premier type consiste à rejeter l'hypothèse bien qu'elle soit vraie. Une erreur de second type consiste à accepter l'hypothèse bien qu'elle soit fausse. Ces erreurs peuvent être probabilisées en fonction des observations. Un test possible consiste dès lors à accepter l'hypothèse si les erreurs des deux types sont inférieures à des seuils conventionnels.

12. Application aux sciences empiriques

12.1. Considérations générales

Nombre de modèles des sciences empiriques sont de nature stochastique. Deux interprétations sont directement associées aux aléas introduits dans la modélisation. L'interprétation ontologique suppose que le phénomène décrit est réellement de nature indéterministe. L'interprétation épistémique suppose que le modélisateur ne connaît pas la structure véritable du modèle (variables explicatives omises, forme analytique d'une relation erronée) et introduit des aléas pour combler sa méconnaissance. Une troisième interprétation intervient cette fois dans la confrontation du modèle aux données. Plus précisément, l'interprétation méthodologique introduit des probabilités pour tenir compte des erreurs de mesure dans les données.

Ainsi, les probabilités sont objectives dans nombre de circonstances, y compris en sciences sociales où les conditions des expériences sont difficilement répétables et indépendantes. En physique, la mécanique quantique introduit des probabilités (irréductibles) supposées intrinsèques concernant l'état instantané des particules

élémentaires. De même, la mécanique statistique introduit des probabilités liées à la distribution (trop complexe) des molécules dans un gaz. En biologie, des probabilités affectent les mutations aléatoires qui affectent les gènes d'une espèce animale ou encore les erreurs de transcription du génome. En sciences sociales, Durkheim définit selon une approche populationnelle la probabilité de suicide d'un individu en fonction de la religion qu'il pratique.

Les probabilités subjectives sont utilisées de façon plus exceptionnelle, sauf en matière de prévision pour apprécier l'incertitude des scénarios proposés. En physique, les prévisions météorologiques sont affectées de probabilités qui traduisent la méconnaissance des lois climatiques. En économie, les prévisions macroéconomiques sont soumises à des incertitudes d'origines multiples et sont également probabilisées. Elles permettent de construire des « cônes d'incertitude » à partir du présent, qui traduisent le fait que l'incertitude augmente avec l'horizon de prévision.

12.2. Cas de la théorie des jeux

Un domaine privilégié d'utilisation des probabilités en sciences sociales est la théorie des jeux, cadre général d'analyse des relations stratégiques entre les acteurs. Elle distingue trois sources principales d'incertitude qui affectent, certes le modélisateur, mais aussi les acteurs représentés dans le modèle. Ces incertitudes ont systématiquement été représentées sous une forme probabiliste, quelle que soient leur interprétation. Ce n'est que récemment que des mesures non probabilistes ont été introduites pour traduire la troisième forme d'incertitude (fonctions de croyances).

L'incertitude objective du modélisateur traduit un indéterminisme intrinsèque du système. Elle s'exerce dans le comportement des acteurs du fait de leur libre arbitre ou d'une main tremblante (décalage entre l'intention d'action et l'action elle-même). Elle intervient dans des phénomènes collectifs issus de nombreuses chaînes causales comme pour les innovations ou les chocs macroéconomiques extérieurs.

L'incertitude subjective du modélisateur provient de sa méconnaissance profonde de phénomènes complexes. Elle intervient déjà au niveau des aléas impondérables introduits dans les fonctions de production ou les fonctions de coût, qui mélangent des facteurs techniques et humains. Elle intervient aussi au niveau des

fonctions de comportement des acteurs qui ne prennent en compte qu'un nombre réduit de facteurs explicatifs. En particulier, elle intervient sur les états mentaux qui conditionnent les comportements (croyances, préférences).

L'incertitude subjective des acteurs porte sur les états de la nature et les « types » d'autrui (i.e. les caractéristiques qui déterminent son choix). Elle affecte leurs informations factuelles (observations passées), structurelles (structures permanentes) ou futures (prévisions). Elle s'exerce sur le contexte passif (états de la nature), les caractéristiques de leurs adversaires (actions, types) et même sur leurs propres caractéristiques (actions, types). On suppose en général que les acteurs adoptent une probabilité *a priori* de nature objective (commune à tous) sur les variables incertaines (*common prior assumption*). Ils reçoivent des messages de nature subjective sur ces variables au sens où ils diffèrent d'un individu à l'autre (ensembles d'information). Ils forgent enfin une probabilité *a posteriori* de nature subjective à partir des éléments précédents.

À titre d'exemple, on peut examiner l'interprétation donnée à une stratégie mixte mise en œuvre par un acteur, soit une distribution de probabilités sur ses stratégies pures. La première interprétation considère qu'il s'agit d'un comportement volontairement aléatoire de l'acteur qui peut s'expliquer de diverses façons. La deuxième interprétation renvoie au comportement moyen d'une population dont les membres poursuivent des stratégies pures différentes. La troisième interprétation renvoie à la vision incertaine que possède un acteur du comportement de son adversaire.

Une analyse similaire peut être faite des modèles de préférences incluant un élément stochastique, développés par la théorie et la psychologie de la décision et appliqués parfois en théorie des jeux. La première interprétation renvoie à des « états d'esprit » alternatifs que l'acteur peut avoir (en fonction éventuellement de circonstances extérieures). La deuxième interprétation renvoie à la dispersion des préférences au sein d'une population. La troisième interprétation traduit l'incertitude qui s'impose à l'acteur sur le « type » d'autrui (ce type résumant ses déterminants de choix).

13. Application à l'épistémologie

En philosophie, les probabilités sont utilisées principalement, mais pas exclusivement¹⁸, en théorie de la confirmation et du raisonnement inductif. Depuis les travaux fondateurs de Hempel (1945), cette théorie se donne pour objectif de clarifier et de codifier les notions de confirmation et d'infirmité telles qu'elles interviennent dans le raisonnement scientifique et ordinaire. On pose généralement la question de la manière suivante : si H est une hypothèse et si E est une proposition qui rapporte un certain nombre de données empiriques, à quelles conditions peut-on considérer que E confirme (resp. infirme) H ?¹⁹

Les théories de la confirmation ont d'abord été formulées dans un cadre logique, sans faire intervenir de probabilités. Elles adoptent aujourd'hui un cadre plus souvent probabiliste et s'appuient sur une interprétation épistémique des probabilités. On regroupe les travaux se rattachant à cette tendance dominante sous l'appellation de *théorie bayésienne de la confirmation* (TBC)²⁰.

La TBC opte pour un concept *incrémental* (et non pas absolu) de confirmation : il y a confirmation de H par E non pas si E rend H très probable, mais si E augmente la probabilité de H , i.e.

$$E \text{ confirme } H \text{ si } P(H / E) > P(H)$$

À ce concept qualitatif de confirmation, on fait souvent correspondre un concept quantitatif qui varie selon la mesure que l'on choisit. La popularité de la TBC provient de sa capacité à rendre compte d'un grand nombre d'intuitions concernant la confirmation. Elle permet de fournir des éléments de réponse, plus ou moins convaincants, à des problèmes confirmationnels classiques, comme le paradoxe des corbeaux ou le problème de Duhem-Quine. Elle conduit aussi aux résultats suivants :

¹⁸ Elles jouent également un rôle important dans les théories de la causalité et de la connaissance.

¹⁹ Pour une présentation générale, voir Zwirn et Zwirn (1996).

²⁰ Pour des discussions approfondies, voir Earman (1992) et Howson et Urbach (1993). Pour une présentation plus succincte, voir Cozic (2011).

- (1) toutes choses égales par ailleurs, plus une donnée E est probable étant donné une hypothèse H , plus H sera confirmée par E
- (2) toutes choses égales par ailleurs, *moins* la donnée E est probable *a priori*, plus l'hypothèse H sera confirmée par E (« principe de surprise »)
- (3) E confirme H si et seulement si $P(E | H) > P(E | \neg H)$

L'une des sources majeures de difficulté pour la TBC tient au fait qu'elle repose sur une interprétation subjective des probabilités. Rien n'empêche, de son point de vue, qu'une donnée E confirme H pour un individu i mais pas pour un autre individu j . Selon certains, ce libéralisme excessif empêche la TBC de rendre compte de l'objectivité à l'œuvre dans le raisonnement scientifique. Une réponse possible consiste à restreindre les applications de la TBC à certains cas « favorables » où les probabilités individuelles sont similaires, voire identiques. Il doit en être ainsi pour la probabilité conditionnelle $P(E/H)$, mais aussi pour $P(E)$. Il en est ainsi si les probabilités subjectives découlent de probabilités objectives par le principe d'alignement (ou celui de focalisation) ou si les probabilités subjectives convergent lorsque les informations sont suffisamment riches (voir §8). Cependant, on peut remettre en cause la TBC et caractériser le concept de confirmation uniquement en matière de vraisemblance (*likelihoodism*).

14. Conclusion

Deux interprétations extrêmes des probabilités se démarquent par leur caractère opérationnel. Les probabilités fréquentielles sont inhérentes aux phénomènes et sont obtenues par la simple mesure de leur fréquence d'occurrence dans des expériences répétées. Les probabilités subjectives sont des degrés de croyance des acteurs sur des phénomènes et sont obtenues par révélation à partir de leurs actions. Les autres types de probabilités se situent dans l'entre-deux et s'appuient sur des méthodologies moins rigoureuses. Pour les plus objectives, il n'y a plus véritablement d'expériences séquentielles, mais de simples données séquentielles ou populationnelles. Pour les plus subjectives, il n'y a plus forcément

d'acteurs bien identifiés, mais des estimations très libres de probabilités qui permettent des calculs.

Les approches épistémiques et subjectives des probabilités sont souvent qualifiées de « bayésiennes ». En fait, ce terme admet des acceptions multiples qu'il est possible de distinguer à partir des usages des probabilités dans le comportement d'un acteur quelconque. Au premier niveau, un acteur est bayésien s'il exprime l'incertitude sur son environnement sous forme de probabilités épistémiques. Au deuxième niveau, un acteur est bayésien s'il effectue ses raisonnements conformément au calcul des probabilités, en particulier en s'appuyant en dynamique sur la règle de Bayes. Au troisième niveau, un acteur est bayésien s'il prend ses décisions en fonction du critère de maximisation de l'espérance d'utilité. Les approches bayésiennes s'appliquent aux trois niveaux à un décideur quelconque qui perçoit et agit sur son environnement. Elles s'appliquent aussi au modélisateur, en particulier au statisticien et à l'épistémologue qui sont bayésiens au premier et deuxième niveau.

De nombreux ponts ont été établis entre les approches ontologique/objective et épistémique/subjective. Ces ponts ont surtout conduit à faire de l'approche objective un cas particulier ou asymptotique de l'approche subjective. Cela est vrai dans le calcul des probabilités pour lequel l'approche objective apparaît comme un cas limite de l'approche subjective en information complète. Cela est vrai aussi en statistiques pour lesquelles la statistique classique apparaît comme un cas particulier de la statistique bayésienne.

Références

- Adams, E. (1975), *The Logic of Conditionals*, Dordrecht, Hollande : Reidel.
- Carnap, R. (1950/1962), *Logical Foundations of Probability*, Chicago : University of Chicago Press.
- Courgeau, D. (2012), *Probability and Social Science : Methodological Relationships between the two Approaches*, Dordrecht, Holland, and New York : Springer.
- Cozic, M. (2011), « Confirmation et induction », in A. Barberousse, D. Bonnay et M. Cozic (eds), *Précis de philosophie des sciences*, Paris : Vuibert.
- Earman, J. (1992), *Bayes or Bust ? A Critical Examination of Bayesian Confirmation Theory*, Cambridge, MA : MIT Press.
- Fine, T. (1973), *Theories of Probability: An Examination of Foundations*, New York : Academic Press.
- de Finetti, B. (1937), « La prévision : ses lois logiques, ses sources subjectives », *Annales de l'I.H.P.*, 7(1), 1-68.
- de Finetti, B. (1974), *Theory of Probability*, 1, New York : Wiley, 1990.
- Fisher, R.A. (1925), *Statistical Methods for Research Workers*, Biological Monographs and Manuals Series, Edinburgh : Oliver and Boyd.
- Gillies, D. (2000), *Philosophical Theories of Probability*, London : Routledge.
- Hajek, A. (2010), « Interpretations of Probability », *The Stanford Encyclopedia of Philosophy (Spring 2010 Edition)*, in Edward N. Zalta (ed.),
URL = <http://plato.stanford.edu/archives/spr2010/entries/probability-interpret/>
- Howson, C. et P. Urbach, (1993), *Scientific Reasoning: The Bayesian Approach*, La Salle, IL : Open Court.
- Hoefler, C. (2007), « The Third Way on Objective Probability: A Sceptic's Guide to Objective Chance », *Mind*, 116(463), 549-596.
- Jaynes, E. (2003), *Probability Theory: The Logic of Science*, Cambridge : Cambridge University Press.
- Jeffrey, R. (1992), *Probability and the Art of Judgment*, Cambridge : Cambridge University Press.
- Keynes, J.M. (1921), *A Treatise on Probability*, London : Macmillan.
- Kolmogorov, A.N. (1933/1950), *Foundations of Probability*, New York : Chelsea Publishing Company.
- Laplace, P.S. (1812), *Théorie Analytique des Probabilités*, Paris : Coursier.
- Levi, I. (2010), « Probability Logic, Logical Probability and Inductive Support », *Synthese*, 172, 97-118.
- Lewis, D. (1976), « Probabilities of Conditionals and Conditional Probabilities », *The Philosophical Review*, LXXXV(3), 297-315.
- Lewis, D. (1980/1986), « A Subjectivist's Guide to Objective Science », *Philosophical Papers*, 2 (19), 83-132.
- Lewis, D. (1999), « Why Conditionalize? », *Papers in Metaphysics and Epistemology*, 403-7.

- Loewer, B. (2001), « Determinism and Chance », *Studies in the History and Philosophy of Modern Physics*, 32, 309–320.
- Maher, P. (2006), « The Concept of Inductive Probability », *Erkenntnis*, 65, 185-206.
- Maher, P. (2009), « Physical Probability », in C. Glymour, W. Wei et D. Westerstahl (éds), *Logic, Methodology and Philosophy of Science*, London : College Publications.
- Miller, D. (1966), *Critical Rationalism. A Restatement and Defence*, Chicago and Lasalle, IL : Open Court.
- von Mises, (1928/57), *Probability, Statistics and Truth*, New York : Macmillan.
- von Plato, J. (1983), « The Method of Arbitrary Functions », *British Journal for the Philosophy of Science*, 34, 37-47.
- Poincaré, R. (1896/1912), *Calcul des probabilités*, Paris : Gauthier-Villars.
- Popper, K. (1959), « The Propensity Interpretation of Probability », *British Journal for the Philosophy of Science*, 10, 25-42.
- Ramsey, F.P. (1931), « Truth and Probability », in R. Braithwaite (éd.), *The Foundations of Mathematics and other Logical Essays*, London : Routledge et Keegan Paul.
- Reichenbach, H. (1949/1971), *The Theory of Probability*, Berkeley : University of California Press.
- Savage, L. (1954/1972), *The Foundations of Statistics*, New York : John Wiley and Sons.
- Strevens, M. (2011), « Probability out of Determinism », in C. Beisbart et S. Hartmann (éds), *Probabilities in Physics*, Oxford : Oxford University Press.
- Suppes, P. (1966), « Probabilistic Inference and the Concept of Total Evidence », in J. Hintikka et P. Suppes (éds), *Aspects of Inductive Inference*, Amsterdam : North Holland, 49-55.
- Wagner, C. (2004), « *Modus Tollens* Probabilized », *British Journal of Philosophy of Science*, 50(4), 747-53.
- Walliser, B. et D. Zwirn (2002), « Can Bayes' Rule Be Justified by Cognitive Rationality? », *Theory and Decision*, 53, 95-135.
- Walliser, B. et D. Zwirn (2011), « Change Rules for Hierarchical Beliefs », *International Journal of Approximate Reasoning*, 52(2), 166-83.
- Williamson, J. (2009), « Philosophies of Probability », in A. Irvine, (éd.), *Handbook of the Philosophy of Science*, 4, 493-533.
- Zwirn, D. et H. Zwirn (1996), « Metaconfirmation », *Theory and Decision*, 41, 195-228.
- Zwirn, H. (2002), « Principe de Miller et révision des croyances probabilistes », in P. Livet (éd.), *Révision des croyances*, Paris : Lavoisier.