IHPST

Cahiers de Recherche

Série « Décision, Rationalité, Interaction »

Cahier DRI-2009-02

Confirmation et induction

Mikaël Cozic

IHPST Éditions
13, rue du Four
75006 Paris

Cahiers de recherche de l'IHPST

Série « Décision, Rationalité, Interaction »

Sous la responsabilité scientifique de :
Jacques Dubucs, Mikael Cozic,
Philippe Mongin.

Confirmation et induction[1]

Mikaël Cozic[2]

Résumé

La présente étude (à paraître dans A. Barberousse & ali., Précis de philosophie des sciences, vol.1, Vuibert) est consacrée aux théories contemporaines de la confirmation. On y présente l'objectif général que poursuivent, depuis le texte fondateur de Hempel (1945), ces théories, ainsi que les principales tentatives proposées pour atteindre cet objectif. On se concentre sur les fondements, les réalisations et les difficultés de la théorie qui domine le paysage philosophique actuel : la théorie bayésienne de la confirmation.

Mots-clés

confirmation, induction, théorie bayésienne de la confirmation, problème des données connues, logique inductive.

Abstract

This contribution (to appear in A. Barberousse & ali., Précis de philosophie des sciences, vol.1, Vuibert) is devoted to contemporary confirmation theories. We introduce the general aim of these theories as it is conceived since Hempel (1945)'s seminal paper, and the main existing tentatives. We focus mainly on the foundations, achievements and issues of the theory which currently dominates the philosophical landscape : bayesian confirmation theory.

Keywords

confirmation, induction, bayesian confirmation theory, old evidence problem, inductive logic.

Classification JEL

[1] Je tiens à remercier les participants et les organisateurs (Isabelle Drouet et Thierry Martin) du séminaire « Probabilité, Décision, Incertitude » où j'ai pu aborder une partie des questions que traite ce chapitre Je remercie vivement Anouk Barberousse, Isabelle Drouet et Ph. Mongin pour leurs relectures attentives de versions préliminaires de ce chapitre.

[2] IHPST, GREGHEC & DEC

Cahier DRI-2009-02

1 Introduction

Les hypothèses et théories des sciences empiriques sont, en principe du moins, confrontées à des données empiriques. On évalue ces hypothèses et ces théories en bonne partie à partir du résultat de cette confrontation. Il arrive que des données parlent en faveur d'une hypothèse ; il arrive également que des données soient défavorables à une hypothèse ; ou encore que des données soient plus favorables à une hypothèse qu'à une autre. On considère, par exemple, que l'avance du périhélie de Mercure parle *en faveur* de la théorie de la relativité générale et *en défaveur* de la théorie newtonienne ; ou que les données paléontologiques parlent *en faveur* de la théorie de l'évolution. Ces notions intuitives, qui semblent guider les scientifiques dans le développement et l'évaluation de leurs travaux, la philosophie des sciences les thématise sous le concept général de *confirmation*. Nous allons commencer par caractériser sommairement le concept de confirmation et la façon dont il est traité par l'épistémologie contemporaine avant d'entrer plus avant dans les théories de la confirmation.

1.1 Confirmation et théories de la confirmation

L'analyse philosophique de la confirmation se développe en général dans un cadre fortement idéalisé. On distingue les énoncés qui expriment des données empiriques, que l'on note canoniquement E. On note H une hypothèse ou une théorie, sans approfondir plus avant les différences entre les deux concepts. La discussion porte ainsi sur la question de savoir comment caractériser la confirmation qu'une donnée E apporte (ou n'apporte pas) à une hypothèse H. Plus précisément, on introduit dans ce contexte les quatre concepts cardinaux de confirmation, infirmation, vérification et réfutation. A titre de première caractérisation, on dira de données favorables à une hypothèse qu'elles la *confirment* ; de données défavorables à une hypothèse qu'elles l'*infirment*. « Favorables » et « défavorables » sont évidemment des notions très vagues. La notion de confirmation contient, nous semble-t-il, l'idée qu'une donnée est favorable à une hypothèse en ce sens qu'elle « supporte » ou qu'elle renforce notre confiance[1] dans la *vérité* de l'hypothèse H. C'est ce qui distingue, par exemple, la confirmation de la corroboration de Popper (voir ci-après). On peut concevoir les deux concepts, également célèbres, de *vérification* et de *réfutation* (ou falsification) comme des cas-limites de confirmation. Des données vérifient une hypothèse si elles la confirment maximalement, c'est-à-dire si elles établissent que l'hypothèse est vraie. A l'opposé, des données réfutent une hypothèse si elles la infirment maximalement, c'est-à-dire si elles établissent que l'hypothèse est fausse.

Certains ont contesté la légitimité et l'intérêt d'un concept comme celui de confirmation, nous y reviendrons. A supposer cependant qu'un tel concept guide le raisonnement scientifique, il est clair que son usage fait appel à des principes qui sont essentiellement tacites. La situation est analogue à celle du raisonnement mathématique : les mathématiciens, quand ils établissent leurs résultats, font appel à des principes logiques qu'ils n'explicitent pas ou peu. C'est à la logique (déductive) qu'il revient de dégager, de codifier et d'analyser les principes du raisonnement mathématique. De la même façon, on peut concevoir l'étude par le philosophe des sciences de la confirmation comme consistant, en partie du moins, à dégager, codifier et analyser les principes des raisonnements qui font appel au concept de confirmation[2]. Ainsi que le dit Hempel (1945), l'objectif d'une *théorie de la confirmation* est de fournir « une approximation raisonnablement proche de la conception de

[1] A ce stade, nous ne voulons pas prendre parti sur la question de savoir si la notion de confirmation est « subjective ». Nous n'excluons donc pas que le « renforcement de confiance » en question soit fondé objectivement.

la confirmation qui est implicite dans la procédure scientifique et dans la discussion méthodologique ».

1.2 Confirmation et déduction

La logique mathématique moderne a codifié avec succès le raisonnement déductif : elle a caractérisé rigoureusement l'idée intuitive selon laquelle un ensemble de prémisses Γ a pour conséquence logique un énoncé A si, et seulement si, il est impossible que les prémisses contenues dans Γ soient vraies tandis que A serait faux. La logique propositionnelle ou la logique du premier ordre donnent des exemples d'une telle caractérisation. La relation de conséquence logique joue un rôle important dans le traitement conceptuel et formel de la relation de confirmation. Tout d'abord, la *vérification* d'une hypothèse H par une donnée E (ou par un ensemble fini de données $E_1, \ldots, E_n$) correspond au cas où E implique logiquement H. La *réfutation* d'une hypothèse H par une donnée E correspond à celui où E implique logiquement $\neg H$[3]. Si un macro-économiste défend l'hypothèse H = « la croissance française s'élèvera à 1.5 % en 2009 », alors normalement on sera en mesure, à la fin de l'année 2009, d'obtenir un ensemble fini de données qui vérifieront ou réfuteront H. Mais, nous l'avons déjà dit, vérification et réfutation ne sont que des cas-limites. En toute généralité, la relation de confirmation diffère de celle de conséquence logique. En effet, une donnée E peut confirmer (resp. infirmer) une hypothèse H sans que H (resp. $\neg H$) soit conséquence logique de E. C'est même, à vrai dire, le cas normal : si H est un énoncé universel du type « Tous les P sont Q », alors en général on considère qu'un énoncé E du type « a est P et Q » (qu'on appelle une « instance positive ») confirme H alors que « a est P et Q » n'a évidemment pas pour conséquence logique « Tous les P sont Q » : il se pourrait que E soit vraie sans que H le soit. Popper a fait valoir avec force que l'essentiel des hypothèses scientifiques, dans la mesure où leur forme logique est celle d'énoncés universels (et où leur domaine de quantification n'est pas fini[4]), ne sont pas *vérifiables* : un ensemble fini de données empiriques ne peut logiquement impliquer un énoncé de ce genre. De leur côté, Duhem et Quine ont fait valoir que, bien souvent, des hypothèses scientifiques isolées ne sont pas *réfutables* par des données empiriques car il faut leur adjoindre des *hypothèses auxiliaires* pour qu'elles aient des implications observables (« problème de Duhem-Quine »). Le cœur des théories de la confirmation réside dans ce qui se passe « hors » des cas-limites de la vérification et de la réfutation : si E n'a pas pour conséquence logique H (resp. $\neg H$), dans quelles conditions peut-on dire que E confirme (resp. infirme) H ?

1.3 Déduction et induction

On distingue souvent le raisonnement déductif du raisonnement inductif, que l'on illustre par certaines formes typiques de raisonnement[5]. La première de ces formes est (1) la *généralisation* ou *induction énumérative* par laquelle on infère un énoncé universel comme

[2] Carnap prend l'analogie à son compte : dans « Inductive Logic and Science », il affirme que l'objectif de la logique inductive est analogue à celui de la logique déductive : il n'est pas de proposer de nouvelles façons de raisonner, mais d'*expliciter* nos manières habituelles de raisonner.

[3] Voir Mongin (à paraître).

[4] Popper (1959) parle d' « universalité numérique » pour les énoncés universels dont le domaine de quantification (implicite ou explicite) est fini (par exemple, « Tous les êtres humains vivant aujourd'hui font moins de trois mètres de hauteur »), d' « universalité au sens strict » pour les autres (par exemple, « L'énergie des oscillateurs harmoniques ne tombe jamais en dessous d'un certain point »).

[5] Nous nous inspirons ici de Vickers (2006).

« Tous les *P* sont *Q* »

d'un ensemble d'instances positives de cet énoncé

« a_1 est *P* et *Q* »,

« a_2 est *P* et *Q* »,

…,

« a_n est *P* et *Q* ».

Une autre forme de raisonnement inductif qui est souvent évoquée est (2) l'*inférence singulière*: on infère

« *b* est *Q* »

de

« a_1 est *P* et *Q* »,

« a_2 est *P* et *Q* »,

…,

« a_n est *P* et *Q* », et

« *b* est *P* »

Le raisonnement inductif est parfois réduit à l'une de ces deux formes, ou aux deux[6]. Ce n'est pas satisfaisant du point de vue conceptuel : le raisonnement inductif n'est, en toute généralité, ni un raisonnement qui irait du particulier au général, ni un raisonnement qui irait du passé vers le futur. La caractéristique centrale d'un raisonnement inductif est qu'il est *ampliatif* : ses prémisses n'impliquent pas logiquement sa conclusion et il y a, comme on dit parfois, « plus » dans la conclusion que dans les prémisses. L'induction au sens étroit correspond aux formes célèbres de raisonnement inductif comme la généralisation ou l'inférence singulière ; l'induction au sens large correspond au raisonnement ampliatif en

[6] Voir par exemple Mill (1843, Livre III, Chap.II, § 1) : « L'induction est le processus par lequel nous concluons que ce qui est vrai de certains individus d'une classe est vrai pour la classe entière, ou que ce qui est vrai à certains moments sera toujours vrai dans des circonstances similaires. » Il faut toutefois noter que, pour Mill, il y a véritablement induction quand les prémisses ne parcourent pas exhaustivement la classe considérée. C'est le cas quand la conclusion exprime une « universalité au sens strict » au sens où Popper emploie ce terme (voir note précédente).

général. Il existe de très nombreuses familles de raisonnement ampliatif. Le raisonnement statistique en offre des exemples célèbres[7]. (3) L'*inférence directe* consiste à inférer une proposition sur un échantillon à partir d'une proposition portant sur la population entière :

<u>80 % des malades réagissent bien au traitement Viralyse</u>

80% des malades de l'Hôpital Velpeau réagissent bien au traitement Viralyse

(4) L'*inférence prédictive* consiste à inférer une proposition sur un échantillon à partir d'une proposition portant sur un autre échantillon. :

<u>80 % des malades de l'Hôpital Velpeau réagissent bien au traitement Viralyse</u>

80 % des malades de l'Hôpital Urgo réagissent bien au traitement Viralyse

(5) L'*inférence inverse* consiste à inférer une proposition sur la population entière à partir d'une inférence sur un échantillon.

<u>80 % des malades de l'Hôpital Velpeau réagissent bien au traitement Viralyse</u>

80 % des malades réagissent bien au traitement Viralyse

La logique déductive confond dans la classe des inférences non valides l'ensemble des inférences ampliatives. Du point de vue intuitif pourtant, certaines inférences ampliatives sont meilleures que d'autres : dans certains cas, les prémisses confèrent une confiance très forte dans la conclusion, dans d'autres non. Considérons par exemple (2) l'inférence singulière: pour une même conclusion « *b* est *Q* », il semble que si l'on compare une inférence qui se base sur deux instances positives à une inférence qui se base sur ces deux instances positives et sur mille autres, la seconde inférence est clairement à son avantage. Considérons par ailleurs l'inférence (2') qui infère

« b est $\neg\ Q$ »

de

« a_1 est P et Q »,

« a_2 est P et Q »,

…,

« a_n est P et Q », et

« b est P »

Du point de vue de la logique déductive, (2) et (2') ont un statut semblable : ce ne sont pas des schémas d'inférence valides, la vérité de leurs prémisses n'entraîne pas

[7] Nous utilisons dans ce qui suit la terminologie de Carnap (1950/1962, §44).

nécessairement celle de leur conclusion. Pourtant, on se fierait beaucoup plus volontiers à (2) qu'à (2').

1.4 Induction et confirmation

L'induction (au sens large) et la confirmation sont apparemment des notions extrêmement proches l'une de l'autre. Comparons par exemple les prémisses *P* et la conclusion *C* d'un raisonnement inductif à la donnée empirique *E* et l'hypothèse *H* d'une relation de confirmation. (i) En général, *P* n'implique pas logiquement *C*, de la même façon que *E* n'implique pas logiquement *H*. (ii) *P* et *E* donnent en principe une certaine confiance dans la vérité, respectivement, de *C* et de *H*. (iii) Cette confiance est affaire de degré et peut être plus ou moins grande[8].

Pour Carnap (1950/1962), le problème de l'induction est essentiellement le même que le problème de la relation de confirmation entre hypothèses et données. En première analyse, il y a cependant certaines différences qu'il convient de noter. Tout d'abord, le raisonnement ampliatif n'est en principe nullement limité à des prémisses qui seraient des données empiriques et des conclusions qui seraient des hypothèses ou des théories. La notion de confirmation comporte donc, en ce sens, des restrictions de domaine par rapport à la notion plus générale d'inférence inductive ou ampliative. Ensuite, quand nous disons qu'une donnée *E* confirme une hypothèse *H*, il n'est pas clair que nous voulions dire que l'inférence de *E* à *H* est un bon raisonnement ampliatif ou que la force inductive de cette inférence est élevée. Nous pouvons vouloir dire que *E* augmente notre confiance en *H*[9]. Cette seconde remarque n'est toutefois pas décisive, dans la mesure où il se peut que de telles ambiguïtés se trouvent à la fois dans la notion d'inférence ampliative et dans celle de confirmation. Quoi qu'il en soit des contours exacts de nos concepts pré-théoriques, le noyau commun (i)-(iii) à l'induction et à la confirmation nous semble philosophiquement crucial : (i)-(iii) consiste à affirmer l'existence d'une notion de *support inductif*. Les théories du raisonnement inductif et celles de la confirmation cherchent précisément à saisir cette notion de support inductif.

1.5 Popper contre l'induction et la confirmation

Nous venons de voir qu'il existe des liens étroits entre la confirmation et le raisonnement inductif. Avant d'entamer l'exposition et la discussion des théories de la confirmation, il nous faut préciser que la notion de confirmation – plus précisément, la thèse selon laquelle il existe un support inductif -, ne fait pas l'unanimité parmi les philosophes des sciences contemporains.

Parmi eux, Popper est l'un de ceux qui s'y est opposé de la façon la plus spectaculaire : « Le mieux que nous puissions dire relativement à une hypothèse est qu'elle a été jusqu'à présent capable de prouver sa valeur et qu'elle été plus féconde que d'autres, bien qu'en principe l'on ne puisse jamais la justifier, la vérifier ni même prouver qu'elle est probable. Cette évaluation de l'hypothèse repose seulement sur les conséquences *déductives* (les prédictions) que l'on peut en tirer : *il n'est même pas nécessaire de mentionner l'induction.* »[10] Le raisonnement scientifique est en effet pour Popper essentiellement déductif : il s'agit de déduire (au sens strict) les conséquences observationnelles d'une

[8] Notons que toutes les théories de la confirmation n'ont pas l'ambition de rendre compte du caractère graduel de la confirmation.

[9] Nous reviendrons plus tard sur cette distinction en contrastant le concept absolu et le concept incrémental de confirmation.

[10] Popper, 1959, trad.fr. (modifiée) p.321.

hypothèse et ensuite de comparer ces conséquences aux données empiriques[11]. S'il y a désaccord entre les conséquences observationnelles et les données empiriques, l'hypothèse H est réfutée. Jusqu'à ce point, la logique seule suffit[12]. Que se passe-t-il si *H n'est* pas réfutée par les données empiriques ? L'hypothèse fondamentale des théoriciens de la confirmation est que, du point de vue épistémologique, il *peut* se passer quelque chose d'important : il se peut que H soit confirmée et que notre confiance en la vérité de H s'en trouve renforcée. Pour Popper, il n'y a rien de tel. Si H survit à un ou plusieurs tests, alors H est « *corroborée* » - Popper emploie ce terme pour marquer une différence avec la notion de confirmation[13]. Plus H survit à des tests empiriques, plus ces tests sont sévères et plus H se met en danger lors de tels tests (plus H est « réfutable »), plus le *degré de corroboration* de H est élevé. Mais le degré de corroboration ne reflète pas notre confiance dans la vérité de H.

Il faut bien mesurer à quel point l'anti-inductivisme de Popper est radical : l'hypothèse H peut bien survivre à un grand nombre de tests, du point de vue popperien il n'y aura pas de raison supplémentaire d'avoir confiance dans la vérité de H. Pour un partisan convaincu de Popper, notre chapitre devrait probablement s'arrêter ici puisque l'hypothèse de travail fondamentale des théories de la confirmation est précisément que des données empiriques peuvent augmenter notre confiance dans la vérité d'une hypothèse sans nécessairement l'impliquer. La position de Popper a néanmoins essuyé bon nombre de critiques. Nous nous contenterons d'en citer une parmi les plus fameuses, que l'on doit à W. Salmon (1981). Salmon se place dans un contexte pratique, où un agent doit prendre des décisions sur la base de son évaluation de différentes hypothèses. La corroboration d'une hypothèse porte exclusivement sur ses performances passées, si ce n'était pas le cas la notion comporterait une dimension inductive. Il objecte alors qu'on ne voit pas comment une telle notion pourrait rationnellement fonder les prédictions pertinentes pour la situation de décision dans laquelle se trouve l'agent : même si H_1 est très fortement corroborée alors que H_2 l'est très peu, on voit mal ce qui dans la théorie de Popper contraint notre agent à s'appuyer sur H_1 plutôt que sur H_2 pour guider son action puisque rien ne contraint sa confiance dans les prévisions basées sur H_1 et H_2. *A contrario*, l'une des forces majeures de la théorie bayésienne de la confirmation que nous présenterons ci-après est qu'elle est intégrée dans une théorie de l'action rationnelle (la théorie bayésienne de la décision).

1.6 Menu

Ce chapitre est consacré aux tentatives qui ont été faites, depuis une soixantaine d'années, pour élaborer une théorie de la confirmation. Nous allons présenter, discuter et illustrer les principales théories en présence. La section 2 portera sur les célèbres paradoxes de la confirmation et exposera les deux principales théories qualitatives de la confirmation, la théorie instantialiste et la théorie hypothético-déductive. La section 3 exposera les principes fondamentaux du bayésianisme, qui sert de fondements à la théorie de la confirmation qui

---

[11] Popper, 1959, trad.fr, pp. 28 et *sq*.

[12] Il faut bien distinguer le déductivisme popperien et la théorie hypothético-déductive de la confirmation (THDC) que nous présenterons ci-après. Ces deux conceptions ne font appel, parmi leurs notions primitives, qu'aux concepts de la logique déductive. Et toutes deux ne prennent en compte que les implications logiques des hypothèses auxquelles on s'intéresse. Mais Popper s'en tient au raisonnement déductif pur tandis que la THDC construit une notion non-déductive à partir du concept de conséquence logique (la HD-confirmation).

[13] Voir la note qui précède la section 79 : « J'ai introduit les termes de corroboration et particulièrement de degré de corroboration dans mon ouvrage parce que je souhaitais disposer d'un terme *neutre* pour exprimer le degré auquel une hypothèse a résisté à des tests sévères et a ainsi « fait ses preuves ». Par « neutre », j'entends un terme ne préjugeant pas de la question de savoir si, en résistant à ces tests, l'hypothèse devient « plus probable » au sens du calcul des probabilités. » (p. 256)

domine le paysage philosophique actuel, la théorie bayésienne de la confirmation. La section 4 est consacrée à la théorie bayésienne. La dernière section aborde, du point de vue bayésien, les questions de la justification et de l'objectivité de la confirmation et du raisonnement inductif.

2 Instantialisme et hypothético-déductivisme

Nous allons commencer notre étude des théories de la confirmation en présentant deux théories élémentaires de la confirmation, l'instantialisme (Hempel) et l'hypothético-déductivisme. Ces deux théories sont des théories *qualitatives* de la confirmation : elles n'élaborent pas de mesure de la confirmation, mais un critère qui, pour une donnée empirique E et une hypothèse H, permet simplement de dire si E confirme H. Avant de présenter ces théories, nous allons voir que, même dans ce cadre élémentaire, les théories de la confirmation doivent surmonter de redoutables difficultés. Nous donnerons deux exemples de ces difficultés : le paradoxe des corbeaux et le paradoxe de Hempel, qui sont tous les deux exposés et analysés dans la contribution fondamentale d'Hempel (1945).

2.1 Le paradoxe des corbeaux

La construction d'une théorie de la confirmation, en dépit, peut-être, des apparences, n'est pas chose triviale. La manifestation la plus spectaculaire des difficultés engendrées par la théorie de la confirmation est le célèbre *paradoxe des corbeaux* qui montre à quel point il peut être difficile de faire coexister certaines propriétés intuitives de la notion de confirmation. Supposons que l'hypothèse considérée H soit un énoncé universel de la forme :

« Tous les corbeaux sont noirs »,

ce que l'on symbolise canoniquement par

$$\forall x(\mathrm{C}x \rightarrow \mathrm{N}x)$$

en logique du premier ordre. Du point de vue confirmationnel, il semble naturel de considérer que si l'on observe une entité qui possède à la fois les propriétés (exprimées par les prédicats) C et N, alors cette observation confirme l'hypothèse H. Dans la symbolisation canonique, cela signifie qu'un énoncé comme $(\mathrm{C}a \wedge \mathrm{N}a)$ confirme $\forall x(\mathrm{C}x \rightarrow \mathrm{N}x)$. Pour reprendre l'exemple de Hempel (1945) : l'observation d'un corbeau qui est noir confirme l'hypothèse selon laquelle tous les corbeaux sont noirs. Rappelons que $(Ca \wedge \mathrm{N}a)$ est une *instance positive* associée à l'énoncé $\forall x(\mathrm{C}x \rightarrow \mathrm{N}x)$. Le principe que nous venons de formuler est quant à lui généralement appelé le *Critère de Nicod* : il affirme qu'une instance positive confirme l'énoncé universel associé.

L'une des contraintes les plus naturelles sur la relation de confirmation est certainement la *Condition d'Equivalence* : si une donnée E confirme (resp. infirme) une hypothèse H, alors elle confirme (infirme) tout énoncé H' qui est logiquement équivalent à H. La Condition d'Equivalence a un attrait normatif extrêmement fort : la rejeter signifierait, comme le dit

Hempel (1945), que la relation de confirmation dépend de la manière dont l'hypothèse est exprimée. Mais l'acceptation conjointe de ces deux principes (le Critère de Nicod et la Condition d'Equivalence) conduit à des conclusions paradoxales.

Considérons en effet l'énoncé universel « Tous les corbeaux sont noirs ». Cet énoncé est équivalent à l'énoncé « Toutes les choses non-noires sont des non-corbeaux » ($\forall(\neg$ N$x \rightarrow \neg$ Cx)). Par conséquent, en vertu de la Condition d'Equivalence, une donnée *E* confirme « Tous les corbeaux sont noirs » ssi elle confirme « Toutes les choses non-noires sont des non-corbeaux ». Considérons ensuite un énoncé qui implique le fait que l'objet *a* est à la fois non-noir et non-corbeau – par exemple, « *a* est une chaussette blanche ». En vertu du Critère de Nicod, « *a* est une chaussette blanche » confirme « Toutes les choses non-noires sont des non-corbeaux ». On en déduit que « *a* est une chaussette blanche » confirme du même coup « Tous les corbeaux sont noirs », ce qui est pour le moins contre-intuitif.

2.2 Le paradoxe de Hempel

On doit également à Hempel (1945) le second paradoxe que voici. Hempel considère quatre propriétés du concept de confirmation qui semblent particulièrement plausibles :

- (C1) *La Condition de Conséquence* : si *E* implique *H*, alors *E* confirme *H*[14]

- (C2) *La Condition de Cohérence* : si *E* confirme *H* et *H'*, alors *H* et *H'* ne sont pas logiquement incompatibles

- (C3) *La Condition de Conséquence Spéciale* : si *E* confirme *H* et si *H* implique *H'*, alors *E* confirme *H'*

- (C4) *La Condition de Conséquence Inverse* : si *E* confirme *H* et *H*' implique *H*, alors *E* confirme *H'*[15]

Si ces quatre propriétés sont plausibles, il est souhaitable qu'une théorie de la confirmation les satisfasse simultanément. Mais Hempel montre que ce n'est pas possible. En effet, (C1)-(C4) ne peuvent être simultanément acceptées sans que le concept de confirmation se voit trivialisé – n'importe quelle donnée *E* confirmerait n'importe quelle hypothèse *H*. La

[14] La Condition de Conséquence signifie, comme le dit Hempel, que l'implication logique est un cas particulier de confirmation.

[15] Soit *H'* la théorie newtonienne de la gravitation et *H* la première loi de Kepler selon laquelle la trajectoire des planètes du système solaire forme une ellipse dont le soleil est l'un des foyers. Supposons, en simplifiant, que *H'* implique *H*. La Condition de Conséquence Inverse dit dans ce cas que toutes les observations qui confirment la première loi de Kepler confirment également la théorie newtonienne de la gravitation.

preuve est très simple. Soit E une donnée quelconque et H une hypothèse quelconque. En vertu de la Condition de Conséquence, E se confirme lui-même. ($E \wedge H$) implique logiquement E donc en vertu de la Condition de Conséquence Inverse, E confirme ($E \wedge H$). Mais ($E \wedge H$) implique H donc en vertu de la Condition de Conséquence Spéciale, E confirme H. Par conséquent, une théorie satisfaisante de la confirmation doit rejeter au moins l'une des propriétés (C1)-(C4).

2.3 L'instantialisme hempelien

Les théories instantialistes de la confirmation (TIC) accordent une importance fondamentale au Critère de Nicod, c'est-à-dire à l'idée qu'un énoncé comme « Tous les *C* sont *N* » est confirmé par ses instances positives. La théorie de Hempel est une forme sophistiquée d'instantialisme qui introduit la notion originale de *développement*. On parle de *développement* d'une hypothèse H pour un ensemble fini d'individus[16] I. Le développement d'une hypothèse exprime ce que H affirmerait s'il n'existait que les individus de l'ensemble I. Par exemple, si $I = \{a, b\}$ et si $H = \forall x \mathrm{P}x$, alors le développement de H est ($\mathrm{P}a \wedge \mathrm{P}b$). De manière analogue, si $H' = \exists x \mathrm{P}x$, alors le développement de H' est ($\mathrm{P}a \vee \mathrm{P}b$). La caractérisation de la confirmation que Hempel propose est la suivante :

- E **H-confirme directement** H ssi E a pour conséquence logique le développement de H relativement aux individus mentionnés par E

- E **H-confirme** H ssi H est conséquence logique d'un ensemble d'énoncés dont chaque élément est H-confirmé directement par E

Quelques commentaires sur cette caractérisation s'imposent. (i) On remarquera tout d'abord que la théorie hempelienne a une portée beaucoup plus large que l'instantialisme rudimentaire qui porte sur les énoncés du type « Tous les *C* sont *N* ». (ii) Il faut ensuite bien saisir le contenu du concept de H-confirmation directe. Les données E circonscrivent en quelque sorte un domaine logique auquel l'hypothèse H est (provisoirement) restreinte - le développement de H pour l'ensemble des individus apparaissant dans E est précisément la restriction de H au domaine logique circonscrit par E. La H-confirmation directe n'exige pas seulement que la donnée E soit compatible avec H restreinte au domaine qu'elle circonscrit. Elle exige en outre que la donnée E *implique* la restriction de H au domaine qu'elle circonscrit, ce qui est beaucoup plus fort. Considérons un exemple élémentaire : si $H = \forall x \mathrm{P}x$, alors $E_1 = \mathrm{P}a$ H-confirme directement H puisque le domaine circonscrit par E est $\{a\}$. Or, le développement de H relativement à $\{a\}$ est $\mathrm{P}a$ qui est bien impliqué par (puisque identique à) la donnée E_1. En revanche, la donnée $E_2 = (\mathrm{P}a \wedge \mathrm{Q}b)$ où Q est un prédicat quelconque distinct de P ne H-confirme pas directement H puisque le domaine circonscrit par E_2 est $\{a,b\}$ et que E_2 *n'implique pas* le développement correspondant de H soit ($\mathrm{P}a \wedge \mathrm{P}b$). On peut justifier la H-confirmation directe en faisant valoir que tous les prédicats en jeu sont censés exprimer des propriétés observables ; par conséquent, en idéalisant quelque peu, si un observateur constate que b a une propriété (exprimée par) Q quelconque, il est en position de déterminer si b a la propriété (exprimée par) P. (iii) La notion de H-confirmation (par contraste avec la H-confirmation *directe*) permet d'étendre sensiblement la portée de la

[16] Il s'agit d'individus au sens logique, c'est-à-dire d'éléments du domaine de quantification.

théorie et d'en faire une version très libérale de l'instantialisme. Considérons en effet $E_3 = (Pa \wedge Pb)$ et de nouveau $H = \forall x Px$. E_3 H-confirme directement H, mais pas $H' = Pc$. H' est impliquée par H, donc H' est H-confirmée (indirectement) par E_3. Intuitivement, le fait que a et b aient la propriété (exprimée par) P nous donne confiance dans le fait que l'entité c que nous n'avons pas observée aura, elle aussi, la propriété (exprimée par) P. La notion de H-confirmation permet de rendre compte des formes de raisonnements inductifs comme l'inférence singulière que nous avons décrite dans la Section 1.

Qu'en est-il des deux paradoxes que nous avons présentés, le paradoxe des corbeaux et le paradoxe de Hempel ? Commençons par le second. On peut montrer que la théorie de Hempel satisfait les conditions (C1)-(C3). Comme manifestement elle n'est pas triviale (il existe, heureusement, certaines données qui ne H-confirment pas certaines hypothèses), il faut en conclure que cette théorie *ne satisfait pas* (C4) selon laquelle, si E confirme H et H' implique H, alors E confirme H'[17]. Passons maintenant au paradoxe des corbeaux et considérons de nouveau le cas problématique de la donnée suivante que l'ornithologue d'intérieur peut obtenir à peu de frais : « a est une chaussette blanche », que l'on exprimera par $E = (\neg Ca \wedge \neg Na)$. L'hypothèse est $H = \forall(Cx \rightarrow Nx)$. Le domaine circonscrit par E est $\{a\}$ donc le développement approprié de H est $(Ca \rightarrow Na)$. Or $(\neg Ca \wedge \neg Na)$ a pour conséquence logique $(Ca \rightarrow Na)$. Donc « a est une chaussette blanche » H-confirme (directement) « Tous les corbeaux sont noirs ». On aboutit donc à ce qui faisait figure de conséquence contre-intuitive de l'acceptation conjointe du Critère de Nicod et de la Condition d'Equivalence. Hempel a évidemment parfaitement conscience que sa théorie conduit à accepter également cette conséquence, et reconnaît sans difficultés que cette conséquence est contre-intuitive. Mais, selon lui, ce sont nos intuitions confirmationnelles qui nous égarent : nous sommes victimes d'une « illusion psychologique » qu'il s'agit de dissiper. Il y aurait en effet deux biais qui rendent contre-intuitif le pouvoir confirmationnel de « a est une chaussette blanche » relativement à « Tous les corbeaux sont noirs ». Le premier biais concerne l'interprétation des énoncés universels conditionnels du type « Tous les P sont Q ». D'après Hempel, nous aurions l'impression qu'un énoncé de ce type n'affirme quelque chose qu'à propos des entités qui sont des P, alors qu'en réalité il affirme quelque chose à propos de toutes les entités. Le second biais invoqué par Hempel est à la fois plus intéressant car il est propre à la confirmation, mais malheureusement exposé moins clairement que le premier : quand nous envisageons le pouvoir confirmationnel de $(\neg Ca \wedge \neg Na)$ relativement à $\forall(Cx \rightarrow Nx)$, nous avons tendance à lui substituer le pouvoir confirmationnel de $\neg Na$ à propos d'un objet a dont on sait déjà qu'il ne satisfait pas le prédicat C. Autrement dit, *on supposerait un certain arrière-plan de croyances* qui contient $\neg Ca$. Dans ce cas, savoir si a est noir ou pas n'a plus d'importance du point de vue confirmationnel. Par contraste, supposons que nos informations soient acquises en deux temps : nous apprenons d'abord que a est non-noir, ensuite que ce n'est pas un corbeau. Ce scénario semble bien confirmer l'hypothèse selon laquelle tous les corbeaux sont noirs. L'idée d'Hempel est intéressante, mais il faut souligner qu'il n'est pas sûr qu'elle soit compatible avec sa propre théorie de la confirmation (pour une analyse détaillée, voir Fitelson & Hawthorne, 2006).

2.4 Difficultés de la théorie hempelienne

La théorie hempelienne est l'une des manières les plus élégantes et les plus convaincantes de rendre justice à l'intuition instantialiste qui commande le Critère de Nicod. Elle fait néanmoins face à de sévères difficultés[18].

[17] Cela illustre parfaitement la différence entre une théorie instantialiste à la Hempel et les théories hypothético-déductives de la confirmation puisque (C4) est une conséquence immédiate de ce type de théories

[18] Nous suivons ici largement la remarquable discussion menée par J. Earman (Earman, 1992, chap.3).

La première de ces difficultés concerne les conditions (C1)-(C3) que Hempel considère comme de bonnes contraintes pour une théorie de la confirmation et que sa propre théorie de la confirmation valide. Carnap a en effet attiré l'attention sur le fait que Hempel semble mélanger deux concepts de confirmation (Carnap, 1962, § 87). Selon le *concept absolu* de confirmation, E confirme H si E donne de bonnes raisons de penser que H est vraie. Mais la théorie de Hempel n'est certainement pas une théorie du concept absolu de confirmation puisqu'une instance positive d'un énoncé $\forall x(\mathrm{C}x \rightarrow \mathrm{N}x)$ H-confirme (directement) cet énoncé. S'agirait-il alors du *concept incrémental* de confirmation ? Selon ce concept, E confirme H si E augmente notre confiance dans la vérité de H. Mais si l'on se laisse guider par ce concept, la Condition de Conséquence Spéciale (C3) ne semble pas totalement convaincante : le fait que E augmente notre confiance en H et que H implique H' n'implique pas que E augmente notre confiance en H'. Supposons par exemple que $E = \mathrm{P}a$, $H = (\mathrm{P}a \wedge \mathrm{Q}b)$ et $H' = \mathrm{Q}b$. Dans ce cas, apprendre E augmente bien notre confiance dans H, H' est bien impliqué par H, mais en toute généralité on ne voit pas pourquoi apprendre E augmenterait notre confiance en H'. La Condition de Conséquence Spéciale semble en revanche bien plus convaincante quand on s'attache au concept absolu de confirmation : si E donne de bonnes raisons de penser que H est vrai, alors il donne de bonnes raisons de penser qu'une conséquence H' de H est vraie.

Le second niveau de difficulté concerne les performances spécifiques de la théorie hempélienne. Il y a d'abord certaines conséquences contre-intuitives de la théorie. Par exemple (Earman, 1992), l'ensemble des observations Ra_ia_j pour $i = 1,2,\ldots,10^9$ et $j = 1,2,\ldots 10^9-1$ ne H-confirme pas l'hypothèse $\forall x\forall y Rxy$ puisqu'elles n'impliquent pas le développement de $\forall x\forall y Rxy$ pour les individus concernés (il « manque » $Ra_{10^9}a_{10^9}$). Pourtant, l'observation $\mathrm{R}a_1a_1$ H-confirme $\forall x\forall y Rxy$! Deuxièmement, le statut des termes théoriques (ou non-observationnels) est tout sauf clair. Hempel consacre une partie de son essai (Hempel, 1945, section 7) à une critique de la conception hypothético-déductive de la confirmation. Il attire alors l'attention, à juste titre, sur l'omniprésence des termes théoriques dans les hypothèses et théories de la science moderne. Mais on voit mal comment sa propre théorie est capable de rendre compte du pouvoir confirmationnel de données empiriques relativement à des hypothèses qui contiennent des termes théoriques. Dans l'exposition technique de sa théorie de la confirmation (Hempel, 1943), Hempel se facilite la tâche puisqu'il considère un langage qui ne contient que des prédicats exprimant des propriétés et relations observables (Hempel, 1943, p.126)[19].

La troisième difficulté concerne le type plus général de théorie de la confirmation auquel la théorie hempelienne appartient, à savoir une théorie purement syntaxique de la confirmation. Ce type de théorie doit en effet surmonter une difficulté découverte par N. Goodman, la célèbre « *nouvelle énigme de l'induction* » également nommée le « *paradoxe des émeraudes vreues* » (Goodman, 1946, 1955). Considérons les deux hypothèses suivantes :

H_1 : « Toutes les émeraudes sont vertes », $\forall x(\mathrm{E}x \rightarrow \mathrm{V}x)$

H_2 : « Toutes les émeraudes sont vreues », $\forall x(\mathrm{E}x \rightarrow \mathrm{VR}x)$

[19] Notons toutefois que l'une des principales tentatives contemporaines, la théorie du « *bootstrap* » de C. Glymour (1980), peut être conçue, de ce point de vue, comme une amélioration de la théorie hempélienne. Nous ne présenterons pas cette théorie sophistiquée et ingénieuse, qui a fait l'objet de nombreuses discussions. Voir notamment Christensen (1990).

Par définition, une chose est « vreue » ssi, (a) si elle a été observée avant t, alors elle est verte et (b) sinon, elle est bleue. Il suit de cette définition que si a a été observée avant t, alors elle est verte ssi elle est vreue. Supposons que a est observée avant t, et que a soit une émeraude verte. Alors on obtient comme donnée $E = (\mathrm{E}a \wedge \mathrm{V}a \wedge \mathrm{VR}a)$. Il en résulte que *E H-confirme (directement) à la fois H_1 et H_2* (voir Fitelson 2008 pour une reconstruction rigoureuse). Cette conclusion est manifestement contre-intuitive : on a du mal à se persuader que E confirme H_2. Par ailleurs, pour une émeraude observée en t ou après, les deux hypothèses font des prédictions incompatibles – cette émeraude sera verte d'après H_1 et bleue d'après H_2. Goodman conçoit sa « nouvelle énigme de l'induction » comme un argument contre les théories « syntaxiques » [20] de la confirmation, c'est-à-dire contre les théories de la confirmation qui construisent la relation de confirmation à partir de la *forme logique* des énoncés en jeu. La forme logique de H_1 *et* H_2 est en effet symétrique relativement à E. Mais leur comportement confirmationnel est, intuitivement, extrêmement différent. La conclusion qu'en tire Goodman est qu'une théorie de la confirmation qui repose sur la forme logique « manque » donc quelque chose d'essentiel du point de vue de l'objectif même qu'elle se fixe[21]. Goodman appelle « *projetables* » une hypothèse qui se laisse confirmer par ses instances positives, et sa thèse revient à affirmer que la forme logique d'une hypothèse ne permet pas d'en déterminer sa « projetabilité ».

2.5 Les théories hypothético-déductives de la confirmation (THDC)

Hempel prend soin de distinguer sa théorie de la confirmation des théories qu'il appelle « prédictionnistes », et qui correspondent à peu près à ce qu'on appelle en général les *théories hypothetico-déductives* de la confirmation (THDC). L'idée centrale de ces théories est la suivante : soit une hypothèse H et des croyances d'arrière-plan K[22]. Supposons que H et K impliquent (déductivement) une certaine conséquence observationnelle E. Dans ce cas E HD-confirme H (relativement aux croyances d'arrière-plan K) :

- E **HD-confirme** H relativement à K ssi $(H \wedge K)$ implique logiquement E

Considérons l'exemple suivant. Selon la Loi d'Ohm, la tension (U) d'un conducteur ohmique est égale au produit de sa résistance (R) et de l'intensité du courant qui le traverse (I) :

$$U = R.I$$

[20] Une remarque importante : on dit souvent que la nouvelle énigme de l'induction est un défi pour une théorie syntaxique de la confirmation ou du raisonnement inductif, et on pense souvent à Hempel ou Carnap. Mais il n'est pas certain que ce soit la dimension syntaxique, au sens où l'entendent les logiciens, qui soit réellement pertinente. Carnap (« On Inductive Logic ») affirme que pour lui tant la notion de conséquence déductive (l'objet de la logique déductive) que celle de degré de confirmation (l'objet de la logique inductive) sont des notions *sémantiques*. Le point principal, nous semble-t-il, est qu'il s'agit de théories fondées sur la *forme logique* des énoncés en jeu, et seulement sur leur forme logique, par opposition à leur contenu.

[21] « La confirmation d'une hypothèse par l'une de ses instances dépend assez largement de caractéristiques de l'hypothèse qui ne se ramènent pas à sa forme syntaxique. » (Goodman, 1955, pp. 71-2 ; notre traduction).

[22] L'exemple qui suit donnera une idée du rôle de ces croyances d'arrière-plan.

Supposons que l'on connaisse, pour un conducteur donné, sa résistance R et la tension à ses bornes U. On peut alors prédire une valeur pour l'intensité du courant qui le traverse. Si cette valeur est bien la valeur qui est mesurée, alors la Loi d'Ohm sera HD-confirmée par la mesure de l'intensité relativement à la donnée de la tension et de la résistance. Bien sûr, nous simplifions considérablement : les croyances d'arrière-plan sont beaucoup plus vastes que les *données* auxquelles nous les avons réduites. Elles contiennent également ce que l'on appelle des *hypothèses auxiliaires* comme par exemple l'hypothèse selon laquelle l'ampèremètre qui sert à mesurer l'intensité est fiable. La définition de la HD-confirmation que nous venons de donner est à certains égards contre-intuitive : supposons que K implique déjà logiquement la donnée E. Il est clair que dans ce cas E HD-confirmera H, résultat pour le moins étrange. On amende en général la définition originale de la manière suivante :

- E **HD-confirme** H relativement à K ssi (i) $(H \wedge K)$ implique logiquement E et (ii) K n'implique pas logiquement E

De manière générale, la relation de confirmation devient donc en quelque sorte la *Inverse* de la relation de conséquence logique. L'une des forces de la THDC est qu'elle semble rejoindre assez largement la pratique méthodologique des sciences empiriques : elle restitue l'idée que pour évaluer une théorie ou une hypothèse, on en extrait d'abord certaines « prédictions » et que lorsque ces prédictions sont correctes, la confiance dans la théorie ou l'hypothèse s'en trouve confortée. Voici comment Huygens distingue sa méthode de celle des Géomètres dans la Préface du *Traité de la Lumière* (1690) :

> « …au lieu que les Géomètres prouvent leurs propositions par des principes certains et inconstestables, ici les principes se vérifient par les conclusions qu'on en tire ; la nature de ces choses ne souffrant pas que cela se fasse autrement. Il est possible toutefois d'y arriver à un degré de vraisemblance, qui bien souvent ne cède guère à une évidence entière. »[23]

La THDC valide la condition (C4) ou *Condition de Conséquence Inverse* de Hempel : si E HD-confirme H et H' implique H, alors E HD-confirme H' puisque H' implique E par transitivité de la conséquence logique (pour simplifier nous laissons de côté les croyances d'arrière-plan). La relation de HD-confirmation est donc préservée par renforcement logique de l'hypothèse, alors que la réciproque n'est bien sûr pas vraie.

La conception hypothético-déductive de la confirmation, dans la formulation élémentaire que nous venons de proposer, rencontre un grand nombre de difficultés. Des raffinements sont régulièrement proposés depuis celui de Hempel (1945), mais on ne dispose d'aucune formulation stable aujourd'hui[24]. (i) La première de ces difficultés est le *problème de la conjonction non-pertinente*[25] : si E HD-confirme H, alors pour n'importe quelle

[23] *Op.cit*, p. 3 (nous avons modernisé quelque peu la langue). Nous empruntons cet exemple à Maher (2004). La suite de l'extrait est particulièrement intéressante également : « Savoir lorsque les choses, qu'on a démontrées par ces principes supposés, se rapportent parfaitement aux phénomènes que l'expérience a fait remarquer ; surtout quand il y en a grand nombre, et encore principalement quand on se forme et prévoit des phénomènes nouveaux, qui doivent suivre des hypothèses qu'on emploie, et qu'on trouve qu'en cela l'effet répond à notre attente. »

[24] Pour des tentatives récentes de raffinement de la THDC, voir Schurz (1991), Gemes (1998, 2005).

[25] Appelé également « problème de la confirmation sélective » par Gemes, 1998.

hypothèse *H'*, *E* confirme la conjonction de *H* et de *H'*. C'est une conséquence qui découle de la propriété de monotonie de la relation de conséquence logique. (ii) La seconde difficulté est duale de la première, il s'agit du problème de la disjonction non-pertinente : si *E* HD-confirme *H*, alors pour n'importe quelle autre donnée *E'*, la disjonction de *E* et de *E'* confirme *H*. (iii). La troisième difficulté est le *problème des hypothèses concurrentes* : bien souvent, lorsque *E* confirme une hypothèses *H*, *E* confirme également un très grand nombre d'autres hypothèses mutuellement incompatibles. C'est le cas par exemple quand *E* consiste en des observations de deux variables *x*, *y* et que l'hypothèse porte sur la relation qui existe entre les deux variables : pour n'importe quel ensemble fini de couples (*x*, *y*), il existe une infinité de fonctions capables d'engendrer les couples de l'ensemble. C'est une conséquence fâcheuse pour l'application de la THDC aux disciplines quantitatives.

Mentionnons enfin un dernier inconvénient de la THDC, *qui vaut également pour une théorie instantialiste à la Hempel* : elle n'est pas capable de traiter des hypothèses où figurent explicitement des probabilités « objectives » - propension, chance, fréquence relative. Considérons par exemple *H* = « il y a une chance sur deux pour qu'un noyau de radium 224 se désintègre au cours d'une période de 3.5 jours » et supposons pour simplifier que l'on puisse facilement déterminer pour un noyau de radium s'il s'est désintégré ou non pendant un certain intervalle de temps. *H* n'implique rien à propos d'un certain noyau *a* de radium 224 qui puisse être vérifié ou réfuté par l'observation de *a* pendant l'intervalle de temps approprié.

3 Le bayésianisme

Les théories de la confirmation que nous avons présentées pour le moment sont des théories exclusivement *qualitatives*. Il ne s'agit jamais de déterminer *le degré* de confirmation conféré par telle donnée à telle hypothèse, ce qui en fait manifestement des théories assez pauvres. Nous allons passer maintenant à la principale théorie de la confirmation que nous allons considérer dans ce chapitre, la théorie bayésienne de la confirmation (TBC) qui permet d'aborder la confirmation qualitativement *et* quantitativement. La théorie bayésienne de la confirmation est fondée sur l'épistémologie bayésienne ou *bayésianisme* dont nous allons maintenant exposer les principales idées. Cela nous permettre notamment d'introduire un outil fondamental de la TBC qui était absent des théories que nous avons abordées : le calcul des probabilités.

3.1 Degrés de croyances et calcul des probabilités

Le bayésianisme a une histoire complexe et se décline sous bien des variantes ; mais on peut ramener à trois thèses le cœur de la l'épistémologie bayésienne contemporaine :

- (B1) *le gradualisme* : une épistémologie adéquate doit considérer les *degrés de croyance* et non pas seulement les croyances « pleines ». L'attitude épistémique des agents vis-à-vis de propositions est affaire de degrés qui reflètent la confiance qu'ils ont à l'égard de la vérité des propositions.

- (B2) *le probabilisme* : les degrés de croyance d'un agent rationnel se laissent représenter par une distribution de probabilités.

- (B3) *la révision par conditionnalisation* : les croyances d'un agent rationnel sont révisées par conditionnalisation.

Nous allons dans la suite de cette section commenter (B1) et (B2), nous consacrerons la section suivante à (B3). Les théories instantialistes ou hypothetico-déductives, sous leur forme usuelle, ne laissent pas de place aux degrés de confirmation. Il est tout au plus question de la confiance que la vérité d'une donnée peut nous donner en la vérité d'une hypothèse. L'hypothèse fondamentale de l'épistémologie bayésienne, le gradualisme (B1), est qu'il faut prendre en compte et expliciter toute la palette des degrés de croyance que nous pouvons entretenir à propos des données, hypothèses, théories, etc. C'est manifestement une hypothèse très vague qu'il faut qualifier. La thèse probabiliste (B2) affirme précisément que les degrés de croyance d'un agent rationnel se conforment aux axiomes du calcul des probabilités. Supposons que les croyances d'un agent portent sur un ensemble d'énoncés et que l'on indique génériquement par $P(H)$ le degré de croyance que l'agent entretient vis-à-vis de l'énoncé H. Alors la thèse fondamentale affirme que P constitue une distribution de probabilités, c'est-à-dire que P obéit aux axiomes suivants :

- (A1) $P(H) \geq 0$ pour tout H

- (A2) $P(H) = 1$ si H est une vérité logique

- (A3) $P(H_1 \vee H_2) = P(H_1) + P(H_2)$ si H_1 et H_2 sont logiquement incompatibles

L'axiome (A1) exprime le fait que le degré de croyance minimal est représenté par 0 tandis que l'axiome (A2) exprime le fait que le degré de croyance maximal est représenté par 1. Tout énoncé se voit attribuer un degré de croyance compris entre 0 et 1. L'axiome (A3) est l'axiome central, l'axiome d'*additivité*[26]. Un certain nombre de propriétés découlent très directement de ces axiomes :

- $P(H) = 1 - P(\neg H)$

- $P(H) = 0$ si H est une contradiction logique

- Si H_1 et H_2 sont logiquement équivalents, alors $P(H_1) = P(H_2)$

- $P(H_1) = P(H_1 \wedge H_2) + P(H_1 \wedge \neg H_2)$

[26] Nous fournissons ici une présentation simplifiée et adaptée à la littérature philosophique du calcul des probabilités. La théorie mathématique contemporaine des probabilités attribue les probabilités à des ensembles (des « événements ») et non pas à des énoncés. En outre elle suppose en général non seulement l'additivité finie (comme nous l'avons fait) mais l'additivité dénombrable. Le lecteur philosophe qui voudra s'initier aux probabilités pourra consulter Skyrms (1966) ou Hacking (2001).

Tous les bayésiens s'accordent sur la thèse selon laquelle si un agent est rationnel, alors ses degrés de croyance obéissent à (A1)-(A3). Le bayésianisme radical ajoute la réciproque : si les degrés de croyance d'un agent obéissent à (A1)-(A3), alors cet agent est rationnel. Autrement dit, pour ce qui est des croyances, la rationalité n'impose aucune autre norme que celles exprimées par les axiomes du calcul des probabilités. Pour le bayésianisme radical, un agent n'est en particulier pas tenu d'aligner ses degrés de croyance sur les probabilités objectives (fréquence relative, chance, propension) - si de telles choses existent - dont il pourrait être informé.

3.2 La conditionnalisation et le théorème de Bayes

On peut considérer (B2) comme la thèse bayésienne statique ou synchronique. Par contraste, la thèse (B3) est une thèse dynamique ou diachronique qui porte sur le changement des degrés de croyance. (B3) affirme en effet qu'un agent rationnel doit réviser ses degrés de croyance par conditionnalisation : quand il apprend que E est le cas, son degré de croyance en H passe de la probabilité initiale (ou *a priori*) – celle qu'il a *avant* de prendre en compte une information - $P(H)$ à la probabilité *a posteriori* $P(H \mid E)$ qui est définie ainsi :

- $P(H \mid E) =_{def} P(E \wedge H) / P(E)$ où $P(E) > 0$

Nous laissons le lecteur vérifier que $P(. \mid E)$ satisfait (A1)-(A3) donc est une distribution de probabilités. Il faut insister sur deux caractéristiques de la conditionnalisation qui ont suscité de nombreuses discussions. La première caractéristique tient dans le fait que la conditionnalisation est *partielle* : si la donnée E a une probabilité initiale nulle, alors la conditionnalisation ne contraint pas la nouvelle distribution de probabilités. La seconde est qu'elle s'applique aux données que l'on considère comme certaines. Des objections philosophiques peuvent s'élever à cet endroit : est-on jamais absolument certain de la vérité d'une donnée ? L'une des principales figures du bayésianisme contemporain, R. Jeffrey, a proposé une généralisation de la conditionnalisation, la « *règle de Jeffrey* », qui permet de réviser les croyances partielles de l'agent à partir de données auxquelles on accorde une probabilité quelconque (pas nécessairement maximale)[27]. Dans ce qui suit, nous nous en tiendrons cependant à l'idéalisation usuelle qui consiste à faire comme si les données sur la base desquelles on révise ses croyances étaient certaines.

Le *théorème de Bayes*[28] est une conséquence immédiate de la définition de la probabilité conditionnelle ; il s'énonce comme suit :

- (TB1) $P(H \mid E) = [P(E \mid H).P(H)]/P(E)$ où $P(H)$, $P(E) > 0$

Dans le contexte de la théorie de la confirmation, le théorème de Bayes nous indique comment déterminer la probabilité d'une hypothèse H compte tenu d'une donnée E à partir des probabilités initiales de E et de H et de la probabilité de E étant donné (la vérité de) l'hypothèse H $P(E \mid H)$. On parle parfois de la *vraisemblance* (*likelihood*) de H pour désigner $P(E \mid H)$. Il s'agit du degré auquel l'hypothèse H *prédit* la donnée E. Il est aisé de

[27] Supposons qu'une observation fasse passer la probabilité que E soit vraie de $P(E)$ (la probabilité initiale) à $P^*(E)$. Comment déterminer de manière générale la nouvelle distribution de probabilité $P^*(.)$? La règle de Jeffrey dit que pour toute proposition H, $P^*(H) = P(H \mid E).P^*(E) + P(H \mid \neg E).P^*(\neg E)$. Il est aisé de vérifier que dans le cas-limite où $P^*(E)=1$, la règle de Jeffrey se ramène à la conditionnalisation usuelle.

[28] Voir Hacking (2001), chap. 15 et Joyce (2003). Hacking (2001) contient de nombreux exemples commentés et exercices.

voir que si H implique logiquement E, alors $P(E | H)$ est maximale ; si H implique $\neg E$, alors $P(E | H)$ est nulle. Il est parfois délicat de supposer connue $P(E)$ mais l'on peut s'en passer si l'on connaît $P(E | \neg H)$[29] :

- (TB2) $P(H | E) = [P(E | H).P(H)] / [P(E | H).P(H) + P(E | \neg H).P(\neg H)]$ où $P(H), P(E) > 0$

Cette seconde forme du théorème de Bayes peut être généralisée au cas où l'on considère n hypothèses exhaustives et mutuellement exclusives $H_1, \ldots, H_n$. Dans ce cas, pour tout H_i ($0 \leq i \leq n$),

- (TB3) $P(H_i | E) = [P(E | H_i).P(H_i)] / \Sigma_j [P(E | H_j).P(H_j)]$ où $P(H_j), P(E) > 0$

3.3 Les justifications du bayésianisme

Pourquoi les degrés de croyance d'un agent rationnel devraient-ils obéir au calcul des probabilités (B2) ? A peu près à la même époque, mais indépendamment l'un de l'autre, De Finetti (1937) et Ramsey (1926) ont construit un argument nommé le *Pari Hollandais* (*Dutch Book*) dont l'objectif est de montrer qu'un agent qui parie sur la base de ses degrés de croyance et dont les degrés de croyance violent le calcul des probabilités peut se voir proposer une série de paris qu'il accepterait alors même qu'ils le mèneraient assurément à une perte monétaire. En d'autres termes, la violation du calcul des probabilités rend un agent *vulnérable* à un Pari Hollandais. On peut démontrer que les degrés de croyance d'un agent obéissent au calcul des probabilités si et seulement s'il est invulnérable à un Pari Hollandais.

Supposons par exemple (i) que Paul croit au degré 0.4 que H est vrai et au degré 0.7 que H n'est pas vrai (ce qui viole le calcul des probabilités); et (ii) que ses croyances sont reflétées dans ses *coefficients de pari*. Cela signifie que Paul est prêt à payer 0.4 m euros pour un pari qui rapporte m euros si H est le cas, et 0 euro sinon. Marie peut alors proposer deux paris à Paul qui lui vaudront une perte certaine : posons par exemple $m = 10$ euros et supposons que Marie propose

- le Pari n°1 sur H (pour 0.4×10 euros), et

- le Pari n°2 sur $\neg H$ (pour 0.7×10 euros).

Si H est le cas, alors Paul obtiendra $10 - (0.4 \times 10 + 0.7 \times 10) = -1$ euro. Si H n'est pas le cas, alors Paul perdra également un euro. Autrement dit, Pierre est perdant dans tous les cas. Un argument similaire mais de nature dynamique, le Paris Hollandais Diachronique, a été proposé par David Lewis pour justifier le recours à la règle de conditionnalisation (Teller

[29] Dans certains contextes (typiquement en statistique médicale), $P(E | H)$ et $P(E | \neg H)$ sont appelés respectivement le *taux de vrais positifs* (ou la sensibilité) et le *taux de faux positifs* : si E est le verdict positif d'un test censé établir la vérité de H (par exemple le verdict positif d'un test de grossesse), $P(H | E)$ est la probabilité pour qu'un verdict soit positif quand H est vraie tandis que $P(E | \neg H)$ est la probabilité pour qu'un verdict soit positif quand H est fausse.

1973, Lewis 1999, chap.23 « Why Conditionalize ? »). L'argument du Pari Hollandais appartient à une famille plus large d'*arguments pragmatiques* en faveur du probabilisme : des arguments qui prétendent montrer que la violation des probabilités engendre de l'irrationalité dans l'action (ou dans la disposition à l'action)[30]. La question de la justification du bayésianisme est une question largement débattue. Certains, en particulier, considèrent que les justifications pragmatiques réduisent les croyances à leur rôle dans l'action et négligent leur dimension épistémologique. C'est ce qui motive la tentative récente de Joyce (1998) de fournir un argument purement épistémique (non pragmatique) en faveur du probabilisme (B2). Joyce caractérise axiomatiquement un ensemble de conditions sur des mesures possibles de *précision* des degrés de croyance et montre que pour toutes les mesures ainsi définies, si les degrés de croyance d'un agent ne satisfont pas le calcul des probabilités, alors il existe des degrés de croyance qui sont strictement plus précis. Par ailleurs, parmi les défenseurs de la théorie bayésienne de la confirmation, vers laquelle nous nous tournons désormais, si certains accordent une certaine importance à cette question de la justification (Howson & Urbach, 1989), d'autres la laissent largement de côté pour se concentrer sur la capacité de la théorie à rendre compte de la pratique scientifique de la confirmation (Strevens (2006), qui se réclame Horwich (1982) et Earman (1992)).

4 La théorie bayésienne de la confirmation (TBC)

4.1 Les différentes notions de confirmation de la TBC

Comme le note Carnap dans la préface à la seconde édition des *Logical Foundations of Probability* (1962), dans un cadre probabiliste il faut distinguer deux notions de confirmation[31] : un concept absolu (« *confirmation as firmness* ») et un concept incrémental (« *confirmation as increase in firmness* ») de confirmation. Il y a confirmation en un sens absolu si la probabilité de H étant donné E est assez forte : $\Pr(H \mid E) > k$. La théorie bayésienne de la confirmation n'adopte pas un tel concept absolu de confirmation, elle adopte le concept incrémental. Il y a confirmation en un sens incrémental si la probabilité de H étant donné E est supérieure à la probabilité initiale de H :

- E **B-confirme** H ssi $P(H \mid E) > P(H)$.
- E **B-infirme** H ssi $P(H \mid E) < P(H)$
- E est **non-pertinent** du point de vue confirmationnel pour H ssi $P(H \mid E) = P(H)$

Autrement dit : E B-confirme H du point de vue d'un certain agent ssi apprendre E augmente la confiance de cet agent en H. Certaines situations épistémiques permettent de comprendre la préférence pour le concept incrémental : supposons par exemple que, pour Paul, apprendre E ferait baisser la probabilité de H : $\Pr(H \mid E) < \Pr(H)$. Et supposons que, malgré cela, la probabilité de H reste au-dessus du seuil k : $\Pr(H \mid E) > k$. Dans ce cas, on aurait confirmation absolue mais pas confirmation incrémentale. Nos jugements confirmationnels spontanés nous font certainement préférer le verdict du concept

[30] L'autre grand type d'arguments pragmatiques se base sur une axiomatique des *préférences*. Les résultats fondamentaux sont dans ce cas fournis par la théorie contemporaine de la décision, comme dans Savage (1954).

[31] Il s'agit bien sûr d'un raffinement de la distinction que nous avons introduite plus haut en discutant la théorie instantialiste de Hempel.

incrémental : nous n'avons pas envie de dire que E, qui rend H moins probable qu'elle ne l'était, confirme H. Deux remarques importantes doivent être formulées à propos de la notion de B-confirmation que nous venons d'introduire. Tout d'abord, les bayésiens insistent souvent sur les croyances d'arrière-plan de l'agent épistémique, que l'on note K. Ils utilisent donc une notion plus fine de confirmation selon laquelle E B*-confirme H *relativement à* K ssi $P(H \mid E \wedge K) > P(H \mid K)$. Par souci de simplicité, nous utiliserons la B-confirmation tant qu'elle suffit à l'analyse. Ensuite, il faut souligner le fait que si la TBC repose sur une théorie quantitative des degrés de croyance, *la notion de B-confirmation est un concept qualitatif de confirmation*. La B-confirmation est muette sur la « quantité » de confirmation qu'une donnée E confère à une hypothèse H. L'une des forces principales de la TBC est qu'elle permet de construire une notion *quantitative* ou une *mesure de confirmation*. On note génériquement une telle mesure $c(H,E)$. Une proposition naturelle consiste à prendre la *différence* entre la probabilité initiale de H et sa probabilité conditionnelle à E :

- $d(H,E) = P(H \mid E) - P(H)$

La mesure d est positive (resp. négative) si E B-confirme (resp. B-infirme) H. Il existe dans la littérature des propositions concurrentes (voir Fitelson 2001) sur lesquelles nous reviendrons ultérieurement.

4.2 Quelques analyses bayésiennes

4.2.1 Le théorème de Bayes et les théories hypothetico-déductives

La popularité de la TBC provient de sa capacité à rendre compte d'un grand nombre d'intuitions confirmationnelles. Rappelons le théorème de Bayes qui découle de la définition de la probabilité conditionnelle : $P(H \mid E) = [P(E \mid H).P(H)]/P(E)$ où $P(H)$, $P(E) > 0$ (TB1). Du théorème de Bayes et de la TBC il découle immédiatement que :

(1) toutes choses égales par ailleurs[32], plus une donnée E est probable étant donné une hypothèse H[33], plus H sera confirmée par E[34].

(2) toutes choses égales par ailleurs, *moins* E est probable *a priori*, plus l'hypothèse H sera confirmée par E. (« principe de surprise », Joyce)

(3) E confirme H si et seulement si $P(E \mid H) > P(E \mid \neg H)$

[32] La clause « toutes choses égales par ailleurs » est indispensable : si les deux autres variables ($P(H)$ et $P(E)$) ne sont pas fixées, l'affirmation peut être fausse. Cela illustre le fait que, selon cette analyse, la confirmation dépend de trois facteurs interdépendants.

[33] Plus H « prédit » E, pour employer la terminologie de la section précédente.

[34] Nous nous basons sur la mesure d.

La propriété (1) reflète certaines relations attendues entre conséquence logique et confirmation. (a) Si E est logiquement incompatible avec H, alors $P(E \mid H) = 0$ et par conséquent la B-infirmation de H par E est maximale. (b) Si E est conséquence logique de H, alors $P(E \mid H) = 1$ et la confirmation de H par E est, toutes choses égales par ailleurs, maximale. La TBC est donc capable de retenir ce qui semble intuitif dans les théories hypothetico-déductives de la confirmation : si E est conséquence logique d'une hypothèse H, alors H est confirmée par E. C'est en effet une propriété élémentaire du calcul des probabilités que si H a pour conséquence logique E, alors $P(H \wedge E) = P(H)$. Par conséquent, dans ce cas,$P(H \mid E) = P(H) / P(E) > P(H)$ si $0<P(H), P(E) <1$[35]. Autrement dit, si E et H ne sont initialement ni certainement vraies, ni certainement fausses, alors H reçoit nécessairement une confirmation du fait que E est le cas. La TBC permet donc de justifier une intuition fondamentale de la THDC et de rendre compte d'une part importante de la pratique scientifique.

Mais la TBC permet également de surmonter certaines difficultés que rencontre la THDC. L'une de ces difficultés, on l'a vu, est le problème de la conjonction non-pertinente : si E HD-confirme H, alors nécessairement E HD-confirme $(H \wedge H')$. La TBC hérite partiellement du problème de la conjonction non-pertinente dans le cas particulier où H implique logiquement E : si $0<P(H), P(H'), P(E) <1$, alors E B-confirme H mais également $(H \wedge H')$. Il faut cependant souligner que cette propriété ne vaut pas en toute généralité (comme c'est le cas avec la HD-confirmation) : il n'est pas vrai que si E B-confirme H, alors pour toute H', E B-confirme $(H \wedge H')$. Contrairement à la conséquence logique, la notion de dépendance probabiliste n'est en effet pas monotone. Quand H implique logiquement E, l'analyse quantitative dont est capable la TBC s'avère par ailleurs fructueuse : si l'on utilise la différence comme mesure du degré de confirmation, alors le degré de confirmation que E confère à H est supérieur à celui qu'il confère à $(H \wedge H')$ (Earman, 1992, 63-5)[36].

La propriété (2) affirme que des données surprenantes ont, toutes choses égales par ailleurs, un fort pouvoir confirmationnel. La restriction est importante : une donnée improbable ne confirme pas nécessairement une hypothèse. Mais si deux données E' et E sont prédites au même degré par l'hypothèse H, alors H reçoit plus de support confirmationnel de la donnée qui a la plus faible probabilité initiale. Les bayésiens voient dans cette propriété une vertu de la TBC[37]. Considérons l'exemple suivant : même si la scarlatine s'accompagne invariablement (supposons-le) d'une forte fièvre et d'une éruption cutanée, l'éruption cutanée de Paul est une meilleure donnée en faveur de l'hypothèse que Paul a la scarlatine car c'est un symptôme bien plus rare qu'une forte fièvre. Du point de vue conceptuel, il est important de noter que, à la différence de la propriété (1), la propriété (2) est propre à la théorie bayésienne et étrangère à la théorie hypothético-déductive. La propriété (3), enfin affirme que E confirme H exactement quand H prédit « plus » E que ne le fait sa négation. Pour le dire autrement, il y aurait plus de chance que E soit vrai si le monde obéissait à l'hypothèse H que s'il n'y obéissait pas.

[35] Nous reviendrons ultérieurement sur le cas où $P(E)=1$.

[36] Voir Fitelson (2002) pour une discussion critique récente du traitement bayésien usuel du problème de la conjonction non-pertinente.

[37] C'est le cas de Howson & Urbach (1989, 86-8).

4.2.2 Le paradoxe des corbeaux[38]

Comment la théorie bayésienne traite-t-elle le paradoxe des corbeaux ? Rappelons qu'une instance positive d'un énoncé comme « Tous les corbeaux sont noirs » $\forall(Cx \rightarrow Nx)$ est de la forme : « *a* est un corbeau et *a* est noir » $(Ca \wedge Na)$. Le Critère de Nicod affirme qu'une instance positive confirme l'énoncé universel associé. Si l'on accepte la Condition d'Equivalence, cela implique que l'instance positive $(\neg Ca \wedge \neg Na)$ de « Tous les non-noirs sont non-corbeaux » confirme « Tous les corbeaux sont noirs ». La Condition d'Equivalence étant automatiquement satisfaite par la TBC, les questions cruciales à laquelle elle doit répondre sont les suivantes :

(Q1) La TBC valide-t-elle le Critère de Nicod ?

(Q2) Y a-t-il des situations où une donnée comme $(\neg Ca \wedge \neg Na)$ B-confirme $\forall(Cx \rightarrow Nx)$?

(Q3) Y a-t-il des différences dans les degrés de confirmation respectifs de $(Ca \wedge Na)$ et $(\neg Ca \wedge \neg Na)$ relativement à $\forall(Cx \rightarrow Nx)$?

Il y a de nombreuses réponses bayésiennes au paradoxe des corbeaux. Concernant la validité du Critère de Nicod (Q1), certains bayésiens ont montré que la réponse était, dans le cas général, négative. Et il existe des situations où la non-validité du Critère de Nicod est, à y bien regarder, parfaitement intuitive. Considérons l'énoncé

« Tous les renards sont situés hors de Paris »

et supposons qu'on ait observé un renard près de la Porte d'Orléans, mais à l'extérieur de Paris. Il s'agit d'une instance positive de l'énoncé initial, mais est-elle à même d'apporter quelque confirmation à l'énoncé « Tous les renards sont situés hors de Paris » ? Dans des conditions normales, il ne semble pas que cela soit le cas : les renards peuvent se déplacer, et si l'on en a aperçu un tout près de Paris, il se peut bien qu'il y en ait dans Paris. L'instance positive de l'énoncé semble donc *diminuer* notre confiance dans la vérité de l'énoncé[39]. Hempel (1967) conteste la capacité de ce genre de contre-exemple à véritablement remettre en

[38] Horwich (1982, pp. 54 et *sq*.), Earman (1992, pp. 69-73), Vranas (2004), Fitelson & Hawthorne (2006), Fitelson (2006).

[39] Un autre contre-exemple est dû à Good (1967) : supposons que nous sachions que notre monde se laisse décrire par l'une des deux hypothèses suivantes. Selon la première, il y a 100 corbeaux noirs, aucun corbeau non-noir et 1 million d'autres oiseaux ; selon la seconde, il y a 1000 corbeaux noirs, 1 corbeau blanc et 1 million d'autres oiseaux. Une instance positive de « Tous les corbeaux sont noirs » pourrait augmenter notre confiance en la seconde hypothèse et donc B-infirmer l'énoncé universel « Tous les corbeaux sont noirs ».

question le Critère de Nicod. Son objection est qu'un tel contre-exemple repose sur un ensemble de croyances d'arrière-plan (ici, sur la géographie parisienne, les renards, etc.) alors que le Critère de Nicod doit plutôt s'entendre de la manière suivante : si l'on se base sur la donnée $E = (Ca \wedge \mathrm{N}a)$ et que l'on ne suppose rien d'autre, alors la donnée confirme nécessairement l'énoncé général $\forall(\mathrm{C}x \rightarrow \mathrm{N}x)$. L'idée de ne rien supposer d'autre que E, ou, de manière équivalente, de supposer un ensemble de croyances d'arrière-plan K dégénéré qui ne contiendrait rien d'autre que les vérités logico-mathématiques, est problématique du point de vue du bayésianisme car elle revient à interdire des différences subjectives qui sont parfaitement permises par la TBC. En revanche, pour l'approche logique des probabilités, à la Carnap (1950/1962, voir ci-dessous), c'est une idée naturelle. Or, dans un cadre néo-carnapien, P. Maher (2004, section 8) a récemment montré que *le Critère de Nicod n'était pas non plus valide* pour le concept incrémental de confirmation de la TBC.

Considérons maintenant la conclusion contre-intuitive du paradoxe des corbeaux : une donnée comme $(\neg\, \mathrm{C}a \wedge \neg\, \mathrm{N}a)$ peut-elle confirmer $\forall(\mathrm{C}x \rightarrow \mathrm{N}x)$ (Q2) ? La TBC répond positivement à cette question. Mais elle permet de rendre compte de l'idée intuitive selon laquelle une instance positive $(Ca \wedge \mathrm{N}a)$ confirme *plus* l'énoncé $\forall(\mathrm{C}x \rightarrow \mathrm{N}x)$ que ne le fait la donnée $(\neg\, \mathrm{C}a \wedge \neg\, \mathrm{N}a)$ et que la donnée $(\neg\, \mathrm{C}a \wedge \neg\, \mathrm{N}a)$ confirme très faiblement l'énoncé $\forall(\mathrm{C}x \rightarrow \mathrm{N}x)$ (voir Vranas (2004) et Fitelson (2006)) (Q3). Les hypothèses qui suffisent à montrer cela sont les suivantes : la probabilité que a soit un corbeau est très faible comparée à la probabilité que a soit noir, et la probabilité que a soit un corbeau ou que a soit non-noir est indépendante de celle de $\forall(\mathrm{C}x \rightarrow \mathrm{N}x)$. Ces hypothèses permettent de montrer que

- $\mathrm{P}(\forall(\mathrm{C}x \rightarrow \mathrm{N}x) \mid (\neg\, \mathrm{C}a \wedge \neg\, \mathrm{N}a)) > \mathrm{P}(\forall(\mathrm{C}x \rightarrow \mathrm{N}x))$

  [*i.e.* $(\neg\, \mathrm{C}a \wedge \neg\, \mathrm{N}a)$ B-confirme $\forall(\mathrm{C}x \rightarrow \mathrm{N}x)$]

- $c((\neg\, \mathrm{C}a \wedge \neg\, \mathrm{N}a), \forall(\mathrm{C}x \rightarrow \mathrm{N}x)) = \varepsilon$ pour un “petit” ε

  [*i.e.* $(\neg\, \mathrm{C}a \wedge \neg\, \mathrm{N}a)$ apporte une faible confirmation à $\forall(\mathrm{C}x \rightarrow \mathrm{N}x)$]

- $\mathrm{P}(\forall(\mathrm{C}x \rightarrow \mathrm{N}x) \mid (\mathrm{C}a \wedge \mathrm{N}a)) > \mathrm{P}(\forall(\mathrm{C}x \rightarrow \mathrm{N}x) \mid (\neg\, \mathrm{C}a \wedge \neg\, \mathrm{N}a))$

  [*i.e.* $(\mathrm{C}a \wedge \mathrm{N}a)$ confirme plus $\forall(\mathrm{C}x \rightarrow \mathrm{N}x)$ que ne le fait $(\neg\, \mathrm{C}a \wedge \neg\, \mathrm{N}a)$]

Même si les hypothèses d'indépendance font débat (voir Vranas, 2004), on voit l'avantage que la TBC tire de la richesse du cadre probabiliste qui permet de distinguer le pouvoir confirmationnel de $(Ca \wedge \mathrm{N}a)$ et celui de $(\neg\, Ca \wedge \neg\, \mathrm{N}a)$.

4.2.3 Le problème de Duhem-Quine

Dorling (1979) et Howson & Urbach (1989) proposent une analyse bayésienne du problème de Duhem-Quine[40]. Rappelons que le problème est le suivant : supposons que l'on soit capable d'extraire certaines conséquences empiriques d'une hypothèse *H*. Dans le cas général, la seule hypothèse *H* ne suffit pas à impliquer de telles conséquences : il faut lui adjoindre des *hypothèses auxiliaires*, disons *A*. Supposons maintenant que les données empiriques contredisent ces conséquences. On peut par exemple supposer que $(H \wedge A)$ implique $\neg E$ et que *E* soit le cas. Du point de vue déductif, cela signifie que la conjonction $(H \wedge A)$ est réfutée. Le problème qui se pose est celui de savoir comment désigner les propositions coupables (ou les plus coupables) dans cette conjonction :

> « la seule chose que nous apprenne l'expérience, c'est que, parmi toutes les propositions qui ont servi à prévoir ce phénomène et à constater qu'il ne se produisait pas, il y a au moins une erreur ; mais où gît cette erreur, c'est ce qu'elle ne nous dit pas. » (Duhem 1906, Partie II, Chap. VI, § II).

On réagit en général de manière *sélective* à des données empiriques qui réfutent un corps de propositions : certaines propositions sont plus « infirmées » que d'autres. (Notons que toutes les propositions ne sont pas infirmées. Il se peut, dans certains cas particuliers, que des propositions voient leur probabilité augmenter). La TBC a les moyens de *décrire* une telle sélectivité. Howson & Urbach (1989) donnent un exemple en provenance de la chimie en considérant l'hypothèse *H* selon laquelle le poids d'un atome quelconque est un multiple entier du poids de l'atome d'hydrogène (Prout, 1815). Les hypothèses auxiliaires *A* consistent essentiellement à supposer que les instruments de mesure sont fiables. Il se trouve que les résultats des mesures prises à l'époque divergeaient significativement de ce qu'ils auraient dû être selon *H*. L'analyse de l'exemple est censée montrer que, même si les chimistes avaient initialement une confiance forte (disons 0.9) en l'hypothèse *H* et assez forte en *A* la fiabilité de leurs instruments (disons 0.6), ils pouvaient parfaitement être justifiés à réviser leurs croyances, après avoir connaissances des résultats des mesures, d'une manière telle que (i) leur confiance dans l'hypothèse centrale restait très forte (0.878) alors que (ii) leur confiance en la fiabilité des instruments de mesure s'effondrait (0.073). Il s'agit donc d'un cas de B-infirmation « légère » de *H* et de B-infirmation massive de *A*. Un exemple dont les conséquences sont analogues avait déjà été proposé par Dorling (1979). De manière générale, si *H* est l'hypothèse examinée, *A* l'ensemble des hypothèses auxiliaires et si *H* et *A* impliquent la négation de *E*, alors la TBC laisse ouvertes de nombreuses possibilités confirmationnelles. Il se peut en effet que

- *H* soit B-infirmée par *E* mais *A* soit B-confirmée (et inversement)
- *H* et *A* soient B-infirmées[41]
- ni *H* ni *A* ne soient B-infirmées[42]

---

[40] Voir Earman (1992, pp. 83 et *sq.*)

[41] Hajek & Joyce (2008).

[42] Voir Salmon (1973) cité par Earman (1992, p.83).

Nous avons donné quelques exemples d'analyse épistémologique bayésienne. Précisons, pour conclure, qu'il en existe de nombreux autres : les bayésiens ont proposé des recontructions de la notion d'*hypothèse ad hoc*, de l'idée que la variété des données empiriques a un fort pouvoir confirmationnel[43], etc.

4.3 Les difficultés de la TBC

Après avoir donné une idée des accomplissements de la théorie bayésienne de la confirmation, accomplissements qui expliquent largement pourquoi elle est, de loin, la théorie la plus répandue aujourd'hui, nous allons maintenant passer aux difficultés qu'elle rencontre. Nous allons examiner deux difficultés célèbres : l'objection de Popper-Miller et le problème des données connues[44].

4.3.1 L'objection de Popper-Miller

Dans un article de 1983, K. Popper et D. Miller élaborent un argument dont l'ambition est d'établir l'impossibilité d'une logique inductive. Cet argument s'attaque directement à la notion d'incrément de probabilité qui est au cœur de la TBC. Supposons que H implique E ; on sait que dans ce cas E B-confirme H ssi $P(H) > 0$ et $P(E) < 1$. Supposons en outre que $P(H \mid E)$ et $P(E) \neq 1$. On montre alors que $P(H \vee \neg E) > P(H \vee \neg E \mid E)$. Autrement dit, la disjonction $(H \vee \neg E)$ est B-infirmée par E.

En quoi ce résultat est-il problématique ? Popper et Miller remarquent que H est logiquement équivalente à l'énoncé $(H \vee E) \wedge (H \vee \neg E)$. Le premier membre de la conjonction suit logiquement de E tandis que le second représenterait, par conséquent, *le contenu de H qui excède E*. Selon cette interprétation, le contenu de H qui excède E est nécessairement B-infirmé par E – quand bien même H serait B-confirmée par E. Popper et Miller en infèrent que l'idée que l'incrément probabiliste $(P(H \mid E) - P(H))$ représente le support inductif conféré à H par E est illusoire et en concluent que « tout support probabiliste est purement déductif ». Une autre façon de présenter cette conclusion consiste à partir du fait que, en vertu des hypothèses initiales,

$$d(H, E) = \mathrm{d}((H \vee \neg E), E) + d(H \vee E, E)$$

Autrement dit, la quantité de support confirmationnel conférée à H par E se laisse additivement décomposer en (i) celle que E confère à $(H \vee \neg E)$ et (ii) celle que E confère à $(H \vee E)$. Le résultat de Popper et Miller implique (sous l'hypothèse que $P(H \mid E)$ et $P(E) \neq 1$) que $\mathrm{d}((H \vee \neg E), E)$ est strictement négatif. Gillies (1986) reformule l'argument en le faisant reposer sur cette décomposition additive[45] : E n'apporterait aucun support inductif au contenu de H qui l'excède.

Les conclusions de l'argument paraissent dévastatrices pour la théorie bayésienne de la confirmation, et en général pour toute théorie probabiliste de la confirmation fondée sur le

[43] Horwich (1982, pp. 118 et *sq*).

[44] Nous renvoyons le lecteur à Earman (1992, chap.4) pour une discussion d'autres difficultés importantes.

[45] Ce que, pour leur part, Popper et Miller *ne font pas*.

critère incrémental. (i) Mais la question est de savoir si l'on peut effectivement identifier ($H \vee \neg E$) au contenu de H qui excède E, comme le proposent Popper et Miller. Ce que rejettent les tenants de la TBC (Jeffrey, 1984 ; Howson & Urbach, 1989, 265). (ii) Les partisans de la TBC font également valoir qu'il est fallacieux d'inférer la conclusion anti-inductiviste de la décomposition additive de $d(H, E)$. Ce n'est pas parce que $d(H, E)$ se décomposerait en deux fonctions qui ne peuvent isolément représenter une notion de support confirmationnel que $d(H, E)$, elle, ne peut représenter le support confirmationnel (conféré par E à H)[46]. Cette objection est notamment avancée par Chihara dans un échange avec Gillies (Chihara & Gillies, 1988). Il existe d'autres décompositions de $d(H, E)$. Par exemple, en toute généralité,

$$d(H, E) = \mathrm{d}((H \wedge \neg E), E) + d(H \wedge E, E)$$

Il paraît donc problématique de s'appuyer sur la décomposition de Popper et Miller pour étayer l'argument. (iii) Eells (1988), enfin, fait remarquer avec subtilité que, même si l'on accepte l'essentiel de l'argument, il ne suit pas du fait que E ne B-confirme que la partie de H qu'elle implique déductivement (soit ($H \vee E$)) que la relation de confirmation probabiliste soit purement déductive. Il est facile de construire une paire d'exemples $[(H_1,E_1), (H_2, E_2)]$ où même si $d(H_1 \vee \neg E_1,E_1) = d(H_2 \vee \neg E_2, E_2) < 0$, E_1 B-confirme H_1 tandis que E_2 B-infirme H_2. Par conséquent, le support conféré par E_i à la partie de H_i varie sensiblement d'un cas à l'autre. D'après Eells, cette variation montre que même si la confirmation opère sur la partie de H logiquement impliquée par E, la confirmation elle-même comporte une dimension essentiellement inductive.

4.3.2 Le problème des données connues (*old evidence*)

Nous allons passer désormais à une difficulté que les bayésiens considèrent souvent comme beaucoup plus épineuse, notamment parce qu'elle touche le cœur de la notion incrémentale de confirmation : le fameux *problème des données connues* (« *old evidence problem* »).

Le problème des données connues a été formulé par C. Glymour (1980) au sein d'une batterie d'arguments destinés à rejeter la TBC (pp. 85 et sq.). Voici comment l'on peut présenter le problème. Durant la seconde moitié du XIXe siècle, l'observation astronomique a montré que l'avance du périhélie de Mercure observée (574 secondes d'arc par siècle) différait sensiblement des prédictions que l'on pouvait en faire sur la base de la théorie newtonienne[47]. Supposons que E soit précisément la donnée de cette avance. Considérons comme hypothèse H la Théorie de la Relativité Générale (TRG), supposons en outre que H implique E et plaçons-nous en 1915, au moment où Einstein la formule. Einstein connaissait les données sur l'avance du périhélie de Mercure : $P_{1915}(E) = 1$. E fut considéré par Einstein et la communauté scientifique comme une donnée empirique très importante en faveur de la TRG. On devrait donc s'attendre à ce qu'une théorie de la confirmation correcte (et convenablement paramétrée) accorde un pouvoir confirmationnel important à E. Mais il découle immédiatement du calcul des probabilités que $P_{1915}(H \mid E) = P_{1915}(H)$. Par conséquent, E ne B-confirme pas H. Le moins que l'on puisse dire, c'est qu'il y a ici une divergence importante entre nos intuitions confirmationnelles et la B-confirmation.

[46] Earman (1992), p. 95 ; Horwson & Urbach, p. 264

[47] La théorie newtonienne ne parvient (par calcul des perturbations sur le système à 2 corps Soleil-Mercure) à rendre compte « que » de 531 secondes d'arc par siècle.

Le problème dépasse bien sûr le simple exemple de la Théorie de la Relativité Générale: à partir du moment où une donnée est connue, elle ne peut ni B-confirmer, ni B-infirmer un quelconque énoncé. On peut donner deux versions du problème des données connues. Dans la *version qualitative*, il tient dans le fait que si une donnée E a une probabilité qui vaut 1, elle ne B-confirme ni ne B-infirme aucune hypothèse. Dans sa *version quantitative*, il tient dans le fait que si la probabilité de E vaut $1 - \varepsilon$, alors $d(H,E)$ est compris entre $-\varepsilon$ et ε. La version quantitative met particulièrement en évidence le fait que le problème des données connues n'est rien d'autre que le « mauvais côté » du « principe de surprise » selon lequel toutes choses égales par ailleurs, *moins E* est probable *a priori*, plus l'hypothèse H sera confirmée par E. Le problème touche donc au cœur de la TBC, de sorte que, comme le dit P. Maher (1996), on reconnaît aujourd'hui qu'une théorie bayésienne de la confirmation aussi simple que celle nous venons de présenter n'est pas tenable. Il faut donc modifier la TBC pour résoudre le problème des données connues. Une difficulté supplémentaire tient dans le fait qu'il semble y avoir *plusieurs* problèmes dans le problème des données connues. Il y a au moins deux problèmes qu'il faut distinguer. Il y a d'abord le *problème de l'incrément*[48] : comment une donnée connue E peut-elle augmenter la confiance en une hypothèse H ? Comment par exemple la considération par Einstein en 1915 de l'avance du périhélie de Mercure peut-elle augmenter sa confiance dans la TRG ? Le second problème est le *problème de la survie*[49] : comment le pouvoir confirmationnel d'une donnée E peut-il survivre à son apprentissage ? Dans la TBC, une donnée ne peut en effet plus confirmer ou infirmer après son apprentissage. Il est déjà difficile de modifier la TBC de manière à résoudre l'un des deux problèmes, mais il est évidemment encore plus délicat de résoudre les deux simultanément.

Considérons par exemple l'approche suivante qui incarne un « *bayésianisme (logiquement) désidéalisé* » : on pourrait considérer que dans une situation comme celle d'Einstein, ce qui augmente sa confiance dans la TRG, c'est le fait qu'il se rend compte que la TRG prédit l'avance du périhélie de Mercure. Autrement dit, Einstein ferait un apprentissage de nature logico-mathématique. Le bayésianisme suppose des agents *logiquement omniscients*, c'est-à-dire des agents qui croient toutes les vérités logiques et toutes les conséquences logiques de leurs croyances; il faut donc l'assouplir quelque peu pour rendre possible la représentation d'un apprentissage logique (Garber, 1983, Jeffrey 1983). Une telle approche permet au mieux de résoudre le problème de l'incrément, mais pas celui de la survie : une fois que la relation logico-mathématique entre H et E est apprise, sa probabilité vaut 1 et elle ne peut plus avoir de pouvoir confirmationnel relativement à H.

Le problème de la survie motive un autre type d'approche, qui incarne cette fois un « *bayésianisme historicisé* ». Supposons en effet que nous nous situions au moment t et que E soit connue en t (donc $P_t(H \mid E) = P_t(H)$). Pour juger du support confirmationnel de E vis-à-vis de H, pourquoi ne pas « remonter » l'histoire épistémique de l'agent jusqu'au moment où il a appris E (disons en $t' < t$) et considérer que E confirme H ssi $P_{t'}(H \mid E) > P_{t'}(H)$. Dans la TBC usuelle, les jugements confirmationnels d'un agent au moment t surviennent sur sa distribution de probabilités en t : si deux agents ont la même distribution de probabilités en t, ils auront les mêmes jugements confirmationnels. La solution que nous venons d'esquisser élargit la base sur laquelle les jugements confirmationnels surviennent puisqu'il s'agit

[48] Le problème de l'incrément correspond approximativement à ce que Garber (1983) appelle le « problème historique des données connues », Eells (1990) « *the problem of new old evidence* », Christensen (1999) le « problème diachronique des données connues », Joyce (1999) le « problème de la nouvelle hypothèse » ou le « problème de l'apprentissage logique ».

[49] Le problème de la survie correspond approximativement à ce que Garber (1983) appelle le « problème anhistorique des données connues », Eells (1990), Christensen (1999) le « problème synchronique des données connues », Joyce (1999) le « problème de la pertinence évidentielle ».

désormais de l'ensemble de l'histoire épistémique de l'agent. Cette solution n'est toutefois pas satisfaisante car elle fait dépendre les jugements confirmationnels des accidents de l'histoire épistémique de l'agent (Christensen 1999, voir aussi Maher 1996). Considérons l'exemple suivant : Paul se promène dans le Bois de Vincennes et à t_1 découvre des excréments de cerf (E_1), ce qui B-confirme fortement l'hypothèse H selon laquelle il y a un cerf dans le Bois de Vincennes. A t_2, il découvre des ramures de cerf (E_2), mais compte tenu du fait que la probabilité de H vient d'être largement augmentée, E_2 B-confirme très faiblement H. Intuitivement, Paul peut considérer au moment actuel t ($> t_2 > t_1$) que E_1 et E_2 confirment H aussi bien l'une que l'autre. Ce n'est pas le verdict que donne le bayésianisme historicisé qui confère à E_1 un bien plus grand pouvoir confirmationnel qu'à E_2. Cela est d'autant plus contre-intuitif que, si le hasard avait fait que Paul découvrît les ramures avant de découvrir les excréments, E_1 aurait eu un pouvoir confirmationnel bien plus faible que E_2.

On peut réagir en proposant une théorie de la confirmation fondée sur un « *bayésianisme contrefactuel* »: la confirmation conférée par E à H serait alors l'incrément probabiliste induit par E dans la distribution de probabilités la plus proche de la distribution actuelle de l'agent où il ne connaît pas la donnée E. Cela revient à considérer la question suivante : la probabilité de H serait-elle augmentée si l'agent ne savait pas initialement, puis apprenait, que E ? La TBC contrefactuelle a été défendue par Howson (1984, 1991) mais n'a pas plus fait consensus que l'approche par l'apprentissage logique évoquée précédemment. D'une part, il n'est pas clair que cette approche puisse résoudre les difficultés du bayésianisme « historicisé » : tout dépend de la façon dont on appréhende l'idée de distribution de probabilités la plus proche de la distribution actuelle. Dans le scénario du cerf du Bois de Vincennes, si E_1 a une probabilité de 1 dans la distribution contrefactuelle de référence pour l'évaluation du pouvoir confirmationnel de E_2,, alors E_2 pourra très bien se voir dotée d'un pouvoir confirmationnel très faible. Mais au moins une forme de symétrie est-elle rétablie, puisque si E_2 a également une probabilité de 1 dans la distribution contrefactuelle de référence pour l'évaluation du pouvoir confirmationnel de E_1, alors E_1 se verra probablement elle aussi dotée d'un faible pouvoir confirmationnel. D'autre part, il semble que le bayésianisme contrefactuel constitue au mieux une solution au problème de la survie. C'est le problème de l'incrément, cette fois, qui est laissé de côté puisqu'il est seulement question d'incrément contrefactuel, pas d'incrément actuel.

Une autre approche a été défendue récemment par Christensen (1999) et Joyce (1999). Elle repose sur une mesure de confirmation différente de $d(.,.)$. On peut en effet remarquer que E B-confirme H ssi $d(H, E) > 0$ mais également (quand $P(E) < 1$) ssi $P(H \mid E) > P(H \mid \neg E)$. Si l'on pose $s(H, E) = P(H \mid E) - P(H \mid \neg E)$, on obtient une nouvelle mesure de confirmation. *Prima facie*, il peut sembler étonnant d'avoir recours à $s(.,.)$ pour aborder le problème des données connues puisque $s(.,.)$ n'est pas définie pour $P(E) = 1$. Mais si l'on s'en tient à la version quantitative du problème (Christensen 1999) ou si l'on modifie le cadre bayésien de manière à autoriser la conditionalisation sur des événements à probabilité nulle (Joyce, 1999), alors cette mesure a des propriétés intéressantes pour le problème des données connues. En effet, contrairement à $d(.,.)$, $s(.,.)$ rend possible le fait qu'une donnée E apporte une confirmation significative voire importante à une hypothèse H alors même que, initialement, la probabilité de E est très proche de 1. De manière plus générale, $s(.,.)$ permet de neutraliser le rôle dans la confirmation de la probabilité initiale de E[50]. Pour des raisons différentes, Christensen et Joyce ne considèrent pas pour autant que *la* bonne théorie de la confirmation est la théorie bayésienne de la confirmation fondée sur $s(.,.)$. Mais leur proposition est loin de faire l'unanimité (Earman 1992 par anticipation, Eells & Fitelson

[50] On peut donner une signification précise à cette idée : $s(H, E)$ est invariant par application de la règle de Jeffrey (voir ci-dessus) à la partition $\{E, \neg E\}$. Cela signifie que si l'on applique la règle à cette partition, la valeur $s(H, E)$ ne changera pas, quelle que soit la nouvelle probabilité de E.

2000). A l'heure qu'il est, il n'y toujours pas de solution bayésienne canonique au problème des données connues, qui reste une difficulté majeure pour la TBC.

5 Bayésianisme, objectivité et problème de l'induction

Nous avons vu que le bayésianisme permet de construire une théorie de la confirmation séduisante sous bien des aspects. Nous avons vu également que le problème de l'induction et la clarification de la notion de confirmation sont intimement liés. Une question que l'on peut donc se poser (et qui est discutée dans la littérature contemporaine, Howson 2001, Strevens 2004) est celle de savoir si le bayésianisme permet de « résoudre » le problème de l'induction. Commençons par clarifier le problème de l'induction.

5.1 Le problème de l'induction

Depuis les pages célèbres que D. Hume (*Traité de la nature humaine*, 1739 ; *Enquête sur l'entendement humain*, 1748) y a consacrées, le problème de l'induction est l'un des problèmes fondamentaux de l'épistémologie et de la philosophie générale des sciences. On se doute que, de sa formulation humienne à la « nouvelle énigme de l'induction » de Goodman, « le » problème de l'induction a connu d'importantes transformations. Nous allons donc commencer par tenter de le clarifier quelque peu.

Le problème de l'induction est souvent conçu comme un problème de *justification* : comment justifier notre confiance dans la vérité de certaines propositions étant donné les informations dont nous disposons (« l'expérience ») ? Si le problème se pose, c'est que les informations dont nous disposons ne nous donnent pas la *garantie logique* de la vérité des propositions qui nous intéressent. Par exemple, quand bien même toutes les mesures que nous ayons jamais prises se conformeraient à la Loi d'Ohm, nous n'aurions pour autant aucune garantie logique que la loi est vraie. Ce qu'il s'agit de justifier, ce n'est pas tant notre croyance dans la vérité de telle ou telle proposition (disons, la Loi d'Ohm), mais le fait que les informations dont nous disposons nous donnent confiance dans la vérité de la Loi d'Ohm. On considère souvent le cas où l'on infère un énoncé universel d'observations empiriques, nécessairement particulières et en nombre fini. Nous avons vu dans la première section que le raisonnement inductif au sens large (ou *raisonnement ampliatif*) débordait largement un type de raisonnement comme la généralisation (ou induction énumérative). Une formulation appropriée du problème de l'induction serait donc la suivante : comment justifier les « bons » raisonnements ampliatifs dont nous faisons usage aussi bien dans la vie quotidienne que dans les sciences ?

J'appellerai le problème de l'induction ainsi formulé, qui est aussi le problème traditionnel de l'induction, *le problème de la justification de l'induction-comme-inférence*. On peut le reformuler ainsi : soit **IND**(P, C) une inférence inductive (ou plus généralement une méthode inductive) qui conduit d'une prémisse P à une conclusion C. Considérons par exemple

P = « Toutes les personnes qui, dans le passé, se sont jetées de la Tour Eiffel sans parachute sont mortes », et

C = « La prochaine personne qui se jettera de la Tour Eiffel sans parachute mourra ».

Dans le cas général, *P* n'implique pas logiquement *C*, donc il est logiquement possible que *P* soit vraie et que *C* soit fausse. Nous n'avons donc pas de garantie logique de la préservation de la vérité de *P* à *C*. Qu'est-ce qui peut, par conséquent, justifier le fait que nous nous fiions à **IND**(.,.) pour passer de *P* à *C* ? L'argument que l'on fait remonter à Hume[51] entend montrer qu'il n'y a pas de justification possible à l'usage de **IND**(.,.). Par hypothèse, ce n'est pas une simple déduction qui nous fait passer de *P* à *C*. On peut soutenir qu'il y a bien une inférence déductive sous-jacente, mais une inférence qui repose sur une prémisse supplémentaire. Considérons la prémisse d'uniformité temporelle de la nature

U = « S'il a toujours été vrai dans le passé que si *x* a la propriété P, *x* a la propriété Q, alors il sera vrai du prochain *x* observé que s'il a la propriété P, il aura également la propriété Q »[52].

Admettons que *P* et *U* impliquent logiquement *C*. Est-on parvenu à justifier notre usage de **IND**(.,.) ? Seulement dans la mesure où la prémisse supplémentaire *U* est elle-même justifiée. Mais comment justifier une telle prémisse ? Il ne s'agit pas d'une vérité logique ou analytique. On peut vouloir faire appel à l'expérience : soit

U' = « il a toujours été vrai dans le passé que, quand il avait été vrai dans le passé que si *x* avait la propriété P, *x* avait la propriété Q, alors il était vrai du *x* observé suivant que s'il avait la propriété P, il avait également la propriété Q »[53].

U' peut donner confiance en *U*. Mais comment ? Pas par un raisonnement déductif : *U* n'est pas conséquence logique de *U'*. Si **IND**(*U'*, *U*) – autrement dit si notre méthode inductive nous fait inférer *U* de *U'* – alors la justification semble circulaire[54].

La littérature contemporaine sur le problème de l'induction présente souvent le problème d'une manière assez différente, et il y a des raisons importantes à cela. Dans ce qui précède, nous avons fait comme si notre méthode inductive était une sorte d'extension de la relation de conséquence logique : **IND** garantirait la vérité d'un certain nombre de

[51] On l'aura compris, nous ne visons pas ici à la rigueur exégétique. Par ailleurs, nous suivons la reconstruction usuelle de Hume qui suppose que toute justification authentique doit être déductive. On trouve une excellente discussion de ce point et une interprétation différente dans Stroud (1977, chap.III).

[52] Le rôle et le statut d'un tel principe d'uniformité sont discutés depuis Hume (1739, I, III, VI).

[53] Voir aussi Mill (1843, Livre III, Chap.III) et Strawson (1952), pp. 251 et sq. sur la « prémisse suprême des inductions »

[54] D. Hume, *Enquête sur l'entendement humain*, 1787 « Nous avons dit que tous les arguments relatifs à l'existence se fondent sur la relation de cause à effet ; que notre connaissance de cette relation dérive entièrement de l'expérience ; et que toutes nos conclusions expérimentales procèdent de la supposition que le futur sera conforme au passé. Tenter de prouver cette dernière supposition par des arguments probables, par des arguments qui concernent l'existence ; c'est donc nécessairement et évidemment tourner dans un cercle et prendre pour accordé le point même en question. » (pp. 93-5)

propositions étant donné un ensemble de prémisses *P*, simplement cet ensemble de propositions est plus large que l'ensemble des conséquences logiques de *P*. Comme l'ont fait valoir avec force Strawson (1952) ou Carnap (1950/1962), ce n'est manifestement pas ainsi que les choses se passent. Le raisonnement inductif est, dans le cas général, affaire de degrés. Les prémisses *P* nous donnent une confiance qui peut être élevée, modérée, faible, etc. en une proposition *C*[55]. Savoir que le traitement Viralyse est efficace sur 80% des personnes atteintes d'une certaine maladie nous donne une certaine confiance dans la proposition selon laquelle « le traitement Viralyse sera efficace sur S. qui est atteint de la maladie ». Pour le dire autrement, la proposition *P soutient* plus ou moins la proposition *C*. Carnap aborde la relation entre la logique déductive et la logique inductive au §43 de ses *Foundations* (1950/1962) et mène une comparaison systématique entre les deux. La logique déductive établit par exemple que *E* a pour conséquence logique *H* tandis que la logique inductive établit par exemple que *H'* est confirmé au degré *r* par *E'*. Carnap met en avant le fait que, de la vérité de *E*, on peut inférer celle de *H*, tandis que de la vérité de *E'*, on ne peut rien inférer concernant celle de *H'*. Il revient sur cette idée au § 44 :

> « le terme 'inférence' dans son usage ordinaire implique une transition d'énoncés donnés à de nouveaux énoncés ou l'acquisition d'un nouvel énoncé à partir d'énoncés déjà acquis. Mais seule l'inférence déductive est une inférence dans ce sens. Si [un observateur] découvre que ses connaissances confirment un autre énoncé à un certain degré, il ne doit pas ajouter cet énoncé à la liste de ses connaissances. Le résultat de son examen inductif ne peut être formulé par l'énoncé seul ; la valeur du degré de confirmation qui a été découverte est une part essentielle du résultat. »

Supposons donc que nous ayons une méthode inductive **IND** qui, étant donné deux propositions *P* et *C*, est capable de déterminer quel est le type ou la quantité de soutien conféré à *C* par *P*. La question que l'on est en droit de se poser est alors : qu'est-ce qui justifie les verdicts de **IND** ? On peut également la formuler ainsi : supposons que **IND** soit une méthode inductive qui délivre des verdicts qui nous semblent raisonnables et soit **IND'** une méthode inductive dont, au contraire, les verdicts nous semblent totalement contre-intuitifs. Qu'est-ce qui peut justifier notre préférence pour **IND** plutôt que pour **IND'** ? Nous appellerons cette seconde formulation du problème de l'induction *le problème de la justification de l'induction-comme-support*[56]. Le genre de difficultés qui se posait pour la justification de l'induction-comme-inférence se pose de nouveau pour la justification de l'induction-comme-support et le fait que, en apparence du moins, nous exigions moins que la certitude de nos méthodes inductives, ne change pas le problème[57]

Terminons en disant un mot de la « nouvelle énigme de l'induction » de N. Goodman, le paradoxe des émeraudes vreues. Nous l'avons abordée précédemment en examinant la théorie

---

[55] Voir aussi Goodman (1955, 61) à propos de ce qui correspond à peu près à ce que nous avons appelé le problème de la justification de l'induction-comme-inférence : « si le problème est d'expliquer comment nous savons que certaines prédictions se révéleront correctes, une réponse suffisante consiste à dire que nous ne savons rien de tel. »

[56] Von Wright (1957) fait une distinction analogue quand il contraste les tentatives pour justifier la méthode inductive comme un genre de raisonnement qui conduit à la certitude avec les tentatives pour justifier la méthode inductive comme un genre de raisonnement qui délivre de la « connaissance probable ».

[57] On trouve dans Skyrms (1966, chap.2) une excellente reconstruction du problème de l'induction pour l'induction-comme-support. Des panoramas des solutions contemporaines au problème de l'induction sont présentés dans le même chapitre, ainsi que, plus récemment, dans Earman & Salmon (1992).

hempelienne de la confirmation ; et nous avons vu que l'une des leçons qu'on pouvait être tenté de tirer du paradoxe est qu'une théorie de la confirmation qui se base uniquement sur la forme logique des énoncés risque de ne pas pouvoir rendre compte de différences intuitives (par exemple entre le comportement inductif du prédicat « vert » et celui du prédicat « vreu »). Cette leçon soulève *le problème* (plus général) *de la construction d'une théorie satisfaisante de l'induction* (Skyrms, 1966, chap. 1 et 4). Comment, en effet, élaborer une théorie du raisonnement inductive précise, rigoureuse, et qui s'accorde avec une bonne partie de nos intuitions pré-théoriques ? *Prima facie*, ce problème est distinct (des deux versions) de celui de la justification de l'induction que nous venons de rencontrer, et, d'une certaine façon, il est logiquement premier. On rencontre le problème de la construction quand on élabore une méthode inductive, et si l'on arrive effectivement à construire une telle méthode, se pose alors le problème de sa justification. Goodman (1955) lie cependant sa « nouvelle énigme de l'induction » à une conception de la justification de l'induction surprenante qui a pour conséquence d'effacer la distinction entre les problèmes de la justification de l'induction et celui de la construction d'une théorie de l'induction. Selon Goodman, en effet, la justification de l'induction doit être analogue à celle de la déduction. Les deux justifications doivent procéder par va-et-vient entre les règles déductives (resp. inductives) et la pratique déductive (resp. inductive) : une inférence déductive (resp. inductive) est justifiée dans la mesure où elle est conforme aux règles de la déduction (resp. de l'induction) ; et les règles de la déduction (resp. de l'induction) sont valides ou correctes si elles se conforment à la pratique déductive (resp. inductive). Il faut comprendre l'idée goodmanienne de manière dynamique : les pratiques inférentielles et les théories sont dans un processus d'ajustement mutuel qui, dans le cas favorable, se stabilise sur ce que la littérature a ensuite appelé un « équilibre réfléchi ». Dans le cas de l'induction, l'idée se formule donc de la manière suivante : « Des prédictions sont justifiées si elles sont conforment aux canons valides de l'induction ; et les canons sont valides s'ils codifient avec précision la pratique inductive acceptée » (Goodman, 1955, 64). Si l'on suit cette idée, l'élaboration d'une théorie du raisonnement inductif justifiera en bonne partie ses règles ou principes inductifs par comparaison avec nos intuitions inductives – et avec la pratique statistique. C'est une idée qui semble être partagée par Carnap (1963) mais également par une bonne partie des partisans de la théorie bayésienne de la confirmation[58].

5.2 Quand Hume rencontre Bayes

Nous pouvons maintenant revenir à la question de savoir si le bayésianisme permet de « résoudre » le problème de l'induction. *Prima facie*, on pourrait croire que tel est le cas. (1) Le bayésianisme fournit un cadre théorique et un critère pour considérer qu'une donnée supporte une hypothèse, voire supporte plus une hypothèse H_1 qu'une hypothèse H_2. (2) La TBC permet de rendre compte de nombreuses intuitions confirmationnelles, et, par conséquent, constitue une solution plausible à ce que nous avons appelé le problème de la construction d'une théorie satisfaisante de l'induction. (3) Le bayésianisme fournit des *justifications* pour l'emploi de son cadre et de son critère, comme par exemple l'argument du

[58] Dans l'une de ses premières contributions au sujet, « On Inductive Logic » (1945), Carnap défend la fonction de confirmation c^* (voir ci-dessous). Outre les axiomes probabilistes usuels, il mentionne un axiome de symétrie qui exige que le degré de confirmation soit invariant par permutation des constantes d'individus – une idée que l'on retrouve en logique *déductive*, notamment chez Tarski, dans la caractérisation des constantes logiques. Parmi les fonctions de confirmation qui satisfont ces propriétés, c^* est la seule à être une distribution équiprobable sur les structures d'état. Pourquoi choisir celle-ci ? Carnap dit que le principal facteur qui justifie ce choix réside dans les conséquences que l'on peut tirer de l'usage de c^*. Cela étant dit, Carnap reconnaît la différence entre construire une théorie de l'induction qui soit conforme à nos intuitions pré-théoriques (il pense que la fonction de confirmation c^* permet d'édifier une telle théorie) et le problème « beaucoup plus difficile », dit-il, qu'est l'authentique problème philosophique de l'induction : qu'est-ce qui justifie les verdicts d'une méthode inductive donnée ?

Pari Hollandais. Ce qui suggère que la TBC fournit également une solution au problème de la justification de l'induction (dans sa variante induction-comme-support).

Mais supposons qu'une donnée E B-confirme une hypothèse H. Cela signifie que, pour l'individu considéré, disons Paul, son degré de croyance en H est inférieur à son degré de croyance en H étant donné E. Dans le cas général, il se pourrait que, pour un autre individu, disons Jean, avec d'autres degrés de croyance, E B-infirme H. Pour autant, ni Jean ni Paul ne se trompent du point de vue bayésien : ils n'ont tout simplement pas les mêmes probabilités subjectives. En cela réside la *subjectivité* de la théorie bayésienne de la confirmation, qui se traduit formellement par le fait que les probabilités *a priori* sont supposées quelconques. Or, on voit mal comment cette subjectivité serait compatible avec l'ambition d'apporter une réponse au problème de la justification de l'induction : on attend typiquement d'une méthode inductive (à justifier) qu'elle délivre des verdicts confirmationnels quand on lui soumet des données E et une hypothèse H. (De la même façon que nos « méthodes déductives » délivrent des verdicts déductifs quand on leur soumet un ensemble de prémisses P et une conclusion (putative) C). Mais la TBC ne délivre en général ces verdicts que lorsque l'on se donne en outre les probabilités subjectives associées à E et H. Les difficultés que nous venons de soulever sont encore très largement débattues, et nous allons dans ce qui suit nous contenter d'exposer certaines des considérations principales qui entrent dans la discussion.

Rappelons-nous des différentes formes du théorème de Bayes (TB1)-(TB3). (TB2) montre que si les vraisemblances $P(E \mid H)$ et $P(E \mid \neg H)$ et la probabilité *a priori* $P(H)$ sont données, alors cela suffit à déterminer la probabilité conditionnelle $P(H \mid E)$. De nombreux bayésiens font valoir l'objectivité des vraisemblances. Premièrement, quand H implique E ou $\neg E$, la vraisemblance est fixée (1 ou 0) et identique pour tous les individus. Bien sûr, ce n'est pas le cas général. Mais il existe une vaste famille de « cas favorables » : quand l'hypothèse H est statistique[59] ou quand elle est mise en relation avec les données empiriques par des hypothèses auxiliaires statistiques. Par exemple, si

H = « il y a une chance sur deux pour qu'un noyau de plutonium 233 se désintègre pendant une période de 20 minutes »,

A = « a est un noyau de plutonium 233 », et

E = « a se désintègre pendant une période de 20 minutes donnée »,

alors relativement l'arrière-plan fourni par A, la probabilité de E étant donné H peut être considérée comme valant un-demi (Hawthorne 2007). Du point de vue philosophique, il est très important de remarquer que rien dans le bayésianisme *n'oblige* à ce que la vraisemblance (et tout degré de croyance en général) s'aligne sur les probabilités « objectives ». Le principe qui consiste à endosser les probabilités « objectives » est discuté sous différentes appellations et dans différentes versions, notamment le « principe d'inférence directe », ou le « Principe

[59] Nous utilisons le terme de manière large pour désigner tout usage de probabilité « objective ».

Principal » (Lewis, 1980)[60]. Supposons pour simplifier que l'hypothèse *H* affirme que *E* a *r* chances d'être vrai, ce que l'on abrège $Ch(E) = r$. Alors on peut formuler le principe de la manière suivante :

$$P(E \mid Ch(E) = r) = r$$ [61]

Hawthorne (2009) affirme que même lorsque les vraisemblances ne peuvent être dérivées par inférence directe, les vraisemblances selon différents membres de la communauté épistémique *doivent* être similaires. Son argument est le suivant : la vraisemblance $P(E \mid H)$ exprime en quelque sorte le contenu empirique (probabiliste) de l'hypothèse *H*. Si les vraisemblances de Paul $P_P(E \mid H)$ et Jean $P_J(E \mid H)$ diffèrent grandement, alors cela implique que Paul et Jean ne sont pas d'accord sur le contenu empirique de *H*. Dans ces conditions, il n'est plus clair que, du point de vue épistémique, Paul et Jean envisagent véritablement la même hypothèse[62]. L'attrait de la fixation des vraisemblances par un principe d'inférence directe est tel que certains partisans de la TBC restreignent son usage précisément au cas où il est possible d'appliquer le principe d'inférence directe (Strevens, 2006). Mais même dans ce genre de cas favorable, pour déterminer la probabilité conditionnelle $P(H \mid E)$, il faut avoir recours à la probabilité *a priori* $P(H)$. Or, on ne voit pas ce qui contraindrait Paul et Jean à avoir les mêmes probabilités *a priori* $P_P(H)$ et $P_J(H)$.

On peut voir resurgir la difficulté lorsque l'on songe à la façon dont la TBC analyse différents problèmes épistémologiques classiques. Considérons le cas suivant[63] : soit *E* un ensemble de données qui sont impliquées par deux hypothèses rivales H_1 et H_2. Dans ce cas, le rapport des probabilités conditionnelles $P(H_1 \mid E) / P(H_2 \mid E) = P(H_1) / P(H_2)$. Par conséquent (le rapport entre) les probabilités *a priori* déterminent directement (le rapport entre) les probabilités *a posteriori* et *les données empiriques n'aident en rien à choisir entre les deux hypothèses rivales*. Si l'on adopte l'interprétation du paradoxe des émeraudes vreues (H_1 : « Toutes les émeraudes sont vertes », H_2 : « Toutes les émeraudes sont vreues »), et si *E* est la conjonction de nos données empiriques actuelles sur la couleur des émeraudes, alors cela signifie que l'on va préférer H_1 à H_2 ssi l'on avait déjà cette préférence *a priori*. Le problème de l'objectivité surgit également dans le traitement du problème de Duhem-Quine, comme le souligne Earman (1992, pp. 83-6). La TBC est sans doute capable de rendre compte de décisions théoriques prises par un ou plusieurs scientifiques suite à une réfutation empirique. Mais il se peut parfaitement que les degrés de croyance de Paul doivent, du point de vue bayésien, lui faire blâmer *H* (l'hypothèse centrale) plutôt que *A* (les hypothèses auxiliaires) tandis que ceux de Jean doivent, au contraire, lui faire blâmer *A* plutôt que *H*. Une authentique solution au problème devrait, semble-t-il, recommander une attitude qui s'applique uniformément à Paul et à Jean. Par ailleurs, si nous venons d'insister sur le problème de l'*accord intersubjectif*, on attend plus encore d'une théorie du raisonnement inductif : on attend en outre la garantie que les normes bayésiennes nous conduisent, d'une manière ou d'une autre, vers la *vérité*.

---

[60] On parle aussi de « principe de Miller » ou, plus récemment, de « principe de coordination probabiliste » (Strevens 2006).

[61] Une part importante de la contribution de Lewis (1980) consiste à préciser le domaine de validité du principe, autrement dit à déterminer des classes de situations épistémiques où il paraît raisonnable d'obéir à $P(E \mid Ch(E) = r) = r$.

[62] Voir aussi Strevens (2006) : « L'entreprise scientifique exige un certain consensus parmi les scientifiques concernant la question de savoir comment les données soutiennent ou pas des hypothèses concurrentes. »

[63] Horwich (1982, p.35).

En réponse aux problèmes que nous venons d'exposer, de nombreux bayésiens mettent en avant une série de résultats de convergence (par conditionnalisation) des probabilités individuelles vers les hypothèses vraies (Halmos 1950, Savage 1954, Blackwell & Dubnis 1961, Gaifman & Snir 1982, Schervish & Seidenfeld 1991). Cette convergence en implique une seconde, celle des différentes probabilités subjectives entre elles (généralement sous l'hypothèse qu'elles assignent des probabilités *a priori* nulles aux mêmes propositions). Elle semble donc résoudre du même coup le problème de l'accord intersubjectif. L'interprétation de ces résultats soulève cependant plusieurs difficultés. Par exemple, la convergence ne vaut en général que « presque sûrement » au sens technique, c'est-à-dire qu'elle n'est pas garantie dans les mondes possibles qui appartiennent à des événements de probabilité nulle. Les propriétés de la probabilité *a priori* restent donc déterminantes, comme le soulignent Earman (1992, chap.6, sec.3-5) et Howson (2000, p.210). En outre, il n'est pas complètement clair que ces résultats de « long terme » aient un impact décisif sur la question de la justification de jugements instantanés comme le sont les jugements confirmationnels.

Howson (2000), l'un des principaux partisans de la TBC, soutient que l'argument humien pour le scepticisme inductif est correct, mais que cela n'empêche pas l'existence d'une logique de l'inférence inductive. Cette logique n'est rien d'autre que la TBC. Comme on l'a vu, la TBC a « besoin » de probabilités *a priori*. Ces probabilités *a priori* codent notamment des engagements inductifs, lesquels ne sont pas justifiés par la TBC. On peut par exemple imaginer que, toutes choses égales par ailleurs, les probabilités *a priori* associées par un individu aux hypothèses complexes soient moins élevées que celles qu'il associe aux hypothèses simples. Howson soutient que, par conséquent, « il existe une authentique logique de l'induction qui montre que le raisonnement inductif est logiquement correct étant donné des prémisses appropriées, mais cette logique ne justifie pas ces prémisses. » Pour le dire dans la terminologie que nous avons mise en place : la TBC ne constitue pas une *méthode inductive* mais permet de mettre en œuvre de manière cohérente des engagements inductifs reflétés dans les probabilités *a priori*.

6 Conclusion

Dans un domaine où les problèmes se rencontrent plus souvent que les solutions, la théorie bayésienne de la confirmation est une espèce rare. Nous avons cependant fait valoir que la TBC rencontrait des difficultés tant du point de vue du problème de la *construction* d'une théorie satisfaisante de l'induction (voir par exemple le problème des données connues) que du point de vue du problème de la *justification* de l'induction (voir la section précédente sur la subjectivité de la TBC). Si la TBC occupe seule le devant de la scène philosophique, c'est donc en partie faute de concurrents sérieux. Il nous semble cependant que ses vertus sont suffisantes pour que les philosophes des sciences s'investissent plus dans ses *applications* à des épisodes choisis de l'histoire des sciences. C'est l'un des mérites de Howson & Urbach (1989) que d'embrayer un tel mouvement, qui est resté à ce jour relativement peu développé.

7 Références

Carnap, R., 1945, « On Inductive Logic », *Philosophy of Science,* 12, 72-97

Carnap, R., 1947, « On the Application of Inductive Logic » , *Philosophy and Phenomenological Research*, 8, 133-47

Carnap, R. (1950/1962), *Logical Foundations of Probability*, Chicago : University of Chicago Press

Carnap, R. (1952), *The Continuum of Inductive Methods*, Chicago : University of Chicago Press

Chihara, C.S. & Gillies, D.A., 1988, « An Interchange on the Popper-Miller Argument », *Philosophical Studies*, 54, 1-8

Christensen, D., 1990, « The Irrelevance of Bootstrapping », *Philosophy of Science*, 57, 644-62

Christensen, D., 1999, « Measuring Confirmation », *Journal of Philosophy*, 96, 437-61

De Finetti, B., 1937, « La prevision : ses lois logiques, ses sources subjectives », *Annales de l'Institut Poincaré*, 7, 1-68

Dorling, J., 1979, « Bayesian Personalism, the Methodology of Scientific Research Programs, and Duhem's Problem », *Studies in the History and Philosophy of Science*, 10, 177-87

Earman, J. (1992), *Bayes or Bust ?A Critical Examination of Bayesian Confirmation Theory*, Cambridge (Mass.): MIT Press

Earman, J. & Salmon, W.C. (1992), « The Confirmation of Scientific Hypotheses » , *in* Salmon, M. & *al.*(eds), *Introduction to the Philosophy of Science*, Indianapolis & Cambridge : Hackett Publishers

Eells, E. 1988, « On the Alleged Impossibility of Inductive Probability », *British Journal for the Philosophy of Science*, 39, 111-16

Eells, E., 1990, « Bayesian Problems of Old Evidence », *in* Wade Savage C. (ed.), *Minnesota Studies in the Philosophy of Science*, vol.14, Minneapolis : University of Minnesota Press, 205-23

Eells, E. & Fitelson, B.,2000, « Measuring Confirmation and Evidence » , *The Journal of Philosophy*, 97 (12), 663-72

Fitelson, B. (2001), *Studies in Bayesian Confirmation Theory*, Thèse de philosophie, Université du Wisconsin

Fitelson, B. 2002, « Putting the Irrelevance Back Into the Problem of Irrelevant Conjunction », *Philosophy of Science*, 69:4, 611-622

Fitelson, B., 2006, « The Paradox of Confirmation », *Philosophy Compass*, 1 (1), 95-113.

Fitelson, B. & Hawthorne, J., 2006, « How Bayesian Confirmation Theory Handles the Paradox of the Ravens » *in* Eells, E. & Fetzer, J., *Probability in Science*, Chicago : Open Court

Garber, D., 1983, « Old Evidence and Logical Omniscience in Bayesian Confirmation Theory » , *in* Earman, J. (ed.), *Minnesota Studies in the Philosophy of Science*, vol.10, Minneapolis : University of Minnesota Press

Glymour, C. (1980), *Theory and Evidence*, Princeton : Princeton University Press

Good, I.J., 1967, « The White Shoe is a Red Herring », *British Journal for the Philosophy of Science*, 12, 63-4

Goodman, N., 1946, « A Query on Confirmation », *Journal of Philosophy*, XLIII (14), 383-5

Goodman, N., 1947, « On Infirmities of Confirmation-Theory », *Philosophy and Phenomenological Research*, 8, 149-51

Goodman, N. (1955), *Fact, Fiction and Forecast*, Cambridge (Mass.) : Harvard University Press ; trad.fr. ,

Hacking, I. (2001), *An Introduction to Probability and Inductive Logic*, Cambridge : Cambridge University Press.

Hajek, A. & Joyce, J., 2008, « Confirmation » , *in* Psillos, S. & Curd, M., *Routledge Companion to the Philosophy of Science*, Routledge : Londres

Hawthorne, J., « Inductive Logic », *The Stanford Encyclopedia of Philosophy (Summer 2007 Edition)*,Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2007/entries/logic-inductive/>.

Hawthorne, J. (2009), « Confirmation Theory », à paraître dans P.S. Bandyopadhyay & M. Forster (eds), *Philosophy of Statistics, Handbook of the Philosophy of Science*, vol.7, Elsevier

Hempel, C.G., 1945, « Studies in the Logic of Confirmation (I) », *Mind*, LIV (213), 12-26

Hempel, C.G., 1945, « Studies in the Logic of Confirmation (II) », *Mind*, LIV (214), 97-121

Hempel, C.G., 1967, « The White Shoe : No Red Herring » , *British Journal for the Philosophy of Science*, 18, 239-40

Horwich, P., (1982), *Probability and Evidence*, Cambridge : Cambridge University Press

Howson, C. & Urbach, P., (1989), *Scientific Reasoning : The Bayesian Approach*, La Salle : Open Court

Hume, D. (1739), *Treatise of Human Nature*

Hume, D. (1748), *Enquiry into the Human Understanding*

Huygens, C. (1690), *Traité de la Lumière*, Van der Aa : Leyden

Jeffrey, J., 1983, « Bayesianims with a Human Face », *in* Earman, J. (ed.), *Minesota Studies in the Philosophy of Science*, vol.10, Minneapolis : University of Minnesota Press

Jeffrey, R. 1984, « The Impossibility of Inductive Probability », *Nature*, 310, 433

Joyce, J., (1999), *Foundations of Causal Decision Theory*, Cambridge : Cambridge University Press

Joyce, J, 2007, « Bayes' Theorem », *The Stanford Encyclopedia of Philosophy (Summer 2007 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2007/entries/bayes-theorem/>.

Lewis, D.K. (1999), *Papers in Metaphysics and Epistemology*, Cambridge : Cambridge University Press

Maher, P., 1999, « Inductive Logic and the Ravens Paradox », *Philosophy of Science*, 66, 50-70

Maher, P., 2004, « Probability Captures the Logic of Scientific Confirmation », *in* Hitchcock, C. (ed.) *Contemporary Debates in the Philosophy of Science*, Oxford : Basil Blackwell, 69-93

Popper, K.R. (1959/1968), *The Logic of Scientific Discovery*, Hutchinson : Londres ; trad.fr. N.Thyssen-Rutten et Ph.Devaux, *La logique de la découverte scientifique*, Paris : Payot, 1978

Popper, K.R. & Miller, D., 1983, « A Proof of the Impossibility of Inductive Probability», *Nature*, 302, 687-8

Quine, W.V.O. (1969), *Ontological Relativity and Other Essays*, New York : Columbia University Press ; trad. fr. par J. Largeault, *Relativité de l'ontologie et autres essais*, Paris : Aubier-Montaigne, 1977

Ramsey, F.P., 1931, « Truth and Probability », *in* Braithwaite, R. (ed.) *Foundations of Mathematics and Other Logical Essays*, Londres : Routledge & Kegan Paul

Savage, L. (1954), *The Foundations of Statistics*, Dover: New-York

Skyrms, B. (1966), *Choice and chance*, Dickenson: Belmont, Cal.

Sober, E., 1994, « No Model, No Inference : A Bayesian Primer on Grue Problem » *in* Stalker, D. (ed.), *Grue ! The New Riddle of Induction*, Chicago : Open Court

Strevens, M., 2004, « Bayesian Confirmation Theory: Inductive Logic, or Mere Inductive Framework? », *Synthese*, 141, 365-79.

Strevens, M. (2006), *Notes on Bayesian Confirmation Theory*, Miméo.

Stroud, B. (1977), *Hume*, Routledge & Kegan Paul : Londres

Teller, P., 1973, « Conditionalization and Observation », *Synthese*, 26(2), 218-58

Talbott, William, « Bayesian Epistemology », *The Stanford Encyclopedia of Philosophy (Fall 2006 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2006/entries/epistemology-bayesian/>.

Vickers, J., 2006, « The Problem of Induction », *The Stanford Encyclopedia of Philosophy (Winter 2006 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2006/entries/induction-problem/>.

Vranas, P., 2004. « Hempel's Raven Paradox: A Lacuna in the Standard Bayesian Solution », *British Journal for the Philosophy of Science*, 55, 545– 60.

Von Wright, G.H. (1965), *The Logical Problem of Induction*, Oxford : Basil Blackwell

Sous la responsabilité scientifique de:
JACQUES DUBUCS, MIKAEL COZIC,
PHILIPPE MONGIN.

Dans la même série

IHPST Éditions
13, rue du Four
75006 Paris