La fondation logique des probabilités chez Carnap

Karine Fradet
Université de Montréal

Résumé

Carnap a eu une participation active dans les discussions entourant la crise des fondements avant de se lancer dans le projet d'établir les fondements logiques des probabilités, ce qu'il présente dans « Logical Foundations of Probability » (1950). Il a montré que les querelles entre l'interprétation des probabilités comme état du monde et comme état de connaissance de l'observateur étaient vaines puisqu'elles ne portaient pas sur le même concept. Pour lui, la logique des probabilités est la logique inductive : une probabilité décrit une relation logique, soit le degré de confirmation d'une hypothèse selon des données observées. Deux ans plus tard, dans « The Continuum of Inductive Methods » (1952), il enrichit son système et démontre qu'il existe une quantité infinie de méthodes inductives. Nous replacerons l'interprétation de Carnap parmi celles discutées à son époque et discuterons de certaines implications.

Carnap occupe un point central dans le développement de la logique inductive. D'un côté, il a montré que les querelles entre l'interprétation des probabilités comme état du monde et comme état de connaissance de l'observateur étaient vaines puisqu'elles ne portaient pas sur le même concept. Ensuite, il s'est penché sur un de ces deux concepts, les probabilités inductives, et a jeté les bases de la logique inductive dans *Logical Foundations of Probability* (1950). Il a ensuite systématisé les différentes méthodes inductives dans *The Continuum of Inductive Methods* (1952) où il présente les différentes méthodes en compétition comme faisant partie d'un système, chaque méthode étant un point sur le continuum des méthodes inductives.

Nous retracerons ici les grandes lignes de ces développements. Nous verrons tout d'abord comment Carnap tranche la querelle entre les tenants de l'interprétation fréquentiste des probabilités et ceux soutenant une interprétation épistémique en éclaircissant les fondements de cette querelle. Ensuite, nous esquisserons la conception laplacienne des probabilités qui est un bon exemple d'interprétation inductive des probabilités. Nous nous pencherons ensuite sur le travail de systématisation des méthodes inductives accompli par Carnap.

1. Deux concepts de probabilités

Carnap souligne à plusieurs reprises (par exemple, 1945b, 1955, 1962, 1973) une confusion autour du concept de « probabilité » : certains y voient une propriété empirique alors que d'autres y voient un état de connaissance (ou plutôt d'ignorance) d'un observateur. Si cette situation crée une confusion, c'est que les deux concepts ne sont pas différenciés : « Practically everyone will say that there is only one scientific meaning [for 'probability']; but when you ask that it be stated, two different answers will come forth. » (Carnap 1955: 1) Selon Carnap, il ne s'agit pas de deux interprétations du même concept, mais bien de deux concepts différents. Si un seul concept préscientifique avait été interprété de deux façons, il s'agirait de deux interprétations. Or, dans le cas de « probabilité », il s'agit de deux concepts préscientifiques différents qui ont donné lieu à deux catégories de concepts scientifiques différents (Carnap 1945b). Il ne s'agit donc pas de deux interprétations en compétition, mais de deux concepts distincts qui ont chacun leur importance et leur place dans les sciences.

Le premier de ces concepts préscientifiques est celui qu'il nomme probabilité$_1$ ou probabilité inductive. Il s'agit d'un concept de nature logique et qui réfère au degré de confirmation d'une hypothèse *h* par rapport à un ensemble de données *e*. Selon Carnap, les premiers écrits sur les probabilités traitent de cette conception inductive des probabilités, par exemple dans le *Ars conjectandi* de Bernouilli (1713) ou la *Théorie analytique des probabilités* de Laplace (1814). Au 20^e siècle, on la retrouve chez des auteurs comme Keynes (1921) et Jeffreys (1939 [1998]).

Le deuxième concept préscientifique est celui qu'il nomme probabilité$_2$ ou probabilité statistique. Il s'agit d'un concept de nature empirique qui réfère à la fréquence relative d'une certaine propriété dans une population. Cette conception des probabilités a pris une place importante en sciences depuis la seconde moitié du 19e siècle et a été développée par des auteurs comme Fisher (1922), von Mises (1928 [1957]) et Reichenbach (1935).

Alors que les probabilités statistiques font référence aux propriétés physiques d'un objet ou d'un système (par exemple, la composition matérielle d'un dé ou la distribution d'un trait dans la population), les probabilités inductives sont relatives à l'information disponible pour un observateur (par exemple, savoir qu'un dé a la forme d'un cube sans connaître la distribution de la masse à l'intérieur). Carnap insiste sur le fait que les deux concepts sont tout aussi objectifs. Les allégations de subjectivité contre le concept de probabilité inductive seraient dues à certaines formulations subjectivistes qui portent à confusion comme celle de « degré de croyance ». Or, selon lui, les auteurs qui utilisent de telles formulations soutiennent malgré tout des théories essentiellement objectivistes (Carnap 1945b).

Le développement des probabilités statistiques étant en plein essor, Carnap se penchera sur les probabilités inductives et il établira les fondations de la logique inductive dans *Logical Foundations of Probability* (1950). Il systématisera ensuite les méthodes inductives qu'il présentera comme n'étant que quelques méthodes possibles, quelques points sur un continuum, dans *The Continuum of Inductive Methods* (1952). Mais d'abord, nous nous tournerons vers Laplace qui a joué un rôle prédominant dans les réflexions de Carnap afin de voir en quoi consistait sa vision des probabilités.

2. Les probabilités inductives avant Carnap : Laplace

Laplace soutient une conception déterministe du monde. Dans un tel cadre métaphysique, l'idée de probabilités n'est possible que dans l'ignorance des conditions initiales de ce monde et des lois qui y opèrent. L'idée même d'une probabilité ne peut donc être due qu'à l'état de nos connaissances qui sont incomplètes : « La probabilité est relative en partie à [notre] ignorance, et en partie à nos connaissances. » (Laplace 1814: iv) Il définit la probabilité d'un événement comme suit :

> La théorie des hasards consiste à réduire tous les événements du même genre, à un certain nombre de cas également possibles, c'est-à-dire, tels que nous soyons également indécis sur leur existence ; et à déterminer le nombre des cas favorables à l'événement dont on cherche la probabilité. Le rapport de ce nombre à celui de tous les cas possibles, est la mesure de cette probabilité qui n'est ainsi qu'une fraction dont le numérateur est le nombre de cas favorables, et dont le dénominateur est le nombre de tous les cas possibles. (Laplace 1814: iv)

On doit donc posséder une liste des cas également possibles, soit ceux où nous sommes également indécis ou pour lesquels « (…) rien ne porte à croire que l'un de ces cas doit arriver plutôt que les autres (…) » (Laplace 1814: 179) et diviser le nombre de cas favorables par le nombre total de cas possibles.

$$P(x) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

eq. 1) La probabilité d'un événement x chez Laplace.

Sa conception des probabilités est clairement celle d'un degré de confirmation, ce qui est sans équivoque un peu plus loin lorsqu'il donne l'exemple des trois urnes.

> Supposons, par exemple, que l'on ait trois urnes, A, B, C, dont l'une ne contienne que des boules noires, tandis que les deux autres ne renferment que des boules blanches. On doit tirer une boule de l'urne C, et l'on demande la probabilité que cette boule sera noire. Si l'on ignore quelle est celle des trois urnes, qui ne renferme que des boules noires, en sorte que l'on n'avait aucune raison de croire qu'elle est plutôt C, que B ou A ; ces trois hypothèses

> paraitront également possibles ; et comme une boule noire ne peut être extraite que dans la première, la probabilité de l'extraire est égale à un tiers. Si l'on sait que l'urne A ne contient que des boules blanches, l'indécision ne porte plus alors que sur les urnes B et C, et la probabilité que la boule extraite de l'urne C sera noire, est un demi. Enfin, cette probabilité se change en certitude, si l'on est assuré que les urnes A et B ne contiennent que des boules blanches. (Laplace 1814: v-vi)

Si le contenu en tant que tel de l'urne C n'est pas une question de probabilité (la situation est déterminée puisqu'elle contient déjà soit uniquement des boules noires, soit uniquement des boules blanches) on peut tout de même parler de la probabilité d'y piger une boule noire relativement à l'ignorance de la personne qui fera la pige. Si nous apprenons que l'urne A ne contient que des boules blanches, le nombre de cas possibles passe alors de 3 à 2. Grâce à cette information, la probabilité de piger une boule noire dans l'urne C passe de 1/3 à 1/2 puisqu'il n'y a maintenant plus que deux cas possibles qui correspondent à l'emplacement probable des boules noires. Finalement, si nous apprenons que l'urne B contient également des boules blanches, il n'y a plus qu'un seul cas possible où peuvent se trouver les boules noires et la probabilité est alors de 1. Encore une fois, ce ne sont pas les propriétés du système physique qui sont modifiées mais les données qui sont disponibles pour l'observateur.

Relativement à cet exemple, Laplace fait une affirmation qui semble aller contre le bon sens dans le calcul des probabilités mais qui sera centrale chez Carnap : « On voit par cet exemple l'influence des événements passés sur la probabilité des événements futurs. » (Laplace 1814: ix-x) Nous y reviendrons plus loin.

L'exemple de l'urne de Laplace illustre également le principe de la raison insuffisante ou principe d'indifférence : bien que la probabilité « réelle » de piger une boule noire dans l'urne C soit entièrement déterminée (elle est soit de 0, soit de 1), par manque d'information et par la symétrie de la situation, on doit attribuer une probabilité d'un tiers à l'événement qui consiste à piger une boule noire dans l'urne C.

Un exemple tiré de Carnap (1955) illustre bien le principe d'indifférence. Un dé est présenté à trois observateurs. Au premier observateur X_1, on donne l'information e_1 que le dé a la forme d'un cube régulier. Par la condition de symétrie et puisque le dé possède six faces, X_1 en conclut que la probabilité de rouler un as au prochain jet est de 1/6. À l'observateur X_2, on fournit en plus l'information e_2 comme quoi le dé est biaisé en faveur d'un des côtés. Sans autre information, X_2 doit également conclure que la probabilité de rouler un as au prochain jet est de 1/6 puisqu'il n'a aucune raison de favoriser un côté plutôt que les autres : toute présomption tant qu'au côté biaisé serait purement arbitraire. Finalement, à X_3, on ajoute l'information e_3 comme quoi le dé est biaisé en faveur de l'as. Celui-ci peut alors conclure que la probabilité de rouler un as au prochain jet est supérieure à 1/6. Dans les trois cas, les propriétés physiques de l'objet ne changent pas, ce n'est que l'information disponible pour les observateurs qui varie et la probabilité est une inférence logique. « Inductive probability characterizes a hypothesis relative to available information; this information may differ from person to person and vary for any person in the course of time. » (Carnap 1955: 4)

Le dernier principe central chez Laplace dont nous discuterons ici est la règle de succession.

> On trouve ainsi qu'un événement étant arrivé de suite, un nombre quelconque de fois ; la probabilité qu'il arrivera encore la fois suivante, est égale à ce nombre augmenté de l'unité, divisé par le même nombre augmenté de deux unités. (Laplace 1814: xiii)

Cet extrait traite d'un événement s'étant produit à chaque reprise mais plus généralement, la probabilité qu'un événement se produise la fois suivante est égale au nombre de fois qu'il s'est produit dans le passé plus 1, divisé par le nombre total de fois où il aurait pu se produire plus 2[1].

3. Les méthodes inductives

Même si on élimine les désaccords à propos de la *vraie* signification du concept de probabilité et qu'on accepte qu'on puisse en parler de façon inductive, il demeure un vaste choix de méthodes

[1] Voir l'équation 3 ci-dessous.

inductives et il n'y a pas de consensus sur l'identité de celle qui s'avère être la *bonne* méthode. Dans cette section, nous verrons ce problème du choix de la méthode inductive à l'aide d'un exemple tiré de l'article « Statistical and Inductive Probability » (Carnap 1955). Nous discuterons ensuite du travail de systématisation des différentes méthodes inductives effectué par Carnap dans *The Continuum of Inductive Methods* (1952) et verrons quatre exemples de ces méthodes telles que définies par le paramètre λ qui permet de les distinguer.

3.1. Le problème du choix de la méthode inductive

L'exemple suivant illustre le problème du choix d'une méthode inductive. Le tableau 1 ci-dessous résume une situation où quatre balles sont tirées d'une urne qui contient des balles blanches et des balles noires dans une proportion inconnue.

On peut s'intéresser à deux sortes de résultats. On peut d'abord regarder les différentes *distributions individuelles*, soit l'ensemble des permutations possibles. Pour quatre piges, cela donne seize distributions individuelles (le résultat « noire, noire, blanche, noire » en est un exemple). On peut également s'intéresser aux différentes *distributions statistiques*, soit le nombre de balles noires et de balles blanches qui feront partie du décompte final, peu importe l'ordre dans lequel elles ont été pigées. Il y a ici cinq distributions statistiques dont le résultat « trois noires, une blanche » est un exemple.

Dans le tableau 1, la colonne « méthode 1 » illustre une situation où le principe d'indifférence est d'abord appliqué aux distributions individuelles. Puisqu'elles sont au nombre de 16, la probabilité initiale de chacune est de 1/16. Pour connaître la probabilité d'une distribution statistique, on peut additionner la probabilité de chacune des distributions individuelles qui forment cette distribution statistique. Par exemple, une seule distribution individuelle correspond au critère de la distribution statistique « quatre noires », celle-ci a donc une probabilité de 1/16 ; par contre, quatre distributions individuelles répondent au critère de la distribution statistique « trois noires, une blanche », celle-ci a donc une probabilité de 4/16.

#	Distributions statistiques		Distributions individuelles	Méthode 1	Méthode 2	
	Nb noirs	Nb blancs		Prob. initiale des distr. individuelles	Prob. initiale des : distr. statistiques	distr. individuelles
1	4	0	●●●●	1/16	1/5	1/5 = 12/60
2	3	1	●●●○	1/16	1/5	1/20 = 3/60
			●●○●	1/16		1/20 = 3/60
			●○●●	1/16		1/20 = 3/60
			○●●●	1/16		1/20 = 3/60
3	2	2	●●○○	1/16	1/5	1/30 = 2/60
			●○●○	1/16		1/30 = 2/60
			●○○●	1/16		1/30 = 2/60
			○●●○	1/16		1/30 = 2/60
			○●○●	1/16		1/30 = 2/60
			○○●●	1/16		1/30 = 2/60
4	1	3	●○○○	1/16	1/5	1/20 = 3/60
			○●○○	1/16		1/20 = 3/60
			○○●○	1/16		1/20 = 3/60
			○○○●	1/16		1/20 = 3/60
5	0	4	○○○○	1/16	1/5	1/5 = 12/60

Tableau 1 : Tableau des distributions individuelles et statistiques pour quatre piges. Tiré de (Carnap 1955: 9).

La colonne « méthode 2 » représente une méthode où le principe d'indifférence est d'abord appliqué aux distributions statistiques. Puisqu'elles sont au nombre de 5, la probabilité initiale de chacune est de 1/5. On divise ensuite cette probabilité par le nombre de distributions individuelles qui peuvent former une certaine distribution statistique pour en connaître la probabilité. Par exemple, la probabilité de la distribution individuelle « noire, noire, noire, noire » est de 1/5 puisqu'elle est la seule à pouvoir former la distribution statistique « quatre noires » ; la probabilité de la distribution individuelle « noire, noire, blanche, noire » est de 1/20 puisqu'il y a quatre distributions individuelles pouvant former la distribution statistique « trois noires, une blanche ».

Notons que ces considérations ne font de sens que si nous parlons bien de probabilités inductives. Dans le cas de probabilité statistiques, s'il s'agit d'une propriété du système que la probabilité de tirer une boule noire soit, par exemple, de 50%. La méthode 1 est celle à laquelle nous sommes habitués, c'est-à-dire que la probabilité de n'importe quelle distribution individuelle est égale au produit des probabilités pour chaque événement, dans ce cas $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}$, ce qui nous donne 1/16. La méthode 2 ne fait aucun sens si on connaît d'avance la probabilité (au sens fréquentiste) de tirer une boule noire.

Peut-on faire cette même analyse en probabilités inductives ? Après tout, avant même de tirer la première balle, puisqu'il y a deux cas possibles et que nous n'avons aucune raison d'en favoriser un, le principe d'indifférence nous dicte d'attribuer une probabilité de 1/2 aux deux événements. Ceci étant déterminé, on pourrait poursuivre notre démarche avec nos calculs habituels tels que celui décrit ci-dessus. En d'autres mots, nous assumons dès le départ par le principe d'indifférence que la fréquence relative des balles noires est de 1/2 et nous faisons les calculs habituels à partir de cette probabilité *a priori*.

Une telle méthode ne possède toutefois pas ce que Carnap juge être la caractéristique fondamentale d'une méthode inductive raisonnable, soit celle d'apprendre de l'expérience : « Inductive thinking is a way of judging hypotheses concerning unknown events. In order to be reasonable, this judging must be guided by our knowledge of observed events. » (Carnap 1955: 12) En effet, si on fixe de façon *a priori* la probabilité de tirer une balle noire pour tous les tirages à venir sans jamais pouvoir réviser ce jugement, on ne peut pas apprendre de l'expérience au fur et à mesure que les résultats empiriques sont disponibles. De même, la méthode 1 ne permet pas de corriger la probabilité en cours de route, au fur et à mesure que les balles sont tirées. Ce défaut devient clair si nous considérons une situation où nous avons tiré 100 balles de l'urne et qu'elles sont toutes noires : selon la méthode 1, on doit toujours attribuer la probabilité de 1/2 à ce que la prochaine balle tirée soit noire.

Bien que la méthode 2 ne semble pas faire de sens du point de vue des probabilités statistiques, il s'agit d'une méthode tout à fait acceptable selon Carnap au point de vue des probabilités inductives. En effet, il est possible que l'urne ne contienne que des balles noires, que des balles blanches, ou toute autre combinaison entre ces deux extrêmes. Les probabilités sont donc d'abord réparties entre les distributions statistiques avant d'être réparties entre les distributions individuelles. La répartition entre les distributions statistiques se fait selon l'équation ci-dessous qui nous donne la probabilité que le nombre de boules blanches tirées en *n* essais soit égal à B_n.

$$P(B_n) = \frac{1}{n+1}$$

eq. 2) Principe d'indifférence appliqué aux distributions statistiques.

Cela revient à attribuer la même probabilité à chacune des distributions statistiques. Il est à noter que le dénominateur est de $n + 1$ et non de n étant donné qu'il faut compter qu'il est possible qu'une distribution statistique ne contienne aucune boule blanche. Comme nous l'avons vu, la méthode 2 répartit ensuite les probabilités selon le principe d'indifférence entre toutes les distributions individuelles qui constituent une distribution statistique.

Dans cet exemple, cette méthode nous donne le même résultat que l'utilisation de la règle de succession de Laplace qui permet de déterminer les probabilités du prochain tirage en tenant compte des résultats des tirages précédents.

$$P(X_{n+1} = B \mid B_n) = \frac{B_n + 1}{n+2}$$

eq. 3) Règle de succession de Laplace.

Cette équation nous donne la probabilité que la balle (X) au prochain tirage (n+1) soit blanche (*B*), sachant que le nombre de balles blanches dans les n tirages jusqu'ici est de B_n. Si on applique cette règle pour chacun des quatre tirages consécutivement, on obtient la probabilité indiquée dans la colonne d'extrême droite pour chacune des distributions individuelles.

Par exemple, si l'on prend la distribution individuelle « blanche, blanche, noire, blanche », on a les probabilités suivantes pour chacun des tours (soit *B* la propriété d'être blanc et *A* la propriété d'être noir).

1er tour : $$P\left(X_{n+1} = B \mid B_n = 0\right) = \frac{0+1}{0+2} = \frac{1}{2}$$

2e tour : $$P\left(X_{n+1} = B \mid B_n = 1\right) = \frac{1+1}{1+2} = \frac{2}{3}$$

3e tour : $$P\left(X_{n+1} = A \mid A_n = 1\right) = \frac{0+1}{2+2} = \frac{1}{4}$$

4e tour : $$P\left(X_{n+1} = B \mid B_n = 2\right) = \frac{2+1}{3+2} = \frac{3}{5}$$

En multipliant les probabilités ainsi calculées pour chacun des tours, on obtient $\frac{1}{2} \times \frac{2}{3} \times \frac{1}{4} \times \frac{3}{5} = \frac{1}{20}$ comme probabilité pour cette distribution individuelle.

La méthode 2 sous-tend donc une réévaluation de la probabilité inductive à chacun des tirages. Elle répond ainsi à la caractéristique d'apprendre de l'expérience qui est fondamentale chez Carnap puisqu'elle réajuste les probabilités *a priori* au fur et à mesure que des résultats empiriques sont disponibles.

Les deux méthodes présentées ci-dessus ne sont que deux méthodes à l'intérieur d'un continuum de méthodes inductives possibles. Voyons maintenant comment Carnap systématise les différentes méthodes de la logique inductive.

3.2. Le continuum des méthodes inductives

Dans *The Continuum of Inductive Methods* (1952), Carnap montre que les différentes méthodes inductives forment un système et qu'elles peuvent être classées sur un continuum. Le degré de confirmation d'une hypothèse *h* relativement à un ensemble de données *e* peut être décrit par différentes méthodes d'induction ayant chacune une fonction de confirmation *c*(*h*,*e*)[2]. Cette fonction a comme paramètres deux propositions et comme résultat une valeur comprise entre 0 et 1, soit le degré de confirmation. Ces propositions font partie d'un système de langage ayant lui-même les caractéristiques décrites dans le tableau 2 ci-dessous[3].

Symbole	**Description**	**Exemple**
N	Nombre de constantes individuelles. Exemple : $a_1, a_2, ..., a_N$	a_1 : première balle a_2 : deuxième balle
π	Nombre de prédicats primitifs. Exemple : $P_1, P_2, ..., P_\pi$	P_1 : est blanc P_2 : est lisse
$k = 2^\pi$	Nombre de propriétés *Q* (conjonction où chaque prédicat primitif apparaît soit nié, soit non nié).	$Q_1 : P_1 \wedge P_2$ $Q_2 : P_1 \wedge \neg P_2$ $Q_3 : \neg P_1 \wedge P_2$ $Q_4 : \neg P_1 \wedge \neg P_2$
M	Propriété moléculaire (formée d'un nombre *w* de propriétés *Q*)	$P_1 \wedge (P_2 \vee \neg P_2) \rightarrow$ $Q_1 \vee Q_2$ $(w = 2)$

Tableau 2 : Propriétés d'un système de langage $\mathfrak{L}_N^\pi$.

On peut finalement fournir une description d'état (*state-description*) pour chacun des systèmes de langage qui consiste en une conjonction de chaque prédicat (ou sa négation) couplé avec chaque individu. Par exemple, pour un système contenant un prédicat et deux individus, les différentes descriptions d'état sont les suivantes :

[2] Carnap discute également des fonctions d'estimation qui ont les mêmes propriétés et qui donnent le même résultat que la fonction de confirmation correspondante pour un même paramètre λ. Nous nous restreindrons donc ici à discuter des fonctions de confirmation.

[3] Ces propriétés sont détaillées dans *Logical Foundations of Probability*. Nous n'en esquissons ici que l'essentiel qui est inclus dans *The Continuum of Inductive Methods*.

#	Description d'état
1	$Pa \wedge Pb$
2	$Pa \wedge \neg Pb$
3	$\neg Pa \wedge Pb$
4	$\neg Pa \wedge \neg Pb$

Tableau 3 : Exemple de description d'état pour $\mathfrak{L}_2^1$

Dans l'exemple de la section précédente, le tableau 1 représente un système $\mathfrak{L}_N^1$, soit un nombre quelconque d'individus possédant chacun une seule propriété (*B*, est blanc) ou sa négation.

Chaque fonction de confirmation doit prendre deux facteurs en compte. Le premier est le facteur empirique, soit le nombre d'individus pour lesquels on a observé une certaine propriété, divisé par le nombre total d'individus observés dans l'échantillon. En gardant la notation établie dans la section précédente, cela nous donne $\frac{B_n}{n}$.

Le second facteur est le facteur logique, soit celui qui correspond à la probabilité établie *a priori*. Ce facteur est $\frac{w}{k}$, soit le nombre de propriétés Q décrivant la conjonction des propriétés recherchées divisé par le nombre total de propriétés Q du système $\mathfrak{L}_N^\pi$. Dans l'exemple du tableau 2 ci-dessus, il y a 4 propriétés Q possibles ($k = 2^\pi$ donc $k = 2^2$). La propriété *M* donnée en exemple nous dit que l'on s'intéresse à une balle qui soit blanche, peu importe qu'elle soit lisse ou non. Deux propriétés Q décrivent une telle situation, soit Q_1 et Q_2. Nous avons donc $w = 2$. Dans ce cas, le facteur logique est donc égal à 2/4, soit 1/2.

Voyons maintenant la valeur de ces deux facteurs pour les données présentées dans le tableau 1. Prenons le cas où les balles tirées étaient respectivement « blanche, blanche, noire et blanche ». Quelle est la probabilité que la prochaine balle tirée soit blanche ? Le facteur empirique est de 3/4 (trois balles blanches sont sorties sur un total de quatre tirages) et le facteur logique est de 1/2 (il n'y a que deux propriétés *Q* possibles et seulement l'une d'elles contient le prédicat avec la valeur recherchée). Comment peut-on concilier les résultats de ces deux facteurs ? De plus, nous avons vu que nous pouvons calculer la probabilité du prochain tirage à l'aide de la règle de succession de Laplace. Or, ce calcul nous donne une probabilité de 4/6. Que devons-nous faire avec tous ces résultats ?

Selon Carnap, les facteurs empirique et logique sont les deux bornes à l'intérieur desquelles doit se situer le résultat de toute méthode de confirmation cohérente. Dans notre exemple, les résultats doivent donc se situer entre 0,5 et 0,75 inclusivement, ce qui est le cas de la règle de succession qui nous donne 0,67. Ce qui différencie les méthodes d'induction entre elles, c'est le poids qu'elles accordent à chacun des deux facteurs. En d'autres mots, les différentes méthodes inductives font une moyenne pondérée (u) entre les deux facteurs (u_1 et u_2) comme dans l'équation ci-dessous. Ce qui les différencie est le poids W qu'elles accordent à chacun des facteurs.

$$u = \frac{W_1 u_1 + W_2 u_2}{W_1 + W_2}$$

eq. 4) La probabilité inductive donnée par une méthode est le résultat d'une moyenne pondérée entre les facteurs empirique et logique.

Par convention, Carnap fixe le poids du facteur empirique au nombre total d'individus observés dans l'échantillon (n). Ce choix a l'avantage de rendre compte du fait que la fiabilité du facteur empirique augmente lorsque la taille de l'échantillon augmente, le poids qu'on désire lui donner augmente donc également. Par exemple, la valeur du facteur empirique risque plus de se rapprocher de la fréquence relative (au sens fréquentiste) des balles dans l'urne si on a tiré 500 balles que si on n'en a tiré que deux.

Le poids du facteur logique est quant à lui représenté par le symbole λ. La forme générale d'une méthode de confirmation sera donc la suivante.

$$u = \frac{n\left(\frac{B_n}{n}\right) + \lambda\left(\frac{w}{k}\right)}{n+\lambda} = \frac{B_n + \lambda\left(\frac{w}{k}\right)}{n+\lambda}$$

eq. 5) Facteurs empirique et logique pondérés par n et λ respectivement.

Les méthodes inductives peuvent donc être caractérisées à l'aide d'un seul paramètre λ qui donne le ratio pour chacun des facteurs. Celui-ci peut prendre n'importe quelle valeur parmi les nombres réels incluant 0 et ∞.

4. Quatre applications

Dans cette section, nous examinerons quatre cas de figure où λ prend chacune des deux valeurs extrêmes, 0 et ∞, ainsi qu'une valeur intermédiaire λ = 2. Finalement, nous verrons la fonction c^* développée par Carnap dont la valeur n'est pas indépendante de la valeur de k.

4.1. λ = 2 ou la règle de succession de Laplace modifiée

Pour pallier à certaines incohérences pouvant résulter de la règle de succession de Laplace, Carnap la modifie afin qu'elle prenne en compte le nombre de prédicats. Celle-ci correspond alors à la méthode λ = 2.

$$P\left(X_{n+1} = B \mid B_n\right) = \frac{B_n + \frac{2w}{k}}{n+2}$$

eq. 6) $c_2\left(h,e\right)$. Règle de succession de Laplace modifiée.

Le facteur logique pondéré $\lambda\left(\frac{w}{k}\right)$ peut être simplifié puisque la méthode de Laplace porte directement sur les prédicats primitifs pour lesquels $\frac{w}{k}=\frac{1}{2}$ puisque $w=\frac{k}{2}$. Cette simplification nous donne alors l'équation 3. Toutefois, telle que formulée à l'équation 6, cette méthode est plus générale et permet une application sur des propriétés moléculaires sans que cela ne cause les incohérences répertoriées dans la littérature.

4.2. λ = ∞ ou la toute-puissance du facteur logique

La méthode $\lambda=\infty$ correspond à la méthode 1 du tableau 1. En simplifiant la fonction, on retrouve le facteur logique, soit le nombre w de propriétés Q correspondant à la situation recherchée relativement au nombre k de propriétés Q du système $\mathfrak{L}_N^\pi$.

$$P\left(X_{n+1}=B \mid B_n\right)=\lim_{\lambda\to\infty}\frac{B_n+\lambda\left(\frac{w}{k}\right)}{n+\lambda}=\frac{\lambda\left(\frac{w}{k}\right)}{\lambda}=\frac{w}{k}$$

eq. 7) $c_\infty(h,e)$. Simplification de la fonction de confirmation pour $\lambda=\infty$.

Le facteur empirique est entièrement absorbé par la valeur λ = ∞. Cette méthode ne tient donc aucunement compte des informations empiriques obtenues. Comme nous l'avons vu, cette méthode est par le fait même insatisfaisante pour Carnap puisqu'elle ne respecte pas le principe selon lequel on doit apprendre de l'expérience.

4.3. λ = 0 ou la règle stricte

La méthode λ = 0 est celle qui n'accorde aucun poids au facteur logique. Elle tient compte du facteur empirique sans porter aucune attention aux probabilités qui pourraient être données *a priori* ou par le principe d'indifférence.

$$P\left(X_{n+1}=B \mid B_n\right)=\frac{B_n+\lambda\left(\frac{w}{k}\right)}{n+\lambda}=\frac{B_n}{n}$$

eq. 8) $c_0(h,e)$. Simplification de la fonction de confirmation pour $\lambda=0$.

Une complication apparente est celle où l'échantillon est vide ($n = 0$) où on a alors 0 au dénominateur. Dans ce cas, il est possible d'approcher l'équation par sa limite lorsque λ tend vers 0, ce qui nous donne $\lim_{\lambda\to 0} c_\lambda(h,e)=\frac{w}{k}$.

Pour de très petits échantillons ($n = 1$ par exemple), considérer uniquement le facteur empirique n'est probablement pas la méthode la plus raisonnable. Par exemple, si on tire une pièce à pile ou face et qu'on obtient le résultat pile au premier jet, selon une telle méthode de confirmation, le degré de confirmation de l'hypothèse « le prochain jet donnera le résultat pile » est alors de 1 ! Il en va de même pour un dé régulier qu'on aurait roulé cinq fois et dont le résultat « 6 » ne serait pas encore sorti : cette méthode de confirmation nous dit que l'hypothèse « le prochain jet donnera un '6' » a un degré de confirmation de 0. Comme le remarque Carnap, ce genre de résultat n'est pas satisfaisant puisque la probabilité d'observer une propriété qui n'est pas encore représentée dans l'échantillon « (…) may be low, but it cannot be regarded as 0 because it is, after all, a *possible case*, i.e., one *logically compatible* with the given evidence. » (Carnap 1952: 42, emphase ajoutée)

4.4. λ = k ou la fonction c* de Carnap

La fonction que Carnap privilégie est la fonction de confirmation *c** où la valeur de λ est déterminée par la valeur de *k*.

$$P\left(X_{n+1}=B\mid B_n\right)=\frac{B_n+\lambda\left(\frac{w}{k}\right)}{n+\lambda}=\frac{B_n+w}{n+k}$$

eq. 9) Fonction *c** de Carnap où $\lambda = k$

Cette fonction remplit les critères d'une bonne fonction de confirmation puisque le facteur logique est pris en compte, tout en accordant de plus en plus de poids au facteur empirique au fur et à mesure que l'échantillon grossit (Carnap 1945a).

De plus, le poids de la grosseur de l'échantillon est pris en compte sur une base relative par rapport au nombre de propriétés observables plutôt que de façon absolue comme c'est le cas dans la règle de succession de Laplace. On accorde le même poids au facteur empirique et au facteur logique lorsque $n = \lambda$.

Par exemple, dans la règle de succession de Laplace modifiée ($\lambda = 2$), les deux facteurs ont le même poids dès que l'échantillon compte deux individus, et ce, peu importe le nombre de propriétés du système. Par exemple, prenons un premier système qui ne contient qu'une seule propriété comme dans le tableau 1 (la couleur, blanche ou noire). Prenons un second système qui contient une dizaine de propriétés, par exemple la couleur, la texture, la présence de pois, la présence d'une craquelure, la présence d'une dépression, etc. Intuitivement, un échantillon de 5 individus nous semblera avoir plus de poids relativement au nombre de propriétés dans le premier système que dans le second. La règle de succession de Laplace modifiée ne nous permet pas de rendre compte de cette intuition alors que pour la règle *c** de Carnap, le facteur empirique aura le même poids que le facteur logique seulement à partir du moment où l'échantillon comptera le même nombre d'individus qu'il y a de propriétés Q possibles. Si $n < k$, le facteur logique sera prépondérant ; si $n > k$, le facteur empirique sera prépondérant. Dans notre exemple, dans le premier système, les deux facteurs auront une valeur pondérale égale après un échantillon de $n = 2$ individus alors que dans le second, leur valeur pondérale sera égale après un échantillon de $n = 2^{10} = 1024$ individus.

Conclusion

En plus d'avoir mis le doigt sur la cause des ambiguïtés entourant le concept de « probabilités », Carnap a construit un système rendant compte des différentes méthodes inductives couramment utilisées et il met en évidence d'autres méthodes inductives possibles, celles-ci formant un continuum.

Toutefois, son système repose sur la logique bivalente où une propriété ne peut prendre que la valeur de « vrai » ou « faux ». Par exemple, une balle pigée dans une urne sera « blanche » ou « non-blanche ». On peut s'en sortir si une seule valeur nous intéresse : il est possible de parler d'un dé si les résultats qui nous intéressent sont « as » et « non-as », quoi qu'ici le facteur logique accorde une probabilité de 1/2 à ces deux valeurs.

Une solution serait de créer une propriété pour chaque valeur possible, par exemple « blanche », « noire », « bleue », « rouge », etc. Toutefois, Carnap précise que les propriétés d'un système doivent être indépendantes : piger une balle rouge ne doit pas impliquer qu'elle sera non-bleue. Il est possible de déformer un peu son système si on ne regarde qu'une seule propriété primitive, mais les choses se gâtent si on regarde deux propriétés (par exemple, la couleur et la texture) ou pour les cas où des propriétés auraient un nombre de valeurs différentes, par exemple la propriété couleur (5 valeurs possibles) et la propriété texture (2 valeurs possibles).

La condition d'indépendance entre les différentes propriétés rend le système très peu flexible et incapable de rendre compte de plusieurs propriétés qui pourraient intéresser l'agent épistémique. Celui-ci pourrait être intéressé à évaluer la probabilité de propositions du type « si la prochaine balle tirée est rouge, alors elle sera lisse », ce qui exprime la dépendance entre deux propriétés.

Références

BERNOULLI, J., 1713. "Ars conjectandi", Basel.

CARNAP, R., 1945a. "On inductive logic". *Philosophy of Science* **12**(2), pp. 72-97.

— 1945b. "The Two Concepts of Probability: The Problem of Probability". *Philosophy and Phenomenological Research* **5**(4), pp. 513-532.

— 1950. *Logical foundations of probability*. Chicago, The University of Chicago Press.

— 1952. *The Continuum of Inductive Methods*. Chicago, University of Chicago Press.

— 1955. "Statistical and inductive probability". *The Galois Institute of Mathematics and Art*.

— 1962. *Logical foundations of probability*, University of Chicago Press Chicago.

— 1973. "Notes on probability and induction". *Synthese* **25**(3), pp. 269-298.

FISHER, R., 1922. "On the mathematical foundations of theoretical statistics". *Philosophical Transactions of the Royal Society of London* Series A, vol. 222, pp. 309-368.

JEFFREYS, S., 1939 [1998]. *Theory of probability*, Oxford University Press, USA.

KEYNES, J., 1921. *A treatise on probability*.

LAPLACE, P., 1814. *Théorie analytique des probabilités*. Paris : Mme. Ve. Courcier.

REICHENBACH, H., 1935. *The theory of probability*.

VON MISES, R., 1928 [1957]. *Probability, Statistics and Truth*. New York : Dover Publications.