Le calcul des probabilités selon Poincaré

(*La Science et l'Hypothèse*, pp.191-213)

Jamil Alioui

décembre 2011

Introduction

La thématique probabiliste

Henri Poincaré (1854-1912) – grand mathématicien, physicien et philosophe français – s'intéresse dans ce chapitre, tiré de *la science et l'hypothèse*, au calcul des probabilités. Il pose un certain nombre de questions intéressantes et profondes, tant philosophiques que mathématiques et physiques.

On verra notamment que l'idée même de la possible existence d'un « calcul des probabilités » n'est pas du tout quelque chose d'évident.

Poincaré s'interroge ensuite de manière épistémologique. Ses préoccupations sont essentiellement liées aux questions que pose le problème philosophique de la « loi vraie ». Il avance notamment que derrière l'induction, pratiquée quotidiennement par les physiciens et les mathématiciens, se dissimule systématiquement un calcul des probabilités.

Organisation de l'exposé

Afin de rester le plus structuré possible vis-à-vis d'un texte dense et déjà très synthétique, je respecterai dans ce commentaire l'ordre établi explicitement par Poincaré dans le chapitre, à savoir un commentaire d'introduction puis six parties – sur la classification des problèmes de probabilité, les mathématiques, la physique, les jeux de hasard, la probabilité des causes et la théorie des erreurs – puis une lecture des conclusions proposées par l'auteur.

J'apporterai, pour terminer, mon commentaire ; une conclusion personnelle sur l'ensemble des questions soulevées.

Des données réelles au modèle théorique

Le propos de Poincaré dans ce chapitre peut se résumer à la question de savoir comment, essentiellement dans la physique et les mathématiques, l'on passe de ce qu'il appelle « probabilité subjective », une intuition, à une « probabilité objective », une loi ; et de savoir ensuite quelle est alors la légitimité d'une telle loi. Le passage de l'intuition, issue de notre psychologie propre, de notre manière d'appréhender le monde, vers une « loi vraie », un modèle qui décrive en lui-même tout un ensemble de phénomènes y compris des phénomènes qui n'existent pas réellement (i.e. qui n'ont pas encore été observés), est *la* question qui anime Poincaré tout au long de ce passage. Comme nous le verrons, cette « objectivation », cette modélisation – que l'auteur résumera essentiellement au passage du paradigme discret, représenté par le signe [1] $\frac{1}{n}\sum^{n}$ relativement à un certain nombre n d'éléments observés réels discrets, vers le paradigme continu représenté par une intégrale $\int_{\mathbb{R}}$ – n'est pas quelque chose d'évident, et est un processus qui met en jeu un certain nombre de principes que nous aborderons et développerons.

Néanmoins, pour illustrer brièvement cette analogie entre discret-continu d'un côté et subjectif-objectif de l'autre, j'utilise généralement l'exemple suivant : si je lâche par terre un certain nombre n d'objets en verre, et qu'un grand nombre d'entre eux se brise en arrivant sur le sol, j'en déduirai subjectivement une loi qui dira « tous les objets en verre que je lâche sur le sol se brisent », eu égard à quelques exceptions du type « objet en verre très solide ». Mais je serai alors limité à une probabilité subjective, elle-même nécessairement limitée par l'impossibilité pratique de briser tous les objets en verre afin d'être certain qu'effectivement, la majeure partie d'entre eux se brise. Si je souhaite dépasser la limite de cette intuition légitime, et obtenir ainsi un réel pouvoir explicatif objectif, je devrai élaborer une loi, sur la base de mes observations, qui explique plus que les n cas observés, c'est à dire qui explique l'ensemble $\mathbb{R}$ du possible.

Le passage du discret au continu (i.e. du subjectif à l'objectif, de $\sum$ à $\int$ ou encore de $n \in \mathbb{N}$ à $\mathbb{R}$), est donc la question centrale de ce chapitre, qui s'articule, comme nous l'avons dit, d'abord par une introduction, puis par l'abord « pratique » d'un certain nombre de problèmes propres et, pour clore, par une conclusion. Cette articulation peut aussi être présentée en d'autres termes, à savoir :

1. la formalisation du problème général de la création d'un modèle à partir du réel ;
2. l'illustration de quelques problèmes caractéristiques et de l'élaboration des modèles y relatifs ;
3. la question de la pertinence et de la légitimité du modèle créé ;
4. une conclusion.

1. Les signes mathématiques sont utilisés, dans ce paragraphe, pour leur valeur symbolique.

Nous verrons ainsi que, si le calcul des probabilités pose un certain nombre de problèmes, il demeure un moyen clairement efficace et justifiable de répondre à un grand nombre de questions très variées, pour autant que le modèle utilisé soit adéquatement fabriqué, la question persistant étant « qu'est-ce qu'un modèle *adéquatement* fabriqué ? ».

Nous ferons accessoirement quelques parallèles avec la théorie statistique contemporaine qui crée, justement, des modélisations à partir d'observations, mais qui possède aussi les outils (i.e. les « tests ») qui permettent de valider ou de rejeter une hypothèse théorique en fonction des données réelles et d'une marge d'erreur, de tolérance, qu'il convient au statisticien de fixer.

1 Le *calcul* des *probabilités*

> *On s'étonnera sans doute de trouver à cette place des réflexions sur le calcul des probabilités. Qu'a-t-il à faire avec la méthode des sciences physiques ?* (p.191)

C'est toute la question ! Rappelons pour commencer que *la science et l'hypothèse* est un texte destiné au grand public, qui vise à faire le bilan de la situation des sciences actuelles ; c'est à dire en 1902, date de publication de l'ouvrage. Il est important, tout au fil de la lecture de ce texte, de bien contextualiser les éléments théoriques en jeu.

Le calcul des probabilités est une science relativement récente. Si la notion actuelle de « probabilité » est certainement née avec les travaux notamment de Pascal [2] (1623-1662) sur les jeux de hasard, il faudra attendre le XIXe siècle pour que la théorie des probabilités en mathématiques prenne la forme qui la fera converger par la suite vers la théorie actuelle, notamment grâce aux travaux démographiques de Adolphe Quetelet (~1840) ou à la physique statistique de J.-C. Maxwell (~1860).

Disons que si Planck avait introduit la notion de quanta en 1899, le paradigme quantique (i.e. l'inscription des probabilités *ontologiquement* au cœur de la physique) n'existe pas encore en 1902. Il faudra attendre des travaux tels que ceux de Heisenberg ou Schrödinger (premier tiers du XXe siècle) qui lieront ontologiquement les probabilités et la physique. Einstein et ses études sur le mouvement brownien de 1905 sont encore clairement inscrits dans le paradigme des probabilités épistémologiques. Rien n'empêche néanmoins d'indiquer qu'il existe deux types de liens entre la physique et les probabilités :

1. un lien épistémologique, lié notamment à la question de la « loi vraie » ;
2. un lien ontologique, explicité tardivement par la physique quantique.

2. Travaux dont on retrouve la trace dans sa correspondance de 1654 avec Pierre de Fermat.

Le problème épistémologique de la « loi vraie »

Dans ce chapitre il est clairement question du premier type de lien entre physique et probabilités, le lien épistémologique. D'ailleurs, pour justifier l'intérêt d'un chapitre dédié au calcul des probabilités dans un livre essentiellement dédié à un bilan des sciences, Poincaré nous dit :

> *Toutes les fois [que le physicien] raisonne par induction, il fait plus ou moins consciemment usage du calcul des probabilités.* (p.191)

Pour illustrer très simplement ce qu'on entend par *raisonnement par induction* disons qu'il s'agit du raisonnement qui entre en jeu lorsque je me mets à croire, après avoir cassé une dizaine de verres en les lâchant par terre, que tout objet lâché tombe. Lorsque Poincaré nous parle du degré de conscience lié à ce raisonnement, il faut comprendre qu'un tel raisonnement, avant d'avoir pu être théorisé par les logiciens et les philosophes, est une nécessité liée au fonctionnement du cerveau humain. Comment envisager un monde où il me faudrait lâcher par terre *tous les objets de l'univers* pour me convaincre que tout objet lâché tombe ? Un tel monde n'est pas imaginable pour la simple raison que le fonctionnement du cerveau ne permet pas la concrétisation d'une telle expérience de pensée.

Les problèmes du calcul des probabilités

> *Le nom seul de calcul des probabilités est un paradoxe : la probabilité, opposée à la certitude, c'est ce qu'on ne sait pas, et comment peut-on calculer ce que l'on ne connait pas ?* (p.192)

Un peu plus loin, Poincaré insiste sur *l'utilité fonctionnelle* de ce calcul pour les physiciens, et montre que la définition d'une probabilité, ou la formalisation d'un problème probabiliste, n'est pas une chose évidente et facile à élaborer. Pour amener ces problèmes, il commence par prendre l'exemple de deux dés A et B que l'on lance. Le tableau des *cas possibles* est le suivant, si l'on place les résultats possibles du dé A en lignes et les résultats possibles du dé B en colonnes :

$$\begin{array}{cccccc}
(1;1) & (1;2) & (1;3) & (1;4) & (1;5) & (1;6) \\
(2;1) & (2;2) & (2;3) & (2;4) & (2;5) & (2;6) \\
(3;1) & (3;2) & (3;3) & (3;4) & (3;5) & (3;6) \\
(4;1) & (4;2) & (4;3) & (4;4) & (4;5) & (4;6) \\
(5;1) & (5;2) & (5;3) & (5;4) & (5;5) & (5;6) \\
(6;1) & (6;2) & (6;3) & (6;4) & (6;5) & (6;6)
\end{array}$$

On comprend ainsi pourquoi la probabilité d'obtenir au moins un 6 est de $\frac{11}{36} \approx 30\%$: il s'agit du nombre de cases qui contiennent au moins un 6 divisé par le nombre de cases total.

Afin de poser la question de la légitimité de la formalisation d'un problème probabiliste par rapport à une autre formalisation, Poincaré propose de considérer cette fois non plus le problème comme la recherche de la probabilité d'obtenir au moins un symbole « 6 » parmi les symboles « 1 », « 2 », « 3 », ..., « 6 » de chaque dé, mais plutôt la recherche du nombre de combinaisons favorables (i.e. qui contiennent au moins un 6) parmi les combinaisons possibles avec deux dés. Ceci nous donne évidemment un autre tableau qui n'est plus le *tableau de tous les possibles*, mais le *tableau des* $\frac{6\times7}{2} = 21$ *combinaisons uniques possibles avec deux jeux de 6 symboles* :

$$\begin{array}{cccccc} (1;1) & (1;2) & (1;3) & (1;4) & (1;5) & (1;6) \\ & (2;2) & (2;3) & (2;4) & (2;5) & (2;6) \\ & & (3;3) & (3;4) & (3;5) & (3;6) \\ & & & (4;4) & (4;5) & (4;6) \\ & & & & (5;5) & (5;6) \\ & & & & & (6;6) \end{array}$$

On a alors 6 combinaisons favorables (la dernière colonne) et la probabilité pourrait être calculée comme $\frac{6}{21} \approx 28\%$.

On comprend alors que pour légitimer la première approche (qui est la bonne approche) par rapport à la deuxième, comme le dit Poincaré, il nous faut définir le probable par le probable selon la définition qu'il propose :

> *La probabilité d'un évènement est le rapport du nombre de cas favorables à cet évènement au nombre total des cas possibles [...], pourvu que ces cas soient également probables.* (p.192)

Cela complique donc en effet les choses.

Les diverses solutions qu'il propose ne semblent alors pas arranger l'affaire. Établir une convention, c'est à dire par exemple – comme dans le problème des dés – établir que les cas possibles doivent être également probables, pose le problème de la nécessité d'établir une convention à priori : pour tous les problèmes de probabilités possibles il faudrait trouver au moins une convention de formalisation, et pouvoir logiquement défendre sa légitimité, comme on vient de défendre celle du premier formalisme de la question des dés.

La « bonne » formalisation

Le « bon sens » lui aussi possède ses limites, comme en témoigne le problème de la corde plus grande que le côté du triangle équilatéral inscrit de la p.193. Ce problème est posé par Joseph Bertrand (1822-1900) – mathématicien, économiste et historien des sciences français – dans son traité *calcul des probabilités* de 1889. Rien ne valant un croquis en géométrie, la figure 1 permet de comprendre immé-

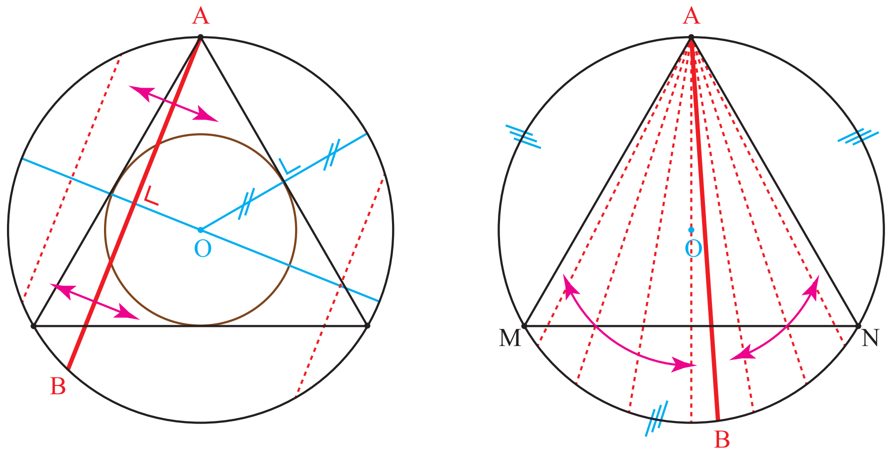

FIGURE 1 – Illustration du problème de Joseph Bertrand.

La probabilité de tracer une corde (ici $\overline{AB}$) « au hasard » plus longue que le côté du triangle équilatéral inscrit est respectivement et paradoxalement de $\frac{1}{2}$ et de $\frac{1}{3}$ selon la compréhension que l'on a de l'expression « au hasard ».

diatement l'objet de ce problème. En appliquant le « bon sens » dans le but de calculer la probabilité qu'une corde dessinée « au hasard » dans un cercle soit plus longue que le côté du triangle équilatéral inscrit, on trouve deux réponses différentes, à savoir $\frac{1}{2}$ et $\frac{1}{3}$. D'autres démonstrations montrent qu'il est encore possible d'obtenir des probabilités de $\frac{1}{4}$ et même de $1-\left(\sqrt{3}/2\right) \approx 0.134$.

Le problème se situe dans la consigne, autrement dit précisément dans la convention de formalisation du problème : qu'est-ce que « au hasard » ? On perçoit ici la limite du paradigme probabiliste où des conceptions différentes d'un geste effectué « au hasard » engendrent des résultats de calcul (i.e. des probabilités) différentes. Un énoncé trop vague ou mal formulé peut avoir des conséquences graves (Bertrand aurait pu, d'ailleurs, se contenter de l'une ou l'autre de ces démonstrations sans même réaliser le problème).

On est alors en droit de conclure que le calcul des probabilités est « une science vaine ». Mais comme nous le rappelle Poincaré au bas de la p.193, c'est le calcul des probabilités « épistémologique » qui nous permet d'établir la validité des lois telles que la loi de Newton suivant que le fait que *cette loi est fausse* est *peu probable*.

Malgré les ennuis que pose la formalisation d'un problème de probabilité, « condamner ce calcul, ce serait condamner la science tout entière » (p.194).

Le calcul des probabilités pour sa valeur pratique

Une autre solution serait d'admettre l'ensemble des sciences et du calcul probabiliste pour sa valeur pratique. Poincaré nous propose :

> *Ce que je sais, ce n'est pas que telle chose est vraie, mais que le mieux pour moi est encore d'agir comme si elle était vraie.* (p.195)

Il s'avère que ce n'est pas une solution. Pour expliciter cela, Poincaré pose deux définitions :

La probabilité subjective est, par exemple, un conseil que je pourrais donner à quelqu'un. Ce conseil étant basé sur mon expérience personnelle, je fais bel et bien recours au calcul des probabilités si je table sur la validité de mon expérience pour la personne qui me demande conseil. Mais je ne lui garantis ainsi pas le succès. Cette probabilité prend forme a priori.

La probabilité objective est, par exemple, le listing de tous les coups réussis et ratés sur une longue période. Un tel listing permet de « [constater] que les évènements se sont répartis conformément aux lois du calcul des probabilités » (p.195). Cette probabilité prend forme a postériori.

Comme le dit Poincaré, le phénomène particulièrement problématique est le passage de la première probabilité à la seconde, qui seule, garantit le succès du calcul. C'est pourquoi considérer le calcul des probabilités juste pour sa valeur pratique est réducteur et dangereux, car si l'on se situe dans la probabilité subjective, la réussite n'est pas assurée. La suite de la discussion essaie de cerner ce problème au travers d'exemples tirés des mathématiques et de la physique, dans une approche que l'auteur qualifie lui-même de « méfiante » mais qui ne veut pas – et ne peut pas – « condamner en bloc ».

2 Les probabilités et l'ignorance

> *Si nous n'étions pas ignorants, il n'y aurait pas de probabilité, il n'y aurait de place que pour la certitude ; mais notre ignorance ne peut être absolue, sans quoi il n'y aurait pas non plus de probabilité, puisqu'il faut encore un peu de lumière pour parvenir même à cette science incertaine.* (p.196)

Il y a donc un lien étroit entre ignorance, connaissance, certitude et incertitude, autrement dit c'est notre ignorance qui cause la nécessité d'un calcul de probabilité et elle est en lien avec lui, ce lien est précisément l'objet de la discussion.

Pour éclaircir le terrain, Poincaré propose dans un premier temps de classifier la généralité – qu'il s'agit de comprendre ici dans le sens de la généralisation, de l'induction – en trois catégories. Il propose dans un deuxième temps de faire la même chose avec l'ignorance.

Ensuite, il traite spécifiquement de ce qu'il nomme *probabilité des causes*. Il s'agit en fait d'une discussion autour de l'induction en elle-même, c'est à dire le passage des cas particuliers connus à une loi générale, un modèle.

Les trois degrés de généralité

Le premier degré de généralité est lié à un nombre fini de cas possibles. C'est le cas du lancer de dés. Les 6 symboles de notre dé sont finis, et quel que soit le symbole sortant, il sera compris entre 1 et 6.

Le second degré de généralité est lié à un nombre infini de cas possibles. C'est le cas du problème posé par Joseph Bertrand où le nombre de cordes possible dans un cercle est infini.

Le troisième degré de généralité est lié à un nombre fini d'observations où l'on ne connait pas le nombre total de cas possibles car il croît en même temps que le nombre d'observations. La partie sur les probabilités en mathématiques, plus bas, explique bien ce type de généralité.

Les trois degrés de l'ignorance

Le premier degré de l'ignorance est l'ignorance que l'on a lorsqu'on lance un dé. On ne connait pas k, le nombre de points du dé sur la face sortante, mais on sait que $1 \leqslant k \leqslant 6$.

Le second degré de l'ignorance est l'ignorance des physiciens lorsqu'ils travaillent, par exemple, uniquement sur la température moyenne d'un gaz. L'ignorance se situe en ceci que l'on ne connait pas (et que l'on ne peut pas connaitre) l'ensemble des valeurs des températures des molécules du gaz T_1, T_2, T_3, ..., T_n pour n molécules, mais on en sait quand même assez pour pouvoir modéliser le système et travailler avec grâce à une température moyenne $\overline{T}$.
Plus généralement, il s'agit de systèmes dont on connait les lois mais dont on ignore les conditions initiales, comme le dit Poincaré :

> *Dans la théorie cinétique des gaz, on suppose que les molécules gazeuses suivent des trajectoires rectilignes et obéissent aux lois du choc des corps élastiques ; mais, comme on ne sait rien de leurs vitesses initiales, on ne sait rien de leurs vitesses actuelles.* (p.197)

Le troisième degré de l'ignorance est l'ignorance totale et l'incapacité de modéliser quoi que ce soit.

La probabilité des causes

Ce qui est nommé ici probabilité des causes correspond essentiellement au domaine d'investigation de l'actuelle statistique. Nous verrons d'ailleurs qu'un parallèle encore plus fort peut être fait entre ce que Poincaré nomme « théorie des erreurs » et les tests statistiques, qui permettent à un résultat d'observation (par exemple une corrélation entre la taille des mains et la taille des pieds d'un groupe d'individus) de passer ou non au stade de « loi générale ». Nous reviendrons là dessus dans la partie consacrée à la théorie des erreurs.

Lorsque je compile un fichier de données – par exemple un fichier qui contient des données du genre taille, poids, couleur des yeux, couleur des cheveux pour un grand nombre d'individus – je peux calculer des corrélations, des covariances, etc. entre ces données. Mais ça n'est pas parce que, dans un fichier de 100 personnes, je trouve une forte corrélation entre – disons – la couleur des yeux et la couleur des cheveux, que je suis en droit d'induire une loi générale qui stipulerait, pour l'ensemble des humains sur la Terre, que la couleur des yeux et la couleur des cheveux sont corrélées. De même, si une banale régression me permet de « prévoir » la valeur d'une variable y^* encore inconnue, en fonction d'un ensemble parfois très grand de variables x_1, x_2, ..., x_n appelées *prédicateurs*, le fait d'affirmer que la variable y^* vaudra telle valeur n'est pas légitimé aussi facilement.

Au stade des observations, une corrélation ou une prédiction ne sont que des *hypothèses* qu'il s'agira de « tester » avec des méthodes *appropriées* pour connaitre la part de corrélation « expliquée », ou celle de pertinence d'une variable prédite, par rapport à la part de « bruit » ; lui-même alors justifiable – par exemple – par un trop faible nombre d'observations, ou par des observations trop localisées, imprécises ou inadéquates (par exemple, si je fais mes observations dans un pays où le 90% de la population a les yeux bruns et les cheveux noirs, mes observations peuvent a priori être considérées comme étant faussées). C'est, en quelque sorte, la question de savoir comment une méthode peut être *appropriée* ou pas qui est l'objet du propos qui vient. Poincaré résume cette question en une phrase :

> *J'ai observé n valeurs de x et les valeurs correspondantes de y ; j'ai constaté que le rapport des secondes aux premières est sensiblement constant. Voilà l'évènement ; quelle est la cause ?* (p.198)

Aujourd'hui on dirait : « voilà les données, l'hypothèse que, de manière générale, y dépend de x est-elle légitime ? ».

3 Probabilités et mathématiques

Poincaré propose d'insister sur la différence et le passage qui existent entre probabilités subjectives et probabilités objectives. Pour ce faire, il présente trois problèmes mathématiques qui s'inscrivent dans le cadre du troisième degré de généralité, c'est à dire celui où le nombre de cas possibles est égal au nombre d'observations (à ne pas confondre avec le second degré où le nombre de cas possibles est simplement infini). Dans ce type de généralité, il est nécessaire d'invoquer certains principes philosophiques et mathématiques pour pouvoir objectiver une probabilité intuitionnée.

Le principe de raison suffisante

Le premier exemple permet de comprendre l'insuffisance explicative du *principe de raison suffisante*. Ce principe est formulé originellement par Leibniz (1646-1716). Rappelons-le ici :

> *Jamais rien n'arrive sans qu'il y ait une cause ou du moins une raison déterminante, c'est-à-dire qui puisse servir à rendre raison a priori pourquoi cela est existant plutôt que non existant et pourquoi cela est ainsi plutôt que de toute autre façon.* (Leibniz, Théodicée, I, 44)

La discussion tourne ici autour du célèbre problème de la quadrature du cercle qui consiste à construire un carré de même aire qu'un cercle donné, à la règle et au compas. C'est un problème impossible à résoudre avec un nombre fini d'opérations à cause de la nature transcendante de π. Sachant qu'avant 1883 il n'avait pas été prouvé que ce problème était insoluble, la question philosophique est ici : « était-il légitime de la part de l'académie – et avant 1883 – de refuser d'entrer en matière *a priori* sur ce sujet ? Et si oui sur quelle(s) base(s) ? ». Poincaré se met à la place des gens de l'académie et, suivant le principe de raison suffisante, propose une réponse :

> *« Pourquoi voulez-vous qu'une valeur particulière d'une fonction transcendante soit un nombre algébrique ; et si π était racine d'une équation algébrique, pourquoi voulez-vous que cette racine soit une période de la fonction* $sin(2x)$ *et qu'il n'en soit pas de même des autres racines de cette même équation ? »* (p.199)

Ce qu'il faut noter dans le principe de Leibniz, c'est l'aspect *a priori* de l'explication. Dans le cas de la quadrature du cercle, l'aspect a priori n'était pas défini et le fait qu'aucune règle n'existait pour légitimer a priori une investigation telle que celle-ci était, selon Poincaré, une raison suffisante pour s'éviter la peine de rentrer en discussion. De manière générale, on peut dire que le principe de raison suffisante est fortement lié à la croyance que les scientifiques ont généralement à la simplicité de l'explication. On retrouve une idée identique dans le fameux Rasoir d'Ockham qui stipule que « les multiples ne doivent pas être utilisés sans nécessité ».

Ce genre d'argument n'est pas forcément conscient et, comme nous le laisse présupposer Poincaré au bas de la p.198, peut se dissimuler derrière des raisons qu'il qualifie de *psychologiques*. Retenons simplement que *raison suffisante* n'implique pas *probabilité objective*, mais y contribue par le biais de la *probabilité subjective*, c'est ce que nous montrent aussi les exemples suivants.

L'intuition

Le second exemple consiste à tirer au sort un logarithme parmi les dix mille premiers et à poser la question suivante : « quelle est la probabilité pour que sa troisième décimale soit un nombre pair ? ». La probabilité subjective, c'est à dire l'observation rapide dans une table – ou de manière contemporaine l'écriture d'un petit logiciel qui fonctionnerait de manière itérative ($+1$ si la troisième décimale est paire et -1 sinon, avec affichage du total en fonction du nombre de logarithmes évalués) – nous montre que $\frac{1}{2}$ semble être la bonne réponse.

Mais comme le dit Poincaré, il existe des théorèmes mathématiques complexes, liés aux dérivées des fonctions, qui permettent de prouver *de manière objective* qu'il est juste de penser que la probabilité de tirer un logarithme dont la troisième décimale est un nombre pair est bel et bien de $\frac{1}{2}$.

> *Et d'ailleurs, comme ce que j'appelais mon intuition n'était qu'un aperçu incomplet d'un véritable raisonnement, on s'explique que l'observation ait confirmé mes prévisions, que la probabilité objective ait été d'accord avec la probabilité subjective.* (p.200)

Cet exemple n'est pas anodin, il nous éclaire sur l'intuition qui serait un raisonnement extrêmement rapide, élaboré sur la base de connaissances diverses, comme par exemple les propriétés de dérivabilité communes aux fonctions continues. On comprend de mieux en mieux le lien entre ces deux types de probabilités, à savoir subjectives et objectives : en jouant de principes mathématiques on arrive à une conclusion qui, dans sa forme, englobe d'un coup tout un ensemble de cas possibles. Cet ensemble est défini par les propriétés que les fonctions en question ont en commun.

La convention

Le troisième exemple est plus compliqué en ceci qu'il met en jeu des notions d'analyse spécifiques et relativement avancées qu'il n'est pas forcément utile de développer en détail ici. Il est plus intéressant de directement se concentrer sur le propos qui est derrière cette démonstration.

La question est la suivante : on aimerait connaitre la valeur probable du sinus d'un produit de deux nombres dont le premier, u, est pris au hasard, et le second, n, est un entier donné très grand. Poincaré nous dit que « ce problème n'a aucun sens par lui-même » et que « pour lui en donner un, il faut une convention », il *faut* convenir que u a une certaine probabilité d'être dans un certain intervalle, sans quoi il est impossible de faire quoi que ce soit.

Ensuite, en usant de certaines propriétés communes aux fonctions continues, en jouant de la périodicité du sinus, et en effectuant un peu de calcul différentiel, on arrive à déduire que $\lim_{n\to\infty} \sin(nu) = 0$.

> *[...] pour définir cette valeur, j'ai eu besoin d'une convention ; mais le résultat reste le même* quelle que soit cette convention. (p.201)

Il est intéressant de voir que le calcul d'une probabilité objective nécessite ici une formalisation – pas forcément celle-ci, mais il en faut une – pour pouvoir exister, mais que cette formalisation, cette « convention », émerge dans l'esprit du mathématicien comme une probabilité subjective, une intuition fondée, pour le coup, sur la connaissance de propriétés communes aux fonctions continues.

A nouveau, on cerne mieux comment l'on passe d'une probabilité subjective à une probabilité objective, et comment un tel passage est possible.

4 Probabilités et physique

Les cas de probabilités physiques problématiques font partie, selon Poincaré, du deuxième degré d'ignorance, c'est à dire celui où l'on admet certains principes comme étant vrais sans connaitre les conditions initiales du système en question. Un seul exemple anime cette partie et illustre bien la nature des diverses problématiques abordées, il s'agit de la question de la distribution actuelle probable des petites planètes sur le zodiaque.

Poincaré commence son exposé en indiquant qu'il n'y a aucune différence de principe entre la considération d'orbites suivant les lois de Kepler et la considération d'orbites circulaires situées dans un même plan.

> *Nous savons [que les planètes] obéissent aux lois de Kepler ; nous pouvons même, sans rien changer à la nature du problème, supposer que leurs orbites sont toutes circulaires et situées dans un même plan et que nous le sachions.* (p.202)

La question de départ est « pourquoi n'hésitons-nous pas à affirmer qu'aujourd'hui la distribution des petites planètes sur le zodiaque est à peu près uniforme ? ». On voit qu'il s'agit à nouveau de la question de la légitimité d'une croyance.

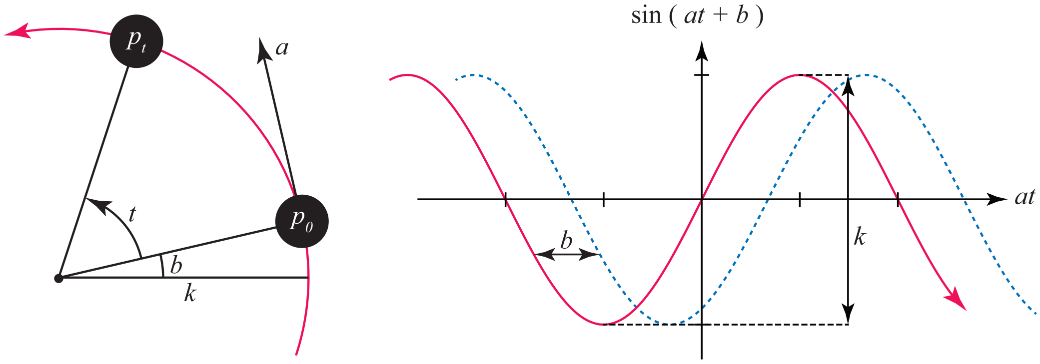

FIGURE 2 – Exemple de la distribution probable des planètes.

Pour comprendre le problème, voyons la figure 2.

> *Soit b la longitude d'une petite planète à l'époque initiale, c'est à dire à l'époque zéro ; soit a son moyen mouvement ; sa longitude à l'époque actuelle, c'est à dire à l'époque t, sera $at+b$. Dire que la distribution actuelle est uniforme, c'est dire que la valeur moyenne des sinus et des cosinus des multiples de $at+b$ est nulle. Pourquoi l'affirmons-nous ?* (p.202)

On peut dire que Poincaré pose la question de la légitimité du formalisme suivant, nécessaire pour pouvoir répondre à la question posée, pour n planètes avec n grand et $k \in \mathbb{R}$ étant un réel quelconque généralisant le principe de base de la figure 2 à l'ensemble des orbites planétaires possibles :

$$\begin{cases} \frac{1}{n}\sum_i^n k\cos\left(a_i t + b_i\right) & = 0 \\ \frac{1}{n}\sum_i^n k\sin\left(a_i t + b_i\right) & = 0 \end{cases}$$

Il faut se rappeler ici qu'il n'est pas question de discuter la pertinence de cet exemple-là particulièrement, mais de traiter *en général* de la formalisation d'un problème de probabilités. Voyons comment Poincaré s'y prend premièrement avec ce problème pour expliquer quelque chose qui sera ensuite généralisé.

La « transformation » du problème, sa formalisation, passe par la représentation de l'ensemble des planètes comme des points dans un plan cartésien dont les axes représenteraient, disons, les mouvements moyens (c'est à dire les vitesses angulaires a) en abscisses et les longitudes au temps t_0 en ordonnées (c'est à dire les valeurs b). Chaque planète devient alors un point $(a;b)$ sur le plan. On « repose » alors la question en ces termes :

> *Quelle est la probabilité pour qu'un ou plusieurs points représentatifs se trouvent dans telle partie du plan ?* (p.202)

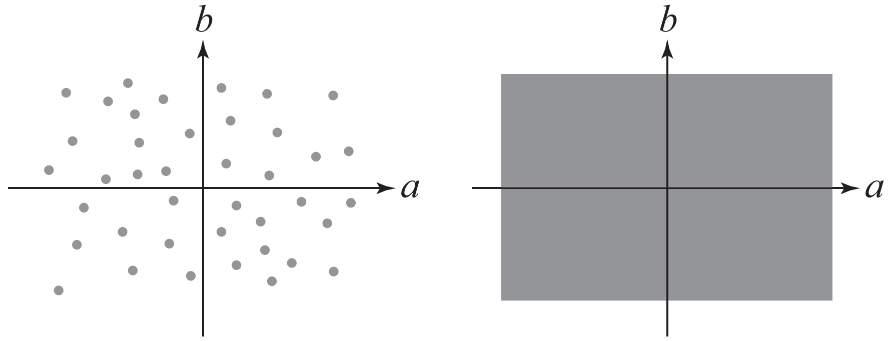

FIGURE 3 – Des données au modèle.

L'auteur nous rappelle alors que nous sommes ici dans le deuxième degré d'ignorance, degré que j'ai choisi de symboliser par la nécessité d'utiliser plutôt une moyenne des températures $\overline{T}$ qu'un immense ensemble de n températures T_1, T_2, ..., T_n correspondant aux n molécules, individuellement insondables, qui constitueraient un volume ou une masse hypothétiques. Poincaré préfère l'exemple de la densité, c'est à dire la distribution des molécules, d'une matière fictive répandue sur un plan ; exemple qu'il nous soumet alors et non sans dessein vis-à-vis du problème qui l'anime ici :

> *Nous conviendrons alors de dire que le nombre probable de points représentatifs qui se trouvent sur une partie du plan est proportionnel à la quantité de matière fictive qui s'y trouve. Si l'on a alors deux régions du plan de même étendue, les probabilités pour qu'un point représentatif de l'une de nos petites planètes se trouve dans l'une ou dans l'autre de ces régions seront entre elles comme les densités moyennes de la matière fictive dans l'une et l'autre région.* (pp.202-203)

On comprend alors, lorsqu'on regarde la figure 3, que nous sommes exactement dans un problème du même type que celui du test statistique, autrement dit du passage des données *réelles discrètes* au modèle *continu imaginaire*. Poincaré explique :

> *Voilà donc deux distributions, l'une réelle, où les points représentatifs sont très nombreux, très serrés, mais discrets comme les molécules de la matière dans l'hypothèse atomique ; l'autre, éloignée de la réalité, où nos points représentatifs sont remplacés par une matière fictive continue. Cette dernière, nous savons qu'elle ne peut être réelle, mais notre ignorance nous condamne à l'adopter.* (p.203)

Finalement, la seule supposition nécessaire à l'appréhension de ce problème est ici la supposition que la distribution des planètes est une fonction continue. Ainsi, Poincaré peut passer de la somme de planètes discrètes à l'intégration de cette surface continue que j'appelle volontiers « fluide planétaire » ! Pour ce faire il découpe la surface en n « pièces » de densité φ et de surface s.

Ce qui est extraordinaire ici, c'est la contingence de φ. Poincaré nous éclaire en disant :

> *[...] quelle que soit l'hypothèse faite, le résultat sera le même et c'est ce qui me tire d'embarras. [...] Je puis choisir φ comme je le veux, sauf une restriction toutefois : cette fonction doit être continue ; et, en effet, au point de vue de la probabilité subjective, le choix d'une fonction discontinue aurait été déraisonnable ; quelle raison pourrais-je avoir, par exemple, de supposer que la longitude initiale peut être égale à 0° juste, mais qu'elle ne peut être comprise entre 0° et 1° ?* (p.204)

On revient ici à la problématique de base, c'est à dire au passage de la probabilité subjective à la probabilité objective. Si l'on a objectivé la distribution probable des planètes en créant un modèle continu sur la base d'un nombre fini d'observations discrètes, on a fait une supposition subjective quant à l'utilisation d'une fonction de densité φ continue. On a présupposé subjectivement, en effet, que les planètes avaient des longitudes de départ comprises dans $\mathbb{R}$, autrement dit qu'elles étaient déjà réparties uniformément à t_0, comme le souligne Poincaré :

> *Si, par exemple, l'on avait pour toutes les petites planètes : $b = \frac{\pi}{2} - at$, toutes les planètes à l'instant t se trouveraient avoir pour longitude $\frac{\pi}{2}$ et la valeur moyenne serait évidemment égale à 1. Pour cela, il faudrait qu'à l'époque 0 les petites planètes eussent été toutes placées sur une sorte de spirale [...]. Tout le monde jugera qu'une pareille distribution initiale est extrêmement improbable. [...] Toutefois, pourquoi jugeons-nous cette distribution initiale improbable ?* (p.205)

A cette interrogation, Poincaré répond par un retour nécessaire au principe de raison suffisante :

> *Il n'y a pas de raison suffisante pour que la cause inconnue qui leur a donné naissance [aux planètes] ait agi suivant une courbe si régulière et pourtant si compliquée, et qui paraitrait précisément avoir été choisie exprès pour que la distribution actuelle ne fût pas uniforme.* (p.206)

Reprenons le cheminement du début. D'un problème probabiliste – à savoir la question de savoir si la distribution probable des planètes, dont les orbites sont considérées comme des cercles, au bout d'un certain temps t, est uniforme – on tire des observations que l'on n'arrive plus à quantifier si leur nombre devient très grand. On n'arrive donc plus à répondre à la question posée sans modéliser la

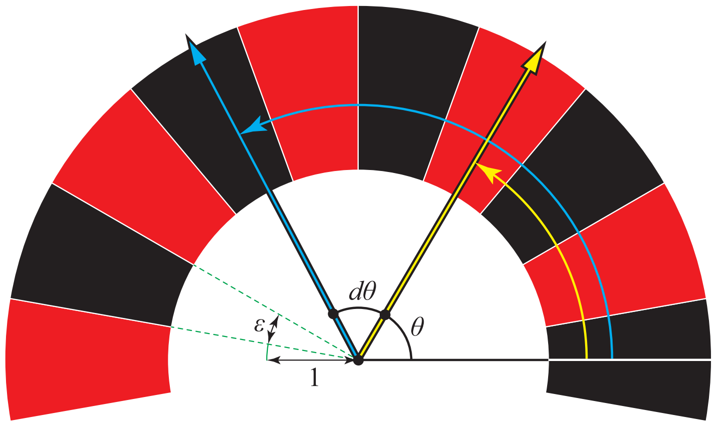

FIGURE 4 – Probabilités dans le jeu de la roulette.

situation. Afin de ce faire, on suppose la distribution continue et uniforme, et l'on justifie cette supposition par le principe de raison suffisante. Cette supposition nous permet de considérer alors la moyenne des positions des planètes pour un nombre de planètes n, tel que $n \to \infty$, et ainsi d'intégrer. On réalise, après cette opération mathématique, que le résultat ne décrit plus une situation réelle (à savoir n planètes « réelles » dont on connait les positions à t_0, les vitesses angulaires a et donc les positions actuelles) mais un modèle imaginaire qui, s'il nous permet de répondre à la question posée, a fait abstraction d'une partie de la réalité.

5 Probabilités et jeux de hasard

L'exemple est ici le jeu de la roulette que l'on trouve encore dans nos casinos aujourd'hui. On aimerait prouver que la probabilité d'arriver soit sur une division rouge soit sur une division noire vaut $\frac{1}{2}$. Pour ce faire, il faut connaitre la probabilité que l'angle de l'aiguille, après lancer, se situe entre θ et $\theta + d\theta$, autrement dit dans un intervalle compris entre les flèches jaune et bleue de la figure 4. Poincaré présuppose cette probabilité comme étant $\varphi(\theta)\,d\theta$, ce qui s'explique par le fait que si $\varphi(\theta)$ nous donne la probabilité d'arriver à l'angle θ après lancer, la multiplication de cette valeur par $d\theta$ nous donnera la probabilité d'arriver dans l'intervalle, probabilité forcément plus grande que celle d'arriver exactement en θ.

A nouveau, on considèrera la fonction $\varphi(\theta)$ comme étant continue, c'est à dire entièrement définie dans $\mathbb{R}$, ce qui fait sens si l'on imagine qu'il n'y a pas d'angle ϑ, tel que $\theta \leqslant \vartheta \leqslant \theta + d\theta$, en lequel l'aiguille n'aurait aucune chance d'arriver.

Ensuite, Poincaré définit ε tel que « la longueur (comptée sur la circonférence

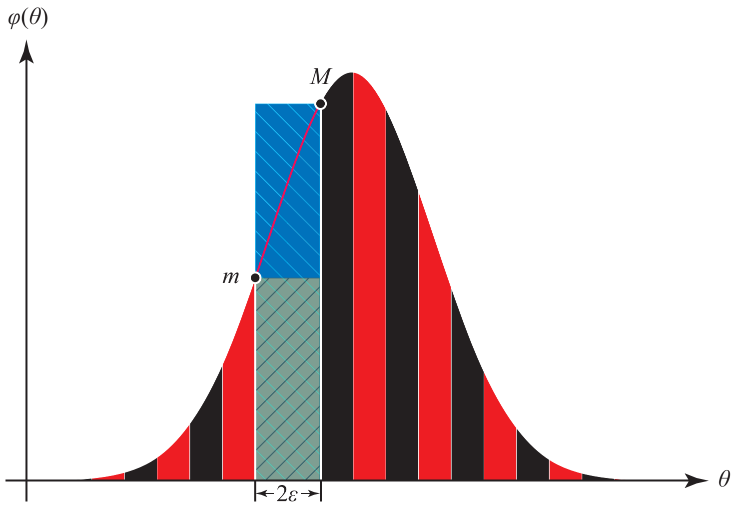

FIGURE 5 – Distribution de probabilité dans la roulette.

de rayon 1) de chaque subdivision rouge ou noire » (p.206). Cette longueur est aussi représentée sur la figure.

> *Il faut calculer l'intégrale de $\varphi(\theta)\,d\theta$ en l'étendant, d'une part, à toutes les divisions rouges, d'autre part à toutes les divisions noires, et comparer les résultats.* (p.206)

Les choses se corsent ensuite : Poincaré définit un certain nombres d'objets mathématiques que je résume ici, relativement à la figure 5 où φ est arbitrairement représentée par une courbe de Gauss. Il est clair que la distribution correcte aurait été une distribution uniforme, mais pour le bien de l'exposé et la compréhension de la contingence de φ, j'ai fait exprès de choisir une courbe de Gauss.

- 2ε contient une division rouge et une division noire ;
- M est la probabilité maximale que θ se situe dans l'intervalle 2ε ;
- m est la probabilité minimale que θ se situe dans l'intervalle 2ε ;
- $\left[\int \varphi(\theta)\,\mathrm{d}\theta\right]_{rouge} < \sum M\varepsilon$ pour les divisions rouges
  autrement dit, la probabilité de tomber sur une division rouge est inférieure à la surface bleue ;
- $\left[\int \varphi(\theta)\,\mathrm{d}\theta\right]_{noir} > \sum m\varepsilon$ pour les divisions noires
  autrement dit, la probabilité de tomber sur une division noire est supérieure à la surface verte ;
- Par corolaire : $\left[\int \varphi(\theta)\,\mathrm{d}\theta\right]_{rouge} - \left[\int \varphi(\theta)\,\mathrm{d}\theta\right]_{noir} < \sum (M-m)\,\varepsilon$ et pour $\varepsilon \to 0$ on obtient zéro et donc une probabilité de $\frac{1}{2}$:

> *Mais, si la fonction φ est supposée continue ; si, d'autre part, l'intervalle ε est très petit par rapport à l'angle total parcouru par l'ai-*

> *guille, la différence $M - m$ sera très petite. La différence des deux intégrales sera donc très petite, et la probabilité sera très voisine de* $\frac{1}{2}$. (p.207)

A nouveau on perçoit la nécessité d'établir une convention, c'est à dire dans ce cas de travailler avec une fonction φ continue. Mais on remarque aussi que cette fonction, dont on ne sait absolument rien, disparait dans le raisonnement.

Afin d'insister sur les répercussions possibles d'une mauvaise formalisation, ou d'une formalisation illégitime d'un problème probabiliste, Poincaré indique :

> *[...] mais [cette loi objective] entraine [les joueurs] dans une singulière erreur, qui a été souvent relevée, et dans laquelle ils retombent toujours. Quand la rouge est sortie, par exemple, six fois de suite, ils mettent sur la noire, croyant jouer à coup sûr ; parce que, disent-ils, il est bien rare que la rouge sorte sept fois de suite. En réalité, leur probabilité de gain reste* $\frac{1}{2}$. (p.207)

On peut expliquer ce fait simplement en rappelant deux choses :

1. ce problème de la roulette a été formalisé par rapport à un seul lancer d'aiguille dont l'angle de départ n'importe pas, car pour n'importe quel angle de départ, cette formalisation ne m'empêche aucunement d'effectuer une rotation pour ramener le système à quelque chose comme ce que l'on voit sur la figure 4 ;
2. la probabilité n'indique qu'une chose : pour un nombre de lancers n, tel que $n \to \infty$, il y aura 50% de divisions rouges et 50% de divisions noires. En aucun cas je ne peux obtenir une quelconque information supplémentaire sur la distribution de ces n lancers dans le temps avec cette formalisation.

6 La probabilité des causes

L'importance de la convention

Comme nous l'avons déjà dit plus haut, nous pouvons ramener ce que Poincaré nomme ici « probabilité des causes » à la statistique. Particulièrement dans ce chapitre, il est question d'une régression. Je résume les deux premiers exemples sans m'éterniser, car le troisième suffit, à mon avis, pour voir où l'auteur veut en venir.

1. Le premier exemple pose la situation et la question suivantes. Lorsque je vois deux étoiles proches l'une de l'autre dans le ciel, j'ai deux explications possibles : la première dit que « ces étoiles sont véritablement proches » et la seconde que « ces étoiles sont très éloignées mais que leur alignement

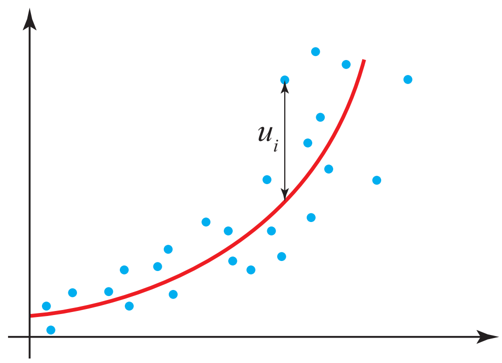

FIGURE 6 – Principe de régression.

En bleu, les observations ; en rouge, la loi prédite ; u_i, un résidu parmi d'autres.

particulier provoque un effet de trompe-l'œil qui me permet de croire à tort à leur proximité ». Quelle est la probabilité pour que la seconde explication soit la bonne ?

2. Le second exemple est celui du jeu de cartes : si je considère mon adversaire comme étant un tricheur avec une probabilité de $\frac{1}{2}$ à priori, s'il donne la première fois et tourne un roi, il y a alors une probabilité de $\frac{8}{9}$ pour qu'il soit effectivement un tricheur. Ce résultat surprenant est expliqué par Poincaré comme suit :

   Si on examine de plus près, on voit qu'on fait le calcul comme si, avant de nous assoir à la table de jeu, *j'avais considéré qu'il y avait une chance sur deux pour que mon adversaire ne fût pas honnête. Hypothèse absurde, puisque, dans ce cas, je n'aurais certainement pas joué avec lui ; et c'est ce qui explique l'absurdité de la conclusion.* (pp.208-209)

La détermination d'une loi expérimentale

Le troisième exemple est très intéressant. Poincaré décrit le processus, en p.209. En résumé cela donne quelque chose comme cela :

1. Je fais un certain nombre d'observations isolées.
2. Je représente ces observations par des points dans un plan.
3. Je fais passer une courbe entre ces points en m'efforçant de m'en écarter le moins possible et, cependant, de conserver à ma courbe une forme régulière.

Poincaré indique :

> *Cette courbe me représentera la loi probable, et j'admets, non seulement qu'elle me fait connaitre les valeurs de la fonction intermédiaires entre celles qui ont été observées, mais encore qu'elle me fait connaitre les valeurs observées elles-mêmes plus exactement que l'observation directe (c'est pour cela que je la fais passer près de mes points et non pas par ces points eux-mêmes).* (p.209)

Là où il devient possible de tirer un parallèle direct avec les statistiques telles qu'elles sont pratiquées aujourd'hui, c'est lorsqu'on prend connaissance de ce que Poincaré nomme clairement « le problème de la probabilité des causes », et qu'il décrit de cette manière : « les effets, ce sont les mesures que j'ai enregistrées ; ils dépendent de la combinaison de deux causes » :

1. la loi véritable du phénomène ;
2. les erreurs d'observation, représentées sur la figure 6 par la valeur u_i si l'on admet la courbe rouge comme étant la loi explicative. Ces valeurs u_i sont appelées communément en statistiques les « résidus » et prennent forme tels que $u_i = y^* - y$ avec y^* la valeur que me prédit ma loi et y la valeur effectivement observée.

En p.210, l'auteur nous propose des explications liées à ces résidus.

– La qualité des instruments de mesure est liée au dessin de la loi (i.e. la courbe rouge). Plus les instruments de mesure sont jugés bons, moins j'ai de liberté de « rehausser » ma fonction.
– Je peux légitimer le tracé d'une courbe (ou d'une droite dans une régression linéaire) car, à nouveau, je suppose qu'une loi représentable par une fonction continue (définie entièrement dans $\mathbb{R}$) est plus probable qu'une loi discontinue, c'est à dire une loi qui exclurait à priori certaines valeurs possibles.

Poincaré observe alors quelque chose d'essentiel :

> *Sans cette croyance [i.e. que la loi est représentable par une fonction continue], le problème dont nous parlons n'aurait aucun sens ; l'interpolation serait impossible ; on ne pourrait déduire une loi d'un nombre fini d'observations ; la science n'existerait pas.* (p.210)

A titre d'exemple, il cite la croyance favorable à la loi la plus simple face à une loi compliquée, croyance qu'il illustre par la confrontation entre la loi de Mariotte et les expériences de Regnault. La loi de Mariotte (~1669) affirme de manière simpliste que pour deux volumes de gaz sous pressions, soumis à la même température, $p_1 \times V_1 = p_2 \times V_2$. Les expériences de Regnault (~1850) ont montré notamment que les liens entre température, pression et volume, pour deux gaz, étaient plus complexes que cela. Mais si l'on a abandonné l'admission systématique de l'explication la plus simple, des « stigmates » persistent, comme le souligne Poincaré :

> *[...] ce qui reste de cette tendance, c'est la croyance à la continuité, et nous venons de voir que, si cette croyance disparaissait à son tour, la science expérimentale deviendrait impossible.* (p.210)

Nous identifions donc clairement, maintenant, ce qui fait le pont entre probabilité subjective et probabilité objective : c'est la croyance systématique en la « continuité » des lois. Mais en plus d'être un « principe » a priori, cette croyance est le fondement même de l'idée de loi de la nature. La question de savoir s'il est légitime d'objectiver une probabilité sur la base de cette croyance ne peut plus se poser de cette manière. Comme nous l'avons déjà dit, si les humains devaient laisser se casser par terre l'ensemble des objets de l'univers pour oser en conclure que tous les objets lâchés tombent et se cassent, on ne pourrait alors plus, en effet, parler de science. La boucle est bouclée. Reste cependant la question des résidus, qui fait l'objet de la partie suivante.

7 La théorie des erreurs

Poincaré commence par poser clairement le problème lié aux résidus de la partie précédente :

> *Il faudrait calculer à postériori la grandeur probable de chaque erreur, et, par conséquent, la valeur probable de la quantité à mesurer. Mais ainsi que je viens de l'expliquer, on ne saurait entreprendre ce calcul, si l'on n'admettait à priori, c'est à dire avant toute observation, une loi de probabilité des erreurs. Y a-t-il une loi des erreurs ?* (p.211)

Deux types d'erreur

Poincaré distingue les erreurs systématiques des erreurs accidentelles.

Les erreurs systématiques : lorsque je mesure avec un mètre trop long, quelle que soit ma précision, la mesure sera systématiquement fausse. Je peux corriger les résultats d'un coup en leur soustrayant à tous l'excès faussement mesuré, par exemple.

Les erreurs accidentelles : lorsque je mesure avec un mètre « avec la bonne longueur », il y a forcément des décimales qui vont m'échapper. Je ne peux rien faire, c'est comme ça.

La courbe de Gauss et la méthode des moindres carrés

La première question que pose l'auteur est la question de savoir si les erreurs accidentelles satisfont à la loi de Gauss, eu égard au fait que les erreurs systématiques ne nous concerneront pas dans les premières questions qui nous animent ici. Il faut comprendre, pour poursuivre la réflexion dans les meilleures conditions, que la courbe de Gauss, habituellement utilisée pour représenter du bruit – c'est à dire une distribution contingente de phénomènes – est liée directement aux résidus par ce que l'on nomme en statistique la « méthode des moindres carrés ».

Si je veux tirer des informations liées à mes résidus, autrement dit, si je souhaite travailler autour de mes erreurs, je dois considérer le fait qu'un résidu peut être à la fois positif et négatif, en ceci que certaines mesures seront au-dessus de la prévision fournie par la loi, et d'autres mesures seront en-dessous de cette prévision. Élever un nombre au carré a plusieurs effets mathématiquement intéressants dans la démarche, dont celui de rendre positives toutes les valeurs sortantes.

Il s'agira alors de minimiser des carrés, ce qui signifie aussi minimiser l'erreur. A titre d'exemple, pour trois mesures où l'on trouverait les valeurs 1, 2 et 3, la méthode des moindres carrés nous dit alors qu'en minimisant l'expression $(x-1)^2+(x-2)^2+(x-3)^2$, on accède à une valeur de x « optimale » selon les mesures effectuées. Dans ce cas particulier, cette valeur optimale vaut $x=\frac{1+2+3}{3}=2$. Pour prouver la pertinence de cette méthode, il est nécessaire d'admettre les mesures comme étant distribuées suivant une loi normale, c'est à dire selon une courbe de Gauss.

La question est alors de savoir si l'usage de la méthode des moindres carrés est légitime, ainsi que de savoir sur *quoi* elle nous renseigne exactement. Pour illustrer cette question, Poincaré oppose un astronome et un physicien.

> *[Les astronomes], outre les erreurs systématiques qu'ils rencontrent comme les physiciens, ont à lutter avec une cause d'erreur extrêmement importante et qui est tout à fait accidentelle ; je veux parler des ondulations atmosphériques. [...] Le physicien, persuadé qu'une bonne mesure vaut mieux que beaucoup de mauvaises, se préoccupe avant tout d'éliminer à force de précautions les dernières erreurs systématiques et l'astronome lui répond : « Mais vous ne pourrez observer ainsi qu'un petit nombre d'étoiles ; les erreurs accidentelles ne disparaitront pas ».* (p.212)

Le prix d'une loi vraie

Si l'on souhaite uniquement connaitre la valeur probable, c'est à dire obtenir par régression une loi qui s'applique « tant bien que mal » dans tel ou tel cas, la règle pratique de la méthode des moindres carrés est appropriée. Mais alors il est nécessaire de faire la différence entre (1) la probabilité que la valeur soit de tant et (2) la nécessité que la valeur soit de tant eu égard à une certaine quantité d'erreurs accidentelles. On réalise alors que le second cas est impossible, c'est ce qu'affirme Poincaré lorsqu'il dit :

> Cela est absolument illégitime *; cela ne serait vrai que si nous étions sûrs que toutes les erreurs systématiques sont éliminées, et nous n'en savons absolument rien.* (p.212)

On peut alors conclure que la régression, ou « méthode des moindres carrés », ou encore « méthode de Gauss », nous permet uniquement d'évaluer la qualité d'une série d'observations par rapport à une autre en comparant les informations que l'on peut trouver dans les erreurs de ces dernières. L'erreur systématique ne peut pas faire l'objet d'un « test statistique » ou d'une analyse objective pertinente en ceci que, comme l'indique justement Poincaré, « notre ignorance à ce sujet [est] absolue » (p.213).

Pour illustrer cette situation dans le calcul statistique contemporain, on peut indiquer que pour un certain nombre *n* de données dont on propose une hypothèse quelconque – disons *n* personnes dont on soupçonne une corrélation entre la couleur des cheveux et celle des yeux –, le test statistique, qui nous donnera la réponse à la question de savoir si cette corrélation observée peut être considérée comme un phénomène universel ou est juste une contingence locale, dépend de plusieurs paramètres dont *n*, mais aussi et surtout d'un seuil de tolérance α. Ce seuil est une marge d'erreur que l'on s'autorise à avoir une fois la règle établie. S'il est clair qu'une courbe de Gauss – ou quelque chose de similaire telle qu'une loi de Poisson – se dissimule derrière les processus mathématiques permettant d'établir la règle absolue de la corrélation entre couleur des yeux et couleur des cheveux, il est aussi clair qu'à un moment donné le statisticien a la nécessité de définir un taux d'erreur acceptable α tel que, par exemple, $\alpha = 5\%$. C'est à dire qu'il aura fabriqué une règle universelle de la corrélation entre couleur des cheveux et couleur des yeux, mais au prix de se tromper nécessairement dans 5% des cas observés. Si le statisticien avait voulu réduire cette marge d'erreur, il aurait peut-être pu le faire dans une moindre mesure, mais passé un certain stade, le test invalide la généralisation de la loi en question. C'est à dire que pour $\alpha = 0.5\%$, par exemple, la loi « générale » n'aurait certainement plus pu exister.

On peut donc conclure en diapason avec Poincaré – même s'il ne disposait pas de l'ensemble de la batterie statistique-mathématique d'aujourd'hui – qu'en principe, la généralisation coûte une certaine marge d'erreur, et il importe au statisticien comme au physicien, c'est à dire à l'ensemble des scientifiques, de fixer cette marge avec intelligence, en gardant les points suivants en tête :

1. La formalisation du problème doit admettre la possibilité d'une erreur systématique.

2. Minimiser les erreurs a priori c'est maximiser la qualité des mesures.
3. Trouver le meilleur consensus entre la loi générale et la marge d'erreur de cette loi, c'est à dire connaître son prix, est un travail nécessaire. La question de savoir si le jeu en vaut la chandelle apparait alors.

Le mot de la fin

Avant d'être une discipline concrète à part entière, les probabilités sont un sujet de réflexion à nette propension philosophique. Lorsqu'on se remet dans le contexte du début du siècle passé, on remarque à quel point Poincaré soulève des questions fondamentales qui animaient déjà Aristote et qui nous animent encore aujourd'hui. La prise de pouvoir des sciences dites « dures » au sein du domaine des sciences humaines, c'est à dire notamment l'application systématique de modèles statistiques à des populations d'humains et la légitimité d'une telle application, sont de bons exemples de questions quotidiennes profondes que peut soulever une réflexion autour des probabilités. Je n'aborderai pas les implications liées à ce qui est communément nommé « libre arbitre » et « déterminisme », car un texte d'une vingtaine de pages ne suffirait pas à couvrir la complexité de ces questions.

Plus concrètement, et pour revenir à la physique, on peut dire que Poincaré est né (ou a publié) juste trop tôt pour pouvoir voir la cristallisation des probabilités dans le réel par le biais de la future physique quantique. Car, parmi les deux types de liens entre probabilités et physique que j'ai noté plus haut (un lien épistémologique et un lien concret), nous ne traitons ici clairement que du premier type, eu égard à l'ensemble des questions philosophiques liées à l'induction logique que la réflexion ici présente caresse sans s'y plonger véritablement.

Dans son dernier paragraphe, l'auteur rappelle quelques points importants de cette démarche « d'induction » que l'on retrouve du début à la fin du chapitre.

1. Pour pouvoir appréhender un calcul de probabilité, pour que ce calcul fasse sens, il faut nécessairement admettre, comme point de départ, une hypothèse ou une convention qui comporte toujours un certain degré d'arbitraire.
2. Le choix de cette convention est guidé par le principe de raison suffisante, seule accroche que l'on a à ce stade de la réflexion.
3. Le principe de raison suffisante se manifeste essentiellement dans la croyance en une forme continue d'explication. Il peut aussi, alternativement, préférer de manière systématique l'explication la plus simple parmi un ensemble d'explication possibles, c'est le Rasoir d'Ockham.
4. La probabilité « objective » apparait lorsque l'hypothèse de départ est rendue contingente au cours du calcul. C'est, comme l'indique Poincaré, les

cas où « le calcul des probabilités peut être appliqué avec profit » (p.213).

Comme nous l'avons déjà dit, nous n'avons fait que passer d'une apparente régularité, ce que Poincaré appelle « probabilité subjective », à une loi générale, la « probabilité objective ». La question de la légitimité d'une telle loi était la question centrale de ce chapitre. Concluons donc en statisticien : la probabilité objective est légitimée, mais toujours au prix d'une certaine quantité d'erreur dont il est impensable de vouloir se débarrasser. Philosophiquement, ce serait comme dire que « nos lois sont vraies jusqu'à preuve du contraire », ce « contraire » étant soit une exception, soit un trop grand nombre d'observations injustifiables, qui nous obligeraient alors à rendre α trop grand pour conserver à notre loi une réelle valeur explicative.

On peut alors se rassurer un peu et utiliser l'outil statistique et le calcul des probabilités, mais les épées de Damoclès d'une mauvaise formalisation, du mauvais choix de convention, d'une foi illégitime en la simplicité de l'univers ou encore d'une erreur systématique inaperçue, ne sont jamais bien loin !

Bibliographie

Livres

- Poincaré, Henri, *La science et l'hypothèse*, Flammarion, 1902/1968
- Poincaré, Henri, *La valeur de la science*, Flammarion, 1905/1970
- Bertrand, Joseph, *Calcul des probabilités*, Gauthier-Villars, 1889
- Bavaud, François, *Modèles et données*, L'Harmattan, 1998
- Esfeld, Michael, *Philosophie des sciences*, Presses polytechniques et universitaires romandes, 2009
- Albert, David, *Quantum Mechanics and Experience*, Harvard University Press, 1994

Articles

- Dellacherie, Claude, « *Pascal et Fermat. La naissance de calcul des probabilités.* », 1994
- Bru, Bernard, « *La courbe de Gauss ou le théorème de Bernoulli raconté aux enfants* », 2006
- Hilgevoord, Jan & Uffink, Jos, « *The Uncertainty Principle* », Stanford Encyclopedia of Philosophy, 2006