

UNE HISTOIRE DES PROBABILITÉS

des origines à 1900



Une histoire des probabilités

des origines à 1900

Jean-Jacques Samuëli
Docteur ès Sciences physiques

Jean-Claude Boudenot
Agréé de Physique



Des mêmes auteurs

- Par-delà les nanosciences et les nanotechnologies*, Jean-Jacques Samueli, Ellipses, 2007
- Trente livres de mathématiques qui ont changé le monde*, Jean-Jacques Samueli, Jean-Claude Boudenot, Ellipses, 2006
- Trente livres de physique qui ont changé le monde*, Jean-Jacques Samueli, Jean-Claude Boudenot, Ellipses, 2007
- H.A. Lorentz (1853-1928). La naissance de la physique moderne*, Jean-Jacques Samueli, Jean-Claude Boudenot, Ellipses, 2005
- H. Poincaré (1854-1912), Physicien*, Jean-Jacques Samueli, Jean-Claude Boudenot, Ellipses, 2005
- Histoire de la Physique et des physiciens*, Jean-Claude Boudenot, Ellipses, 2001
- Max Planck et les quanta*, Jean-Claude Boudenot, Gilles Cohen Tannoudji, Ellipses, 2001
- Instrumentation électronique en Physique nucléaire*, Jean-Jacques Samueli, Jacques Pigneret, Armand Sarazin, Masson, 1968.

ISBN 978-2-7298-4042-6

© Ellipses Édition Marketing S.A., 2009
32, rue Bague 75740 Paris cedex 15



Le code de la propriété intellectuelle n'autorisant, aux termes de l'article L.122-5.2° et 3°a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective », et d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (Art. L.122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit constituerait une contrefaçon sanctionnée par les articles L. 335-2 et suivants du code de la propriété intellectuelle.

www.editions-ellipses.fr

Introduction

Il est très probable, voire évident, que pour comprendre la théorie des probabilités et connaître son évolution depuis ses prémices, le mieux est de lire les textes fondateurs de cette discipline dans leurs premières émissions, afin d'y découvrir l'origine réelle des concepts exposés et l'historique de leur mise en œuvre. Le sujet est sérieux, si l'on en croit Laplace, qui écrivait, au sujet des probabilités¹ : « Ce calcul délicat s'étend aux questions les plus importantes de la vie, qui ne sont en effet, pour la plupart, que des problèmes de probabilité. »

Le présent ouvrage est une initiation ou une introduction à l'histoire des probabilités et non une encyclopédie de cette histoire. Son but essentiel est de recenser les étapes majeures de la conception du calcul des probabilités en présentant et en commentant des extraits des premières éditions de textes fondateurs relatifs à la période pré-axiomatique, comprise depuis l'origine jusqu'à la fin du XIX^e siècle, ou plus exactement jusqu'à l'introduction d'une théorie de la mesure, qui n'est pas ici incluse. On n'a donc pas traité ce que l'on a coutume d'appeler le développement de la théorie moderne des probabilités. Quelques mots relatifs à l'axiomatique moderne du calcul des probabilités sont cependant ajoutés, en forme d'épilogue, à la fin de l'ouvrage. P. Lévy disait, en effet, que « Kolmogorov a donné à l'axiomatique du calcul des probabilités une forme qui semble définitive² ».

Les textes rédigés en latin ou dans une langue étrangère ont été traduits lorsqu'aucune traduction française n'était disponible. Mais notre travail a été simplifié par le fait que nombre d'auteurs majeurs de la théorie des probabilités dont Pascal, Fermat, Montmort, Condorcet, d'Alembert, Buffon, Laplace, Lagrange, Legendre, Poisson, Quetelet, Lacroix, Cournot, Joseph Bertrand, H. Poincaré et bien d'autres, étaient francophones.

Par ailleurs, les liens entre la science des probabilités et celle des statistiques sont multiples mais ne sont pas simples³. Le problème initial des définitions

-
1. *Théorie analytique des probabilités*, Courcier, 1812, page 2 de la dédicace à Napoléon-le-Grand.
 2. P. Lévy, *Les fondements du calcul des probabilités*, Dialectica, tome 3, 1949, p. 55.
 3. Cf. O. B. Sheynin, *The theory of probability: Its definition and its relation to statistics*, AHES, vol. 52, n° 2, 1998, p. 99-108.

est, par lui-même complexe. Pascal fut le premier à proposer un nom au calcul des probabilités qu'il baptisa *Géométrie du hasard*. Mais il n'entrevoyait pas d'autre application de cette *géométrie* en dehors du problème des points dans les jeux de hasard.

Jakob Bernoulli a donné une définition de l'art de conjecturer¹ : « Nous disons *savoir* ou *comprendre* ce qui est certain et indubitable ; *conjecturer* seulement ou *opiner*, tout le reste. *Conjecturer* quelque chose, est mesurer sa probabilité ; et pour cette raison *l'Art de Conjecturer* ou *Stochastique*, est défini par nous comme l'art de mesurer exactement les probabilités des choses, afin que nous puissions toujours choisir dans nos actions, ce qui est préférable ou plus habile ; la sagesse du Philosophe et la prudence du Politique sont d'ailleurs dirigées en ce sens. Les probabilités sont exprimées d'après *le nombre*, et d'après *le poids des arguments*, qui prouvent ou indiquent en quelque sorte que quelque chose arrive, arrivera ou est arrivé. D'autre part, par *Poids*, j'entends la force de prouver. »

Karl Pearson² a affirmé pour sa part que « la statistique est une science qui est l'application de la théorie mathématique à l'interprétation d'observations faites en masse ». Il ajoute qu'il est déplorable que le mot *statistiques* soit aussi utilisé pour ce qui est à proprement parler des données statistiques.

Le présent ouvrage étant consacré spécifiquement aux probabilités on ne trouvera qu'un seul chapitre, assez court, relatif aux statisticiens depuis Ulpien jusqu'à Quetelet.

La théorie des erreurs que l'on peut, éventuellement, considérer comme indépendante de la théorie des probabilités et de celle des statistiques, est, elle, traitée depuis les premiers travaux de Cotes, Simpson et Lambert jusqu'au mémoire de Chebyshev.

Les physiciens se posent généralement la question de savoir si les probabilités sont plus proches des mathématiques que de la physique. En physique, on recherche les lois de la nature, leurs éventuelles démonstrations et leurs domaines de validité. La démarche est quelquefois la même dans le calcul des probabilités, la loi des grands nombres ayant, par exemple, été l'objet de très nombreuses démonstrations bien qu'elle soit considérée par certains comme une loi de la nature. Mais Benoît Mandelbrot a affirmé, en 1973, dans sa théorie des fractales, qu'il n'y a pas une seule forme de hasard. Pour lui, beaucoup de phénomènes naturels obéissent à un autre type de hasard, qu'il appelle hasard sauvage, auquel on ne peut appliquer la loi des grands nombres. Le bruit électronique en $1/f$, par exemple, où l'amplitude d'une

1. *Ars conjectandi*, Bâle, 1713, page 213.

2. *The History of statistics in the 17th and 18th centuries*, London, Griffin, 1978, p. 3.

fréquence lui est inversement proportionnelle, n'obéit ni à la loi des grands nombres ni au théorème central limite¹.

Poisson, dans ses *Recherches sur la probabilité des jugements* (1837), page 7, écrivait :

« Les choses de toutes natures sont soumises à une loi universelle qu'on peut appeler *la loi des grands nombres*. Elle consiste en ce que, si l'on observe des nombres très considérables d'événements d'une même nature, dépendants de causes constantes et de causes qui varient irrégulièrement, tantôt dans un sens, tantôt dans l'autre, c'est-à-dire sans que leur variation soit progressive dans aucun sens déterminé, on trouvera, entre ces nombres, des rapports à très peu près constants. Pour chaque nature de choses, ces rapports auront une valeur spéciale dont ils s'écarteront de moins en moins, à mesure que la série des événements observés augmentera davantage, et qu'ils atteindraient rigoureusement s'il était possible de prolonger cette série à l'infini. »

L'*Essai sur les probabilités* de 1814 de Laplace contenait, par ailleurs, dès les premières pages, l'affirmation d'un déterminisme universel dans toute sa rigueur. Laplace l'énonçait comme suit, son déterminisme étant synonyme de prédictibilité :

« Nous devons envisager l'état présent de l'univers comme l'effet de son état antérieur, et comme la cause de celui qui va suivre. Une intelligence qui, pour un instant donné, connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome : rien ne serait incertain pour elle, et l'avenir, comme le passé, serait présent à ses yeux. L'esprit humain offre, dans la perfection qu'il a su donner à l'astronomie, une faible esquisse de cette intelligence. Ses découvertes en mécanique et en géométrie, jointes à celles de la pesanteur universelle, l'ont mis à portée de comprendre dans les mêmes expressions analytiques les états passés et futurs du système du monde. En appliquant la même méthode à quelques autres objets de ses connaissances, il est parvenu à ramener à des lois générales les phénomènes observés, et à prévoir ceux que les circonstances données doivent faire éclore. »

La mécanique quantique et la théorie du chaos ont mis à mal le déterminisme de Laplace. Mais l'objet de cet ouvrage n'est pas de philosopher ou d'émettre des jugements de valeur. Ce livre tente d'être une introduction à quelques textes fondateurs de la théorie des probabilités.

1. B. Mandelbrot, *Du hasard bénin au hasard sauvage*, Pour la Science, dossier « Le hasard », Paris, 1996.

Les lecteurs cherchant une encyclopédie de l'histoire des probabilités et des statistiques pourront s'orienter, par exemple, vers deux ouvrages récents d'Anders Hald¹, publiés en langue anglaise. Ce sont actuellement, il nous semble, des ouvrages incomparables.

Le lecteur intéressé par l'histoire de l'enseignement des probabilités et des statistiques pourra lire l'excellent article de Robert Meusnier portant ce titre et publié en 2004².

L'appartenance de la théorie des probabilités à la Mathématique n'a, en effet, été reconnue, du moins en France, qu'à partir de 1969 lorsque le groupe *Bourbaki*, c'est-à-dire André Weil cessa de la marginaliser.

Meusnier écrit : « Ce n'est qu'en 1969 que paraît le chapitre 9 Intégration... qui réintroduit les espaces non localement compacts... et la théorie des probabilités, et tente de manière ampoulée de présenter la mesure de Wiener. Ainsi s'est-il écoulé 34 ans depuis la décision du Comité de rédiger un volume sur l'intégration et 32 ans depuis la rédaction du "Diplodocus" qui suivait la voie de De Possel, basée sur le primat de l'intégration sur des espaces abstraits et sa remise en cause par l'intervention du plan Weil basé sur le primat de l'intégration dans des espaces topologiques localement compacts. »

Ayant été jadis un élève de René de Possel et d'André Blanc-Lapierre, je ne peux qu'apprécier le texte de Robert Meusnier.

Enfin, je voudrais ajouter que je déplore infiniment la disparition, survenue le 15 octobre 2008 des suites d'une longue maladie, de mon collègue et ami Jean-Claude Boudenot, co-auteur du présent livre, qui était passionné par l'histoire des sciences.

J.-J. Samuëli,
Novembre 2008.

-
1. A. Hald, *History of probability and statistics and their applications before 1750*, Wiley, 1990 ; A. Hald, *A history of mathematical statistics from 1750 to 1930*, Wiley, 1998.
 2. « Sur l'histoire de l'enseignement des probabilités et des statistiques », in *Histoires de probabilités et statistiques*, Paris, Ellipses, 2004, p. 237-273.

Débuts du calcul des probabilités

1.1. Les probabilités avant 1550

Si la plus ancienne tentative de rédaction d'une histoire des Probabilités est due à Martin Strömer, le traducteur d'Euclide en suédois, qui, sous la direction de Samuel Klingenshierna rédigea, en 1731, un mémoire de thèse de 16 pages intitulé *Specimen Historiae literariae De Arte Conjectandi*, le plus ancien travail consacré au calcul des probabilités et aux statistiques proprement dites est, sans doute, celui d'un auteur romain, Domitius Ulpianus, qui, vers 220 après J.-C. établit une table de valeurs d'annuités pour les assurances-vie qui fut utilisée jusqu'au début du XIX^e siècle.

Selon Guillaume Libri¹, la *Divine Comédie* de Dante comporterait la plus ancienne indication des différentes probabilités des divers lancers possibles avec trois dés. Le texte est contenu dans un commentaire publié à Venise en 1477. Libri indique, de plus, qu'il considère que le mot *hasard* trouve son origine dans le terme, utilisé par Dante, *azari* (points difficiles) qui, lui-même vient de *asar* qui en arabe signifie difficile.

En fait, le calcul relatif aux trois dés est plus ancien.

Un ouvrage de Baudri de Théroutanne² paru en 1615 indique comment l'évêque Wibold de Cambrai inventa, vers 960, un jeu, se voulant édifiant qui consistait à jeter trois dés, chaque combinaison de nombres réalisés étant associée à une vertu. Wibold avait donc énuméré correctement les 56 différentes sorties (sans permutations). Le joueur s'appropriait, croyait-on, la vertu désignée par le hasard. Voici, par exemple, quelques-unes des vertus énumérées par Wibold : 1,1,1 la charité ; 1,1,2 la foi ; 1,1,3 l'espoir ; 1,1,4 la justice ; 1,1,5 la prudence ; ... ; 4,5,5 l'intelligence ; ... ; 6,6,6 l'humilité.

De nos jours, la loi des grands nombres nous permet de conclure que Wibold considérerait qu'une personne seulement sur cinquante-six est intelligente !

1. G. Libri, *Histoire des sciences mathématiques en Italie*, tome II, pages 188-189, Renouard, 1838.
2. Baudri de Théroutanne, *Chronicon Cameracense et Atrebatense sive Historia utriusque ecclesiae...*, Douai : J. Bogard, 1615.

HIERONYMI
CARDANI
MEDIOLANENSIS

Philosophi ac Medici Celeberrimi

OPERA OMNIA:
TAM HACTENUS EXCVSA;

hic tamen multa & emendata: quàm nunquam alijs vís,
ac primò ex Authoris ipsius Autographis edita:

Cura CAROLI SPONII,
DOCTORIS MEDICI COLLEGIO MEDD.
Lectionariorum Aggregati.

TOMVS PRIMVS:

2^{do} CONTINENTE

PHILOGICA. LOGICA. MORALIA.

Etiam vtriusq; Operum, ad eadem vix Antiqui habetur.



LVGDVNI.

Sumptibus IOANNIS ANTONII HVGVEIAN,
& MARCI ANTONII RAVAVD.

M. DC. LXIII.
CVM PRIVILEGIO REGIS

Par ailleurs, un poème du XIV^e siècle en trois chants, intitulé *de Vetula*, traditionnellement attribué à Ovide mais, selon H. Cocheris¹, vraisemblablement œuvre de Richard de Fournival, Chancelier de la cathédrale d'Amiens, écrite au siècle précédent, raconte un épisode de la vie amoureuse de l'auteur de *L'Art d'aimer*. Mais ce même poème comporte de très intéressantes digressions sur les jeux de dés et sur les probabilités. Le nombre de façons de lancer trois dés y est calculé et on peut traduire comme suit le passage correspondant :

« Si les trois nombres affichés sont les mêmes, il y a six possibilités, si deux sont identiques et le troisième différent il y a 30 possibilités, parce que la paire peut être choisie de six façons et l'autre de cinq, enfin, si les trois sont différents il y a 20 façons, parce que 30 fois quatre fait 120, mais chaque possibilité intervient trois fois. Il y a au total 56 possibilités. Mais si les trois nombres sont identiques, il n'y a qu'une possibilité pour chaque nombre, il y en a trois pour deux nombres identiques et six si les trois nombres sont différents. »

Comme l'indique M. Kendall², Fournival, bien qu'il ne l'indique pas explicitement dans son texte, a bien calculé que le nombre total de façons de lancer trois dés est :

$$(6 \times 1) + (30 \times 3) + (20 \times 6) = 216$$

1.2. Cardan, *Liber de ludo aleae*

Jérôme Cardan fut à la fois un mathématicien célèbre et un joueur invétéré. Il écrivit, vers la fin de sa vie, une autobiographie qui fut publiée³ par Gabriel Naudé en 1643. On peut y lire que Cardan joua aux dés durant vingt-cinq ans et qu'il avoue : « je ne veux pas dire seulement de temps en temps au cours de ces années, mais, j'ai honte de le dire, chaque jour. » Il écrit aussi, un peu plus loin : « Pourquoi un joueur de dés qui est écrivain, n'écrit-il pas sur les jeux ? »

Cardan, a lui-même indiqué la date de rédaction de son *Liber de ludo aleae*. [Le livre sur les jeux de hasard]. Plus exactement, on peut déduire qu'il a rédigé le chapitre 20 de son livre en 1564, car, racontant dans ce chapitre un fait survenu en 1526, Cardan indique que depuis ladite époque, trente-huit années se sont écoulées. Il est donc possible que la rédaction du livre ait été achevée peu après, les derniers chapitres étant relativement courts.

Le texte a été imprimé pour la première fois en 1663, près d'un siècle après sa rédaction, dans ses œuvres complètes⁴ en dix volumes (dans le volume 1,

1. H. Cocheris, *La Vieille ou les dernières amours d'Ovide*, poème français du XIV^e siècle traduit du latin de Richard de Fournival par Jean Lefèvre, Paris, 1861.
2. M. G. Kendall, *Biometrika*, vol. 43, n° 1/2 (Jun., 1956), p. 1-14.
3. *De propria vita Hieronymi Cardani Mediolanensis, De propria vita liber*, ex bibliotheca Gab. Naudaei, Parisiis, Iacobum Villery..., 1643.
4. Cardano, Girolamo, *Opera omnia: tam hactenus excusa, cura Caroli Sponii*, Lugduni, Sumptibus Ioannis Antonii Hygvetan, & Marci Antonii Ravard, 1663.

chapitre 10). Un ouvrage consacré à Cardan a été publié en 1953 par O. Ore¹ et comporte une traduction en anglais du *Liber de ludo aleae* due à S.H. Gould. Plus récemment le livre de Cardan a été traduit en italien².

À l'époque de la Renaissance, lors de sa rédaction, le *Liber de Ludo Aleae* de Cardano était à la fois un manuel décrivant les jeux de hasard et un traité avancé de calcul des probabilités. De plus, ce livre comporte de nombreuses considérations relatives à l'équité des jeux, avec le double souci pour le joueur d'être honnête et de ne pas être la victime de tricheurs. Cardan a certes écrit son ouvrage dans un style coloré, mais on peut déceler qu'il a été fortement influencé par l'*Éthique* d'Aristote, en particulier par le concept de justice de ce philosophe.

Le *Liber de ludo aleae* est divisé en trente-deux chapitres. Les huit premiers traitent des genres de jeux (d'adresse, d'habileté ou de hasard), prodiguent des conseils aux joueurs, et soulignent les dangers et les avantages liés au fait de jouer. Les six chapitres suivants abordent l'aspect mathématique des gains aux jeux et la formulation des probabilités simples et composées, particulièrement dans le cas des jeux de dés. Les derniers chapitres sont consacrés aux jeux de cartes, au primero (genre de poker), au backgammon et à l'utilisation des astragales.

Les considérations exposées par Cardan dans ses premiers chapitres sont fort utiles pour découvrir le comportement social dans l'Italie morcelée et ravagée par les guerres du XVI^e siècle. On peut y lire, par exemple :

Au chapitre 2, *Sur les conditions du jeu* :

« L'attention doit être attirée sur [...] les conditions dans lesquelles le jeu est organisé, telles que le montant des mises engagées et l'occasion qui détermine à jouer ; au sujet de ce dernier point il est très important de demander une permission si l'on désire jouer à l'occasion d'un banquet de funérailles. [...]

Mais dans des moments de grande anxiété et de chagrin, le jeu est considéré non seulement comme permis mais même comme bénéfique. Aussi est-il permis aux prisonniers, aux condamnés à mort et aux malades, et en conséquence, la loi permet le jeu lors d'un chagrin [...] Dans mon cas, lorsqu'il me semblait, après une longue maladie, que la mort était proche, je trouvais une grande consolation en jouant aux dés en permanence. »

Au chapitre 3, *Qui peut jouer et quand ?* :

« Si une personne est renommée pour sa sagesse et honorée par un poste de magistrat ou par un autre honneur civil ou s'il fait partie de l'Église, rien n'est pire pour elle que de jouer ; Réciproquement, le jeu est proportionnellement moins reprochable

1. Ore Oystein, *Cardano the gambling scholar*, Princeton University Press, 1953. Notre lecture du texte de Cardan a été faite à partir de la traduction anglaise incluse dans cet ouvrage, c'est pourquoi nous citons les titres des chapitres traduits en anglais.
2. Gerolamo Cardano, *Liber de ludo aleae* (traduz. italiana a cura di Giorgio ORFINO), Ed. Stamperia Ammiano Quinto de' Stampi, Milano 2003.

aux garçons, aux jeunes hommes et aux soldats. Et plus importante est la somme d'argent mise en jeu, plus grande est la disgrâce, ainsi *un haut dignitaire de l'Église, (en fait un cardinal)*, fut sévèrement blâmé parce qu'il avait joué après dîner avec le duc de Milan, l'enjeu étant de cinq mille couronnes ».

Plus loin on trouve ce fort amusant jugement :

« Les hommes de loi et les docteurs retirent peu d'avantages à jouer : d'une part, ils semblent avoir trop de temps libre et d'autre part, s'ils gagnent, ils sont pris pour des joueurs et s'ils perdent ils peuvent paraître aussi peu sérieux dans leur art que dans leur amusement. Les hommes de ces professions prennent le risque d'un même jugement s'ils pratiquent la musique ».

Au chapitre 4, *De l'utilité du jeu et des pertes*, Cardan affirme :

« Le jeu est un très bon test du caractère patient ou impatient d'un homme. Le plus grand avantage que l'on peut retirer du jeu provient du fait de ne pas jouer du tout. »

Le chapitre 6 est intitulé *Le principe fondamental du jeu*. Cardan l'énonce ainsi :

« Le principe le plus fondamental de tous, dans le jeu, est simplement des conditions égales en ce qui concerne les joueurs, les assistants, l'argent, le tablier du jeu et le dé lui-même. Dans la mesure où vous vous écarterez de cette égalité, si c'est en faveur du joueur qui vous est opposé, vous êtes fou, et si c'est en votre faveur, vous êtes injuste. »

Le chapitre 14, *Sur les points combinés*, est consacré aux événements composés que Cardan appelle « points combinés » dans le cas des jeux de dés. Cardan énumère tout d'abord les divers cas possibles puis par addition il dérive les événements composés. Voici sa présentation :

« Nous devons prendre en considération (les points combinés) dans le cas de deux dés parce que le un arrive dans 11 lancers, ainsi que le deux, ainsi que le trois et ainsi de suite pour tous les autres, mais le un ou deux n'arrive pas dans 22 lancers mais seulement en 20. Car le un intervient 11 fois et le deux en 9 fois de plus. Ainsi si le trois est ajouté, on n'aura pas 29 ni 31 mais 27 et plus généralement les nombres de lancers sont donnés dans la table qui suit : »

Deux dés		Total
Cas pour 1 point	11	11
Supplément pour 2	9	20
Supplément pour 3	7	27
Supplément pour 4	5	32
Supplément pour 5	3	35
Supplément pour 6	1	36
Total	36	

Cardan établit ensuite la même table pour trois dés :

Trois dés	
Cas pour 1 point	91
Supplément pour 2	61
Supplément pour 3	37
Supplément pour 4	19
Supplément pour 5	7
Supplément pour 6	1
Total	216

À la fin de son chapitre 14, Cardan donne ce qu'il appelle une règle générale pour l'équité d'un jeu. Il écrit :

« Ainsi, il y a une règle générale, à savoir qu'il faut considérer le circuit total¹ et le nombre total des lancers [de dés] qui représentent en combien de cas le résultat favorable est possible et comparer ce nombre à celui des cas correspondants au reste du circuit, et il faut que les mises mutuelles soient versées conformément à cette proportion, de telle sorte que chacun puisse jouer en termes égaux. »

Le livre présente pour la première fois le concept de probabilité énoncée comme le rapport des cas favorables aux cas possibles lorsque tous les cas sont équiprobables.

Dans les chapitres 14 et 15, Cardan présente la loi des puissances : si p est la probabilité $p = \frac{f}{t}$ d'un événement, c'est-à-dire le rapport du nombre de cas favorables f au nombre de cas total t et $\frac{f}{t-f}$ le rapport du nombre de cas favorables au nombre de cas défavorables pour un tirage (ce rapport est pour un joueur un paramètre plus important que la probabilité), alors pour n répétitions, ce dernier rapport devient $\frac{f^n}{t^n - f^n} = \frac{p^n}{1 - p^n}$. Cardan obtient ce résultat en analysant les jeux de dés et en itérant son raisonnement. Il donne les exemples suivants à la fin du chapitre 15. Pour un jeu avec deux dés, il écrit :

« Disons que les cas favorables sont le 1, le 2, ou le 3 obtenus en trois lancers successifs. Le nombre total du circuit est 36 et le nombre de cas favorables 27. Nous multiplions 36 par lui-même 3 fois ; le résultat est 46 656. Nous multiplions aussi 27 par lui-même 3 fois et le résultat est 19 683. Soustrayons le plus petit du plus grand ; le reste est 26 973 qui donne le rapport du nombre de cas favorables au nombre de cas défavorables lorsqu'il est comparé à 19 683. Ce rapport correspondant au rapport du reste de la soustraction au plus petit nombre est plus grand que $4/3$ et plus petit que $3/2$. »

1. Cardan entend par là le nombre total des cas équiprobables possibles.

Cardan analyse alors le cas d'un jeu avec trois dés :

« De même, il a été établi qu'avec trois dés, une face quelconque, quelle qu'elle soit, a 91 cas favorables d'être tirée une fois sur un circuit total de 216. Donc, si cette face est demandée successivement dans trois tirages, nous devons multiplier le circuit total, et le résultat est 9 324 125. Quand ce dernier nombre est divisé par l'autre nombre, c'est-à-dire 753 571, nous obtenons le rapport des mises devant être jouées ici, un peu plus de 12 à 1. »

[le calcul de Cardan est $\frac{91^3}{216^3 - 91^3} = \frac{753.571}{9.324.125}$, rapport un peu inférieur à 1/12.]

Il est à noter que dans un ouvrage¹ publié en 1539, Cardan a aussi abordé le « problème des partis » que nous étudierons avec la correspondance de Pascal et Fermat. Cardan est le premier à réaliser que la division des enjeux misés dans un jeu interrompu ne dépend que du nombre de parties perdues par chaque joueur.

Le *Liber de ludo aleae* de Cardan n'a été rendu public, comme nous l'avons indiqué plus haut, qu'en 1663. Il n'eut donc qu'une faible influence sur la théorie des probabilités car quelques années plus tôt, en 1657, le mémoire de Huygens *De ratiociniis in Ludo Aleae* avait été inclus dans le livre que Frans van Schooten (1615-1660) publia sous le titre *Exercitationum mathematicorum libri quinque* (Leiden, Elsevir).

Ce mathématicien hollandais avait rencontré Descartes à son arrivée à Leyde et ce dernier lui avait montré les épreuves de sa *Géométrie* (Van Schooten les traduira en latin et publiera cette traduction en 1649, puis, à nouveau, avec ses commentaires en 1660). Afin d'approfondir ses connaissances en mathématiques, van Schooten voyagea à Paris où, muni de lettres de recommandation de Descartes, il rencontra Mersenne et lut les manuscrits de Fermat et de Viète. Il se rendit ensuite en Angleterre avant de rentrer à Leyde en 1643. À la demande de l'éditeur Elzévir, il rapporta une copie des œuvres de Viète. Il rapporta également une copie de celles de Fermat mais ne put persuader Elzévir de les publier, à cause, sans doute, du fait que Descartes avait émis une opinion défavorable sur l'œuvre de Fermat. Christian Huygens, pour sa part, arriva à Leyde en 1645 et devint l'élève de van Schooten. En 1657, van Schooten publia en appendice à son ouvrage intitulé *Exercitationum mathematicorum libri quinque* la première édition du papier de Huygens *De ratiociniis in ludo aleae*. Ce mémoire, le premier ouvrage imprimé traitant des probabilités, deviendra célèbre et Jakob Bernoulli le reproduira avec un commentaire dans la première partie de son *Ars Conjectandi*.

1. *Practica arithmetice, & Mensurandi singularis. In quaque preter alias continentur, versa pagina demonstrabit*, Milan, J.A. Castellioneus pour B. Calusci, 1539.

1.3. Galilée, *Sopra le scoperte dei dadi*

Galilée a rédigé vers 1620 sous le titre *Sopra le scoperte dei dadi*, un mémoire où il calcule la probabilité que la somme des faces de trois dés soit égale à une certaine valeur, cette question lui ayant été posée par le grand-duc de Toscane. La première publication du texte date de 1718 dans une édition collective des Œuvres de Galilée¹.

Nous en donnons ci dessous une traduction :

« Recherche concernant le jeu de dés

Le fait que dans un jeu de dés certains nombres sont plus avantageux que d'autres a une raison évidente, à savoir le fait que les réalisations de ces nombres sont plus aisées et plus fréquentes que d'autres car ils sont plus à même d'être obtenus par une plus grande variété de nombres. Ainsi un 3 et un 18, qui résultent de jets qui ne peuvent être réalisés que d'une seule manière avec trois dés (dans le dernier cas par trois 6 et dans le premier par trois 1, et de nulle autre manière) sont plus difficiles à obtenir que, par exemple 6 ou 7, qui peuvent être obtenus de diverses façons, le 6 avec 1,2,3 ou 2,2,2 ou 1,1,4, et le 7 avec 1,1,5 ou 1,2,4 ou 1,3,3 ou 2,2,3.

Néanmoins, bien que 9 et 12, peuvent être obtenus par le même nombre de manières que 10 et 11, et qu'ils devraient donc être considérés comme de même utilité dans ce jeu, on sait déjà par une longue observation que les joueurs considèrent 10 et 11 comme plus avantageux que 9 et 12. Et il est clair que 9 et 10 peuvent être composés par une égale diversité de nombres (et ceci est également vrai pour 12 et 11) : car 9 est composé de 1,2,6 ou 1,3,5, ou 1,4,4 ou 2,2,5 ou 2,3,4 ou 3,3,3, qui sont six triplets, et 10 de 1,3,6 ou 1,4,5 ou 2,2,6 ou 2,3,5 ou 2,4,4 ou 3,3,4 et pas d'autres manières, et ces dernières sont aussi six triplets.

Maintenant, afin d'obliger la personne qui m'a ordonné d'étudier le problème, je vais exposer mes idées, dans l'espoir non seulement de résoudre ledit problème, mais aussi d'ouvrir la voie à une compréhension précise des raisons pour lesquelles tous les détails du jeu ont été arrangés et ajustés avec grand soin.

Et pour atteindre ce but avec toute la clarté dont je suis capable, je vais commencer par considérer que, puisqu'un dé a six faces, et que lorsqu'il est jeté il peut également tomber sur chacune d'elles, seulement six jets différents les uns des autres peuvent être réalisés avec lui. Mais si en même temps que ce premier dé nous en jetons un second, qui lui aussi a six faces, nous pouvons réaliser 36 jets, chacun étant différent de tous les autres, puisque chaque face du premier dé peut être combinée avec chaque face du second, et en conséquence peut amener six jets différents, et donc il est clair que le nombre de ces combinaisons est 6 fois 6, soit 36. Et si nous ajoutons un troisième dé, puisque chacune de ses six faces peut être combinée avec chacune des 36 combinaisons des deux autres dés, nous trouverons que les combinaisons de trois dés sont au nombre de 6 fois 36, soit 216, chacune d'elles étant différente des autres. Mais parce que les nombres [sommés] de ces combinaisons résultant du jet de trois dés sont seulement au nombre de 16, à savoir

1. *Opere di Galileo Galilei*, Coll'aggiunta di varii trattati dell'istesso autore non più dati alle stampe, Tartini e Franchi, Firenze, 1718. Reproduit in *Opere*, Barbera, Firenze, tome 8, p. 591-594, 1898.

3,4,5 etc. jusqu'à 18, parmi lesquels nous devons partager les 216 jets, il est donc nécessaire que de nombreux jets correspondent à chacune de ces sommes ; et si nous pouvons trouver combien appartiennent à chacune, nous aurons préparé la voie pour trouver ce que nous cherchons. Il sera suffisant d'effectuer une telle recherche pour des sommes de 3 à 10, car ce qui correspond à l'une de ces sommes correspond aussi aux sommes plus grandes¹.

Trois points doivent être notés spécifiquement pour une bonne compréhension de ce qui suit. Le premier est que la somme des points de trois dés qui est composée de trois nombres égaux, ne peut être réalisée que par un jet unique des dés : et ainsi un 3 ne peut être réalisé que par trois faces 1, et un 6, s'il doit être réalisé avec trois 2 ne peut être réalisé que par un jet unique. Deuxièmement : la somme qui résulte de trois nombres dont deux sont identiques et le troisième différent, peut être réalisée par trois jets : par exemple un 4 qui est composé de 2 et de deux 1, peut être obtenu par trois jets différents ; c'est-à-dire quand le premier dé montre 2 et le second ainsi que le troisième montrent 1, ou le second dé un 2 et le premier et le troisième un 1 ; ou le troisième un 2 et le premier et le second un 1. Et également, par exemple un 8, quand il est réalisé avec 3,3,2 peut aussi être réalisé de trois façons ; c'est-à-dire quand le premier dé montre 2 et le second ainsi que le troisième montrent 3, ou le second dé un 2 et le premier et le troisième un 3 ; ou le troisième un 2 et le premier et le second un 3. Troisièmement la somme des points qui résulte de trois nombres différents peut être obtenue de six façons. Par exemple un 8 réalisé par 1,3,4 peut être obtenu par six différents jets : d'abord quand le premier dé montre 1, le second 3, et le troisième 4 ; deuxièmement, quand le premier dé montre toujours 1, mais le second 4 et le troisième 3 ; troisièmement, quand le second dé montre 1, le premier 3 et le troisième 4 ; quatrièmement, quand le second dé montre toujours 1, le premier 4 et le second 3. Donc, nous avons à présent déclaré ces trois points fondamentaux : premièrement que les triplets, c'est-à-dire la somme des jets des trois dés, qui résultent de trois nombres égaux, ne peuvent être produits que d'une manière, deuxièmement que les triplets qui sont composés de deux nombres égaux et d'un troisième différent sont produits de trois façons ; troisièmement que les triplets résultant de trois nombres différents sont produits de six façons. À partir de ces points fondamentaux nous pouvons facilement déduire de combien de façons, ou plutôt en combien de jets différents, toutes les sommes de trois dés peuvent être formées, ce qui sera facilement compris à partir de la table qui suit.

1										
3										
6										
10	631	6	621	6	611	3	511	3	411	3
15	622	3	531	6	521	6	421	6	321	6
21	541	6	522	3	431	6	331	3	222	1
25	532	6	441	3	422	3	322	3		
27	442	3	432	6	332	3				
108	433	3	333	1						
108		27		25		21		15		10
216									6	3
										1

1. NDT : Les résultats sont symétriques pour les sommes 3 à 10 et 11 à 18.

En haut de la table sont notés les points des jets de 10 à 3, et en dessous apparaissent les différents triplets desquels chacun peut résulter ; près desquels sont placés le nombre de façons par lesquelles chaque triplet peut être obtenu, et en dessous est finalement montrée la somme de toutes les façons d'obtenir ces jets. Ainsi, par exemple, dans la première colonne nous avons la somme de 10 points et en dessous ses six triplets de nombres avec lesquels il peut être réalisé qui sont 6,3,1 ; 6,2,2 ; 5,4,1 ; 5,3,2 ; 4,4,2 ; 4,3,3. Et comme le premier triplet 6,3,1 est constitué de trois nombres différents, il peut (comme indiqué ci-dessus) être réalisé par six différents jets de dés, et donc près de ce triplet un 6 est noté : et puisque le second triplet 6,2,2 est constitué de deux nombres égaux et d'un troisième différent, il peut uniquement être produit par trois jets différents, et donc un 3 est noté près de lui. Le troisième triplet 5,4,1, étant constitué de trois nombres différents, peut être réalisé par six jets et donc un 6 lui est accolé ; et ainsi de suite avec les autres triplets. Finalement en bas de la petite colonne de nombres de jets, ceux-ci sont additionnés : on peut ainsi voir que la somme de 10 points peut être réalisée par 27 différents jets alors que la somme de 9 points ne peut l'être que par 25 jets, le 8 par 21, le 7 par 15, le 6 par 10, le 5 par 6, le 4 par 3, et finalement le 3 par 1 jet. Le tout additionné ensemble donne 108. Et il y a le même nombre de jets pour les sommes de points supérieures, à savoir pour les points 11, 12, 13, 14, 15, 16, 17, 18, si bien que l'on arrive au nombre total de tous les jets possibles réalisables avec les six faces de trois dés, qui est de 216. Et à partir de la table, toute personne qui comprend le jeu peut précisément mesurer tous les avantages, aussi faibles qu'ils puissent être des *zare*, des *incontri*, et autres règles spéciales observées dans ce jeu. »

1.4. Biographie de Cardan (1501-1576)

Jérôme Cardan¹ est le fils illégitime de Fazio Cardano et de Chiara Micheria. Son père est avocat à Milan mais son expertise en mathématiques est telle qu'il enseigne aussi la géométrie à l'université de Pavie. Ce n'est que vers la cinquantaine qu'il rencontre Chiara Micheria, qu'il épousera plus tard mais ils donnent naissance à Jérôme dès 1501. Cardan doit à son père ses premières leçons et devient son assistant.

C'est à l'université de Padoue qu'il obtient son doctorat en médecine en 1526. À cette époque, son père meurt et notre futur mathématicien hérite d'un petit legs. Il postule alors pour enseigner à l'université de Milan, ville où vit sa mère. Il n'est toutefois pas accepté, à cause de sa réputation d'homme difficile, intransigeant et sans tact. Il dit lui-même dans son autobiographie : « Ce que j'identifie d'unique et d'exceptionnel parmi mes défauts, c'est l'habitude, dans laquelle je persiste, de dire surtout les choses dont je sais qu'elles vont contrarier mes auditeurs. Je m'en rends bien compte mais pourtant je garde cette habitude obstinément, en sachant combien cela me fait d'ennemis. » L'université de Milan prétexte sa naissance illégitime pour l'écarter. Sur les conseils d'un ami, il s'installe comme médecin à Saccolongo, petite ville située

1. Jérôme Cardan est également connu sous le nom de Girolamo (ou Hieronimo) Cardano, ou encore Hieronymus Cardanus en latin.

à 15 km de Padoue. Vers la fin de 1531 il épouse Lucia, la fille d'un de ses voisins, Aldobello Bandarini, capitaine de la milice locale. Son exercice de la médecine à Saccolongo ne suffisant plus à faire vivre son couple, il déménage en avril 1532 pour Gallarate, près de Milan. La même année, il postule de nouveau à l'université de Milan, mais toujours sans succès.

Pour améliorer ses finances, Cardan se livre aux jeux : dés et échecs principalement, mais il gaspille vite le legs qu'il tient de son père. Il décide alors de jouer en s'aidant du calcul des probabilités et, bénéficiant de cet avantage, il commence à gagner plus qu'il ne perd. Les jeux, à l'époque, ne se pratiquaient pas en très bonne compagnie et un jour, pensant être la victime d'une triche, Cardan taillade le visage de son adversaire avec son couteau. Cardan est resté toute sa vie mordu de jeu, ce qui l'a conduit à une perte importante de temps, d'argent et de réputation. Dès 1533, n'ayant plus guère de goût pour la pratique de la médecine, il fait du jeu son gagne-pain. Mais les choses se passent mal, il est forcé de mettre en gage les bijoux de sa femme et même certains de ses meubles. Les Cardan décident alors de vivre à Milan, mais là les choses empirent et ils sont obligés de vivre à l'hospice !

En 1534, l'espoir revient. Cardan obtient l'ancien poste de son père, conférencier en mathématiques à la fondation Piatti de Milan, poste qui lui laisse beaucoup de temps libre, ce qui lui permet de consulter quelques patients. Ses traitements médicaux ont beaucoup de succès (certains sont même jugés « miraculeux ») ce qui lui vaut une grande réputation. Mais il n'est toujours pas admis à l'université comme professeur (malgré des candidatures répétées en 1536 et 1537), ce qui lui fait dire : « Les choses qui donnent, de nos jours, le plus de réputation à un médecin sont ses façons, ses domestiques, sa calèche, ses vêtements, son élégance, le tout exposé de façon artificielle et insipide. »

Ce n'est qu'en 1539, grâce à la pression de ses admirateurs, que l'université lui ouvre ses portes. La même année il publie ses deux premiers livres mathématiques dont le second est un ouvrage d'arithmétique : *Practica arithmetice et mensurandi singularis* (Pratique arithmétique et mesures simples) particulièrement prometteur. Ces parutions marquent le début d'une période prolifique où Cardan écrit sur différents domaines : médecine, philosophie, astronomie, théologie et, bien sûr, mathématiques.

En 1539, Cardan entre en contact avec Tartaglia, dont la renommée est grandissante car il a réussi à résoudre certaines équations du troisième degré. Cardan demande à Tartaglia de lui indiquer sa méthode, ce dernier accepte à la condition que notre mathématicien jure de ne pas la révéler. Cardan accepte et écrit : « Je vous jure, par les Saintes Évangiles de Dieu, et en tant que vrai homme d'honneur, non seulement de ne jamais éditer vos découvertes, si vous me les enseignez, mais je vous promets également, et je mets en gage

ma foi de vrai chrétien, de les noter de façon codée, de façon qu'après que ma mort personne ne puisse les comprendre. » Toutefois, apprenant que del Ferro avait déjà résolu l'équation, il publie la méthode, ce qui rend Tartaglia furieux. Une période d'intenses travaux mathématiques, qui durera six ans, commence alors pour Cardan : il se penche sur la résolution des équations de degré 3 et de degré 4.

En 1540, Cardan démissionne de son poste de conférencier en mathématiques à la fondation Piatti et le transmet à son assistant Ferrari. De 1540 à 1542 il abandonne ses recherches et se consacre entièrement au jeu d'échec. Dans les années 1543-1552, il enseigne la médecine aux universités de Milan et de Pavie. C'est pendant cette période, en 1545, qu'il publie son grand ouvrage *Ars magna sive de regulis algebraicis liber unus*. Sa femme Lucia meurt en 1546, mais, entièrement focalisé par sa renommée grandissante, il ne semble pas en être très affecté. Il devient recteur de la faculté de médecine et acquiert la réputation d'être le plus grand médecin de son temps. Il reçoit beaucoup d'offres de la part de diverses personnalités d'Europe qui souhaiteraient s'assurer ses services, mais il ne répondra favorablement, en 1552, qu'à l'archevêque de Saint Andrews, John Hamilton, qui souffre d'asthme depuis dix ans. Les soins de Cardan, qui reste auprès de l'homme d'église du 29 juin au 13 septembre, sont efficaces et l'archevêque sera complètement rétabli, dit-on, deux ans plus tard.

À son retour d'Écosse, Cardan est nommé professeur de médecine à l'université de Pavie, il a beaucoup de patients fortunés, et lui-même devient riche. Mais, alors qu'il est au faite de sa renommée un malheur familial survient. Son fils aîné, Giambatista, médecin également depuis 1557, avait épousé secrètement Brandonia di Seroni, une fille que Cardan décrit comme étant « une femme sans valeur et sans scrupule. » Brandonia raille publiquement son mari pour ne pas être le père de leurs trois enfants. Ces railleries conduisent Giambatista à empoisonner sa femme et après son arrestation il avoue son crime. Finalement il est exécuté le 13 avril 1560. Cardan ne se remettra jamais de cette terrible épreuve. Il décide de quitter la région et devient professeur de médecine à Bologne.

Cardan rencontre également des problèmes avec son plus jeune fils Aldo. Ce dernier est joueur et fréquente des individus douteux. Dans son autobiographie, Cardan énumère les quatre grandes tristesses de sa vie : « La première a été mon mariage ; la seconde la mort cruelle de mon fils ; la troisième l'emprisonnement ; la quatrième le caractère indigne de mon plus jeune fils. » En 1570, notre mathématicien fait l'horoscope de Jésus-Christ et écrit un livre à la gloire de Néron. Cela lui vaut d'être emprisonné pour hérésie. Il est toutefois traité avec une certaine clémence et ne reste que quelques mois en prison. Par contre, on lui interdit d'occuper un poste universitaire et de publier ses travaux. À sa libération, Cardan se rend à Rome où il est

reçu chaudement. Il est nommé membre du collège des médecins de cette ville et le pape lui accorde même une pension.

En plus de ses contributions majeures à l'algèbre, Cardan apporte également des contributions importantes aux probabilités, à l'hydrodynamique, à la mécanique et à la géologie. Son livre *Liber de Ludo Aleae*, probablement écrit en 1563, mais publié cent ans plus tard, est une étude systématique du calcul des probabilités, qui devance Pascal ou Fermat de près d'un siècle. Il est également crédité de l'invention du cardan objet auquel il a laissé son nom. Il est également l'auteur de deux encyclopédies « Contenant un peu de tout, de la cosmologie à la construction des machines, de l'utilité des sciences naturelles à la mauvaise influence des démons, des lois de la mécanique à la cryptographie. C'est une mine des faits, vrais ou imaginaires, de notes sur l'état des sciences, de superstitions, de technologie, d'alchimie et de diverses branches de l'occulte. »

On raconte que Cardan a prédit de façon exacte la date de sa mort – mais certains disent que c'est parce qu'il s'est suicidé ! –, toujours est-il qu'il est mort à Rome le 21 septembre 1576.

1.5. Principaux ouvrages de Cardan et références

- Cardano, Girolamo, *Opera omnia : tam hactenus excusa, cura Caroli Sponii*, Lugduni, Sumptibus Ioannis Antonii Hvgvetan, & Marci Antonii Ravard, 1663.
- *Practica arithmetice et mensurandi singularis*, Milan, 1539.
- *Artis magnae, sive de regulis algebraicis liber unus*, Nuremberg, 1545.
- *De subtilitate liber XXI*, Nuremberg, 1550.
- *De rerum varietate libri XVII*, Basel, 1557.

Références

- M. Gliozzi, *Cardano, Girolamo*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 3, p. 64-67, 1970.
- C. B. Boyer, *Cardan and the Pascal triangle*, Amer. Math. Monthly, 57 (1950), 387-390.
- J. Cardano, *Autobiography*, New York, 1930.
- O. Ore, *Cardano, the Gambling Scholar*, Princeton N. J., 1953.
- M. Fierz, *Girolamo Cardano physician, natural philosopher, mathematician, astrologer, and interpreter of dreams*, Boston, 1983.

1.6. Biographie de Galilée (1564-1642)

Galileo Galilei naît à Pise le 15 février 1564, trois jours après la mort de Michel-Ange, dans une famille de petite noblesse. Son père, Vincenzo Galilei, est musicien. Sa mère, Giulia degli Ammannati s'occupe de ses six enfants dont Galileo est l'aîné. La famille Galilei est noble sans être riche. Au début des années 1570 elle s'établit à Florence. Vincenzo oriente son fils vers des études de médecine ; c'est ainsi que Galileo se retrouve à l'Université de Pise. Pour ce qui concerne la philosophie naturelle (*i.e.* les sciences) l'enseignement d'alors doit beaucoup à la philosophie aristotélicienne. C'est à Pise, alors qu'il n'a que 19 ans, que Galilée découvre l'isochronisme des pendules. Son élève et biographe Viviani (1622-1663) indique qu'« il lui vint à l'esprit d'observer les mouvements d'une lampe [le lustre de la cathédrale de Pise] écartée de la verticale, pour voir si, par hasard, les temps d'aller-retour de celle-ci n'étaient pas les mêmes pour les grands arcs, les arcs moyens et les petits ».

En 1585, après quatre ans d'études, Galilée revient à Florence sans diplôme ! Il complète alors sa formation en étudiant *Les Éléments* d'Euclide avec l'aide du mathématicien Ostilio Ricci. La physique le passionne également, mais il préfère Archimède à Aristote. À 22 ans, en 1586, il invente la balance hydrostatique. Entre 1585 et 1589, il donne à Florence et à Sienne des leçons particulières de mathématiques. Lors d'un voyage à Rome, il rencontre le mathématicien Christopher Clavius, dont on sait que les livres, très pédagogiques, avaient inspiré Descartes. Après diverses tentatives, Galilée est nommé, en 1589, professeur de mathématiques à Pise. En 1592 il devient professeur de mathématiques à l'Université de Padoue (situé en République de Venise) où il restera jusqu'en 1610. Il enseigne la géométrie et l'astronomie, mais donne également des leçons particulières touchant à l'arithmétique, les fortifications ou l'optique. En parallèle de ses activités, il rédige un premier traité de mécanique, *Del Motu*, qui ne sera publié que trois cents ans plus tard. De cette époque datent également ses recherches sur la chute des corps. Il utilise pour cela des plans inclinés « soigneusement polis » et des billes « parfaitement rondes et lisses » et grâce à des expériences répétées et méthodiques, il montre que les espaces parcourus sont proportionnels au carré du temps.

Galilée invente le thermoscope, ancêtre du thermomètre (objet maintenant connu sous le nom de « thermomètre de Galilée »), probablement aux alentours des années 1606-1607. Deux ans plus tard, il entreprend la construction de sa lunette astronomique et fait au début de 1610 une série de découvertes qu'il relate dans *Sidereus Nuncius*. Il prend soin de transmettre un exemplaire de son traité au grand-duc de Toscane, Cosme II de Medicis, et obtient peu de temps après le titre de « Premier mathématicien et philosophe du grand-duc » ce qui lui permet de se consacrer à son projet

de construction d'une « science entièrement nouvelle ». Il quitte Padoue et s'installe à Florence. Au printemps 1611, Galilée entre à l'Accademia dei Lincei (Académie du Lynx) où il est admis par son fondateur Federico Cesi, Marchese di Monticelli. Il publie en 1612 son *Discours sur les corps flottants*, qui représente une importante contribution à l'hydrostatique.

À la fin de l'année 1614 des nuages s'amoncellent dans le ciel de Galilée. Un frère dominicain, Tommasco Caccini, fait à Florence un prêche contre Galilée et les mathématiques qu'il utilise pour la défense du système copernicien. Peu après, un autre frère dominicain, Niccolo Lorini, transmet à l'Inquisition une plainte contre Galilée pour les mêmes raisons. En avril 1615, le Cardinal Bellarmín lui suggère de ne traiter le système copernicien qu'en tant qu'hypothèse. En décembre de la même année, Galilée se rend à Rome pour défendre ses idées coperniciennes. Cela lui vaut quelques démêlés avec l'Église, en particulier le pape Paul V demande au Cardinal Bellarmín de dire à Galilée qu'il doit cesser de propager et de soutenir les idées de Copernic. Le 5 mars 1616, l'œuvre de Copernic est mise à l'Index, mais Galilée n'est pas mentionné dans le décret d'interdiction. Néanmoins, il ne renonce pas à défendre ses idées. L'apparition de trois comètes en octobre et novembre 1618 relance le débat. Galilée publie en 1623 *Il sagggiatore* (l'essayeur) dans lequel il mène une réflexion profonde et moderne sur la méthode de la science. Les censeurs romains donnent l'autorisation de publication. Les choses semblent donc s'améliorer, d'autant qu'au mois d'août de la même année, le Cardinal Barberini devient pape et prend le nom d'Urbain VIII. Barberini connaît bien Galilée depuis 1610 et admire son intelligence. Il a d'ailleurs déjà eu, à plusieurs reprises, l'occasion de le soutenir. Galilée dédicace donc son livre *Il sagggiatore* au nouveau pape. En avril 1624, notre physicien se rend à Rome où il obtient six audiences avec Urbain VIII. Ce dernier lui indique qu'il peut parler du système copernicien, tant que celui-ci est présenté comme une simple hypothèse mathématique.

Par ailleurs, Galilée présente en 1624 un microscope aux membres de l'Académie du Lynx et prolonge sa théorie des marées, travail qui est l'amorce de son livre phare : *Dialogue sur les deux principaux systèmes du monde*.

En avril 1630, Galilée achève ce livre. Il le présente au Vatican et obtient la permission de le publier sous l'égide de l'Académie du Lynx. L'impression du livre se termine en février 1632 ; toutefois une commission spéciale est chargée d'examiner l'ouvrage avant de procéder à sa distribution. Mais le rapport de cette commission vaut à Galilée de comparaître devant le tribunal de l'Inquisition à Rome. Grâce au grand-duc Ferdinand II de Médicis, et avec l'autorisation du pape, Galilée séjourne à Rome dans la résidence de l'Ambassadeur de Toscane ; toutefois les contacts avec l'extérieur lui sont interdits. Le procès de Galilée débute le 12 avril 1633. Cette fois Galilée ne persiste pas, il dit lui-même : « Je suis ici pour me soumettre ». Le 30 avril,

Galilée admet que sa défense du système copernicien dans son Dialogue est trop forte. Le procès se termine le mercredi 22 juin 1633 par l'abjuration de Galilée qui doit déclarer à genoux « [...] voulant chasser de l'esprit de vos éminences et de tout chrétien fidèle, ce grave soupçon, justement conçu contre moi, d'un cœur pur et d'une foi sincère, j'abjure, je maudis et je déteste les erreurs et hérésies susdites et en général toute autre erreur ou hérésie contraire à la Sainte Église »¹. Il est condamné à la résidence surveillée, et après un séjour à Rome dans la résidence de l'Ambassadeur de Toscane, puis à Sienne dans celle de l'archevêque de cette ville, Galilée est autorisé, en décembre, à revenir dans sa villa d'Arcetri, près de Florence (en régime de résidence surveillée). Cela n'empêche toutefois pas son travail : dès le mois de juillet 1633 (soit quelques jours après la fin de son procès), Galilée se met à la rédaction d'un nouveau livre : *Discours sur deux nouvelles sciences*. En mai 1636 l'éditeur hollandais Louis Elsevier lui rend visite dans sa résidence d'Arcetri et lui propose d'éditer son livre à Leyde. Mais, peu de temps après, Galilée perd la vue. Il écrit en juillet 1637 à son ami Diodati : « Galilée, votre cher ami et serviteur, est depuis un mois totalement et irrémédiablement aveugle ». En juillet 1638, son livre, dont le titre complet est : *Discours et démonstrations mathématiques concernant deux nouvelles sciences touchant la mécanique et les mouvements locaux*, est édité en Hollande. C'est dans ce livre célèbre, souvent considéré comme le premier livre de physique moderne, que Galilée analyse le problème de la chute des corps en détail avec une précision qui montre qu'il a réalisé de nombreuses expériences. C'est grâce à ses expériences avec un plan incliné qu'il a pu faire ressortir les lois essentielles.

En 1639, il demande sa mise en liberté, mais le pape la lui refuse. Viviani (1622-1703), qui a alors 17 ans, lui sert de secrétaire, devient son élève et est l'un des premiers à pratiquer la nouvelle physique. Galilée accueille deux ans plus tard un second élève qui n'est autre que Evangelista Torricelli (1608-1647).

Après deux mois de maladie, Galilée meurt le 8 janvier 1642, à l'âge de soixante-dix-sept ans. La même année dans un petit village proche de Londres naît Isaac Newton...

L'apport scientifique de Galilée

Cet apport est immense. Galilée est le fondateur de la mécanique, science générale du mouvement. Il peut être considéré comme le père de la science moderne du fait qu'il n'accepte comme critères de vérité que l'expérience et la réflexion. Il y associe l'utilisation des mathématiques, en quelque sorte garantes de l'objectivité. Il expose ce point très clairement dans *Il sagggiatore* :

1. La légende, construite plus tard, veut qu'« au moment de se relever, agité par le remords d'avoir fait un faux serment, les yeux baissés vers la terre, [on prétend] qu'il dit, en frappant du pied, "*Eppur si muove*" [et pourtant elle tourne] » (dictionnaire historique de 1789).

« La philosophie [c'est-à-dire la science] est écrite dans ce très grand livre qui se tient constamment ouvert sous nos yeux, l'univers, et qui ne peut se comprendre que si l'on a préalablement appris à en comprendre la langue et à en connaître les caractères employés pour l'écrire. Ce livre est écrit dans la langue mathématique, ses caractères sont des triangles, des cercles et d'autres figures géométriques, sans l'intermédiaire desquels il est impossible d'en comprendre humainement un seul mot, et sans lesquels on ne fait qu'errer vainement dans un labyrinthe obscur ».

Galilée fait une autre série de prodigieuses découvertes astronomiques à l'aide de sa lunette¹. Il fabrique sa première lunette et la présente aux membres du Sénat de Venise le 21 août 1609. Il en construit une autre de plus forte puissance avec laquelle il observe et dessine la Lune du 30 novembre au 18 décembre 1609. Il découvre que la Lune n'est pas une sphère parfaite mais qu'elle possède un relief accidenté. Du 19 décembre 1609 au 6 janvier 1610, il observe la Voie Lactée et la décrit comme « une masse d'étoiles innombrables assemblées par grappes » alors que l'on pensait jusqu'alors qu'il s'agissait d'une « exhalaison » de l'atmosphère terrestre. Le 7 janvier 1610, à une heure du matin, il remarque que Jupiter est accompagnée de « trois étoiles » qu'il appellera astres médicéens en l'honneur des Médicis, ses protecteurs de Florence. Ces « astres » seront baptisés en 1611 satellites de Jupiter par Kepler. Galilée suit alors Jupiter, découvre sa quatrième lune le 13 janvier, et rédige du 16 janvier au 12 mars son *Sidereus Nuncius*. Le 1^{er} juillet 1610 il observe « la planète la plus lointaine, à triple corps », il s'agit de Saturne, dont Huygens comprendra trente-neuf ans plus tard la véritable constitution avec ses anneaux. En septembre 1610, Galilée découvre les phases de Vénus, analogues à celle de la Lune. Cette découverte est importante car Galilée comprend que l'existence de ces phases de Vénus représente une confirmation éclatante de la justesse du système de Copernic. Galilée observe également la présence de taches sur le Soleil et, en les suivant, comprend que, comme la Terre, le Soleil tourne autour de lui-même. Cette moisson de découvertes spectaculaires fait dire à Mersenne « Galilée ! Qui sera capable de dénombrer ses découvertes ? Avec sa seule lunette, il a presque trouvé plus de choses qu'on en connaissait déjà ».

Galilée, père de la mécanique et brillant astronome, adopte donc la cosmologie de Copernic et construit une nouvelle physique cohérente avec cette cosmologie. Il montre que contrairement au dogme d'Aristote, le ciel, comme la Terre, est altérable. En 1618, trois nouvelles comètes apparaissent. Un an plus tard, Galilée publie le *Discours sur les comètes* pour dénoncer une interprétation de Grassi qui reprenait le système de Tycho Brahé, hybride entre celui de Ptolémée et de Copernic. Les découvertes de Galilée sur les phases

1. Contrairement à une idée reçue, Galilée n'a pas inventé la lunette mais l'a considérablement améliorée et surtout l'a utilisée pour scruter le ciel.

de Vénus et les taches solaires abondent encore dans le sens d'un ciel « altérable ». La mise en cohérence de ce que l'on pourrait appeler la cosmologie physique de Galilée apparaît de toute évidence dans son *Dialogue sur les deux principaux systèmes du monde*, celui de Ptolémée et celui de Copernic.

1.7. Principaux ouvrages de Galilée et références

- *Sidereus Nuncius Magna, longeque admirabilia Spectacula pandens, suspiciendaque proponens unicuique, praesertim vero Philosophis, atque Astronomis...*, Venice, T. Baglionus, 1610.
- *Il saggiatore...*, Roma, appresso Giacomo Mascardi, 1623.
- *Dialogo di Galileo Galilei massimi sistemi del mondo tolemaico, e Copernicano*, Fiorenza, Gio Battista Landini, 1632.
- *Discorsi e dimostrazioni matematiche intorno a due nuove scienze attenenti alla meccanica & movimenti locali*, Leida, Elsevirii, 1638.

Références

- S. Drake, *Dictionary of Scientific Biography*, Charles Scribner's sons, New York, vol. 5, p. 237-249, 1970.
- S. Drake, *Galileo*, Oxford, 1980.
- A. Koyré, *Études d'histoire de la pensée scientifique*, Paris, 1966.
- A. Koyré, *Études Galiléennes*, Paris, 1939.
- G. Favaro, *Bibliographia Galileana*, Venice, 1942.

La correspondance entre Pascal et Fermat et le *Traité du triangle arithmétique* de Pascal

2.1. Pascal, Fermat et le problème des partis

C'est la correspondance entre Pascal et Fermat publiée à Toulouse dans les *Varia Opera* de Fermat en 1679, d'une part, et le chapitre intitulé *l'usage du triangle arithmétique pour déterminer les partis qu'on doit faire entre deux joueurs qui jouent en plusieurs parties*, publié par Pascal dans son *Traité du triangle arithmétique*, d'autre part, qui font que l'on attribue à Pascal (en oubliant souvent Fermat) la paternité du calcul des probabilités.

Voici ce qu'a écrit Poisson¹, en 1837, au tout début de son principal ouvrage relatif aux probabilités :

« Un problème relatif aux jeux de hasard, proposé à un austère janséniste par un homme du monde, a été à l'origine du calcul des probabilités. Il avait pour objet de déterminer la proportion suivant laquelle l'enjeu doit être partagé entre les joueurs, lorsqu'ils conviennent de ne point achever la partie, et qu'il leur reste à prendre, pour la gagner, des nombres de points inégaux². Pascal en donna le premier la solution, mais pour le cas de deux joueurs seulement ; il fut ensuite résolu par Fermat, dans le cas général d'un nombre quelconque de joueurs. Toutefois, les géomètres du XVII^e siècle qui se sont occupés du calcul des probabilités, ne l'ont employé qu'à déterminer les chances de différents jeux de cette époque ; et ce n'est que dans le siècle suivant qu'il a pris toute son extension, et qu'il est devenu une des principales branches des mathématiques [...]. »

En fait, comme le souligne Poisson, les travaux de Pascal ne concernent que les jeux et le mot *probabilité* n'apparaît jamais dans ses écrits mathématiques.

1. Siméon-Denis Poisson, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*, p. 1. Paris, Bachelier, 1837.
2. Problème dit « des partis ».

Dans son *Adresse à l'académie Parisienne des mathématiciens*, Pascal, en 1654, annonçait qu'il rédigeait toute une série d'opuscules dont un sur les probabilités¹ :

« À la très illustre Académie parisienne de Mathématiciens

[...] Et puis un traité tout à fait nouveau, d'une matière absolument inexplorée jusqu'ici, savoir : la répartition du hasard dans les jeux qui lui sont soumis, ce qu'on appelle en français *faire les partis des jeux* ; la fortune incertaine y est si bien maîtrisée par un calcul équitable qu'à chacun des joueurs on assigne toujours exactement ce qui s'accorde avec la justice. Et c'est là certes ce qu'il faut d'autant plus chercher par le raisonnement, qu'il est moins possible d'être éclairé par l'expérience. En effet, les résultats du sort ambigu sont justement attribués à la contingence fortuite plutôt qu'à la nécessité naturelle. C'est pourquoi la question est restée incertaine jusqu'à aujourd'hui ; mais maintenant, demeurée rebelle à l'expérience, elle n'a pu échapper à l'empire de la raison. Et, grâce à la géométrie, nous l'avons réduite avec tant de sûreté à un art exact, qu'elle participe de sa certitude et déjà progresse audacieusement. Ainsi, joignant la rigueur des démonstrations de la science à l'incertitude du hasard, et conciliant ces choses en apparence contraires, elle peut, tirant son nom des deux, s'arroger à bon droit le titre stupéfiant de *Géométrie du hasard*. »

La correspondance sur le sujet des probabilités entre Pascal et Fermat date de 1654. Les lettres de Pascal furent éditées à Toulouse dans les *Varia Opera* de Fermat en 1679 et reproduites avec, en complément, quelques lettres inédites de Fermat dans les *Œuvres de Blaise Pascal* publiées à La Haye, chez Detune en 1789. Cette correspondance occupe les pages 412 à 454 du tome IV de cette édition. Le chevalier de Méré auquel, selon Pascal, « il ne manquait, pour être parfait, que d'être géomètre », et qui, « malgré ce grand défaut, était très bon esprit », avait rencontré des difficultés de jeu qu'il lui avait soumises. Pascal consulta Roberval puis Fermat qui confirma ses résultats. Cette correspondance entre Pascal et Fermat n'a donc eu, en fait, aucune influence sur les travaux de Huygens (rendus publics en 1657) puisqu'elle ne fut publiée que plus tard.

Les lettres échangées entre Pascal et Fermat sont essentiellement relatives au problème des partis, c'est-à-dire au problème du partage équitable des sommes mises en jeu, lorsque ce jeu est interrompu d'un commun accord entre les joueurs.

E. Coumet² a indiqué que Pacioli, Tartaglia et Cardan avaient déjà abordé le même problème sans arriver à des solutions satisfaisantes.

Cette correspondance entre Pascal et Fermat de 1654 comporte une lettre non datée de Fermat à Pascal, une lettre perdue du même au même dans

1. *Œuvres de Blaise Pascal*, tome IV, p. 408-411, La Haye, Detune, 1779. Le texte original est en Latin.

2. Ernest Coumet, *Le problème des partis avant Pascal*, Archives internationales d'histoire des sciences, n° 18, 1965, fascicule 72-73, pages 245-272.

laquelle Fermat a proposé une solution au problème des partis, une lettre de Pascal à Fermat du 29 juillet 1654, lettre bien connue où Pascal décrit sa méthode pour calculer les partis, et une autre lettre de Pascal du 24 août, dans laquelle il critique injustement la méthode des combinaisons de Fermat. Le 29 août Fermat écrit à Pascal, mais il n'a pas encore en mains la lettre de ce dernier en date du 29 juillet. Le 25 septembre 1654, Fermat répond à la lettre du 24 août et justifie sa méthode des combinaisons. Enfin, Pascal écrit à Fermat, le 27 octobre 1654 pour reconnaître la valeur des solutions de son correspondant.

Dans sa lettre à Fermat du 29 juillet 1654, Pascal commente le problème des dés posé par Antoine Gombaud, chevalier de Méré (1610-1685), à savoir : Combien de fois au moins faut-il lancer un dé pour obtenir un six avec une probabilité supérieure à $1/2$? Et combien de fois au moins faut-il lancer deux dés pour obtenir un double six (*sonnez*) avec une probabilité supérieure à $1/2$?

Pascal écrit :

« Si on entreprend de faire un six avec un dé, il y a avantage de l'entreprendre en 4, comme de 671 à 625. Si on entreprend de faire Sonnez avec deux dés, il y a désavantage de l'entreprendre en 24. Et néanmoins 24 est à 36, qui est le nombre de faces des deux dés, comme 4 est à 6, qui est le nombre de faces d'un dé. »

Pascal ajoute, en parlant de Méré :

« Voilà quel était son grand scandale, qui lui faisait dire hautement que les propositions n'étaient pas constantes, et que l'arithmétique se démentait. »

[Si on appelle p , la probabilité d'obtention et $q = 1 - p$, le nombre minimum de jets est n tel que

$$\frac{1 - q^n}{q^n} \geq 1 \quad \text{ou} \quad q^n \leq \frac{1}{2}.$$

Pour un dé $p = 1/6$, on trouve $n = 4$, pour 2 dés, $p = 1/36$ et $n = 25$.]¹

Voici un extrait de la lettre de Pascal à Fermat du 29 juillet 1654 où il donne sa méthode pour résoudre le problème des partis à l'aide d'un exemple², les solutions étant élaborées pas à pas :

« [...] Je reçus hier au soir, de la part de M. de Carcavi, votre lettre sur les partis, que j'admire si fort, que je ne puis vous le dire. [...] Votre méthode est très sûre, et c'est la première qui m'est venue à la pensée dans cette recherche. Mais parce que la peine des combinaisons est excessive, j'en ai trouvé un abrégé et proprement une autre méthode bien plus courte et plus nette que je voudrais vous dire ici en peu de mots [...]

1. Cf. A. Hald, *History of Probability and Statistics... before 1730*, Wiley, 1990, § 14-4.

2. *Œuvres de Blaise Pascal*, tome IV, p. 412-413, La Haye, Detune, 1779.

Voici à peu près comme je fais pour savoir la valeur de chacune des parties, quand deux joueurs jouent, par exemple, en trois parties, et chacun a mis 32 pistoles au jeu :

Posons que le premier en ait deux et l'autre une ; ils jouent maintenant une partie, dont le sort est tel que, si le premier la gagne, il gagne tout l'argent qui est au jeu, savoir, 64 pistoles ; si l'autre la gagne, ils sont deux parties à deux parties, et par conséquent, s'ils veulent se séparer, il faut qu'ils retirent chacun leur mise, savoir, chacun 32 pistoles.

Considérez donc, Monsieur, que si le premier gagne, il lui appartient 64 ; s'il perd, il lui appartient 32. Donc s'ils veulent ne point hasarder cette partie et se séparer sans la jouer, le premier doit dire : "Je suis sûr d'avoir 32 pistoles, car la perte même me les donne ; mais pour les 32 autres, peut-être je les aurai, peut-être vous les aurez ; le hasard est égal ; partageons donc ces 32 pistoles par la moitié et me donnez, outre cela, mes 32 qui me sont sûres." Il aura donc 48 pistoles et l'autre 16.

Posons maintenant que le premier ait deux parties et l'autre point, et ils commencent à jouer une partie. Le sort de cette partie est tel que, si le premier la gagne, il tire tout l'argent, 64 pistoles ; si l'autre la gagne, les voilà revenus au cas précédent, auquel le premier aura deux parties et l'autre une. Or, nous avons déjà montré qu'en ce cas, il appartient à celui qui a les deux parties, 48 pistoles : donc, s'ils veulent ne point jouer cette partie, il doit dire ainsi : "Si je la gagne, je gagnerai tout, qui est 64 ; si je la perds, il m'appartiendra légitimement 48 : donc donnez-moi les 48 qui me sont certaines au cas même que je perde, et partageons les 16 autres par la moitié, puisqu'il y a autant de hasard que vous les gagniez comme moi." Ainsi il y aura 48 et 8, qui sont 56 pistoles.

Posons enfin que le premier n'ait qu'une partie et l'autre point. Vous voyez, Monsieur, que, s'ils commencent une partie nouvelle, le sort en est tel que, si le premier la gagne, il aura deux parties à point, et partant, par le cas précédent, il appartient 56 ; s'il la perd, ils sont partie à partie : donc il lui appartient 32 pistoles. Donc il doit dire : "Si vous voulez ne la pas jouer, donnez-moi 32 pistoles qui me sont sûres, et partageons le reste de 56 par la moitié. De 56 ôtez 32, reste 24 ; partagez donc 24 par la moitié, prenez-en 12, et moi 12, qui avec 32, font 44". »

Pascal a examiné également la solution du problème des partis à l'aide du triangle arithmétique (cf. la fin du présent chapitre).

La lettre de Pascal à Fermat du 24 août 1654 est, elle aussi, relative au problème des partis. Pascal y expose sa vision d'une solution par la méthode des combinaisons et prétend, à tort, que la solution de Fermat (qui avait été communiquée à Pascal par Fermat dans une lettre aujourd'hui perdue) est correcte quand il y a deux joueurs mais n'est pas valable dès que le nombre de joueurs est plus élevé.

Voici un extrait de cette lettre :

« Monsieur,

Je ne pus vous ouvrir ma pensée entière touchant les partis de plusieurs joueurs par l'ordinaire passé, et même j'ai quelque répugnance à le faire, de peur qu'en ceci cette admirable convenance qui était entre nous, et qui m'était si chère, ne

commence à se démentir, car je crains que nous ne soyons de différents avis sur ce sujet. Je vous veux ouvrir toutes mes raisons, et vous me ferez la grâce de me redresser, si j'erre, ou de m'affermir, si j'ai bien rencontré. Je vous le demande tout de bon et sincèrement, car je ne me tiendrai pour certain que quand vous serez de mon côté.

Quand il n'y a que deux joueurs, votre méthode, qui procède par les combinaisons, est très sûre ; mais quand il y en a trois, je crois avoir démonstration qu'elle est mal juste, si ce n'est que vous y procédiez de quelque autre manière que je n'entends pas. Mais la méthode que je vous ai ouverte, et dont je me sers partout, est commune à toutes les conditions imaginables de toutes sortes de partis, au lieu que celle des combinaisons (dont je ne me sers qu'aux rencontres particulières où elle est plus courte que la générale) n'est bonne qu'en ces seules occasions et non pas aux autres.

Je suis sûr que je me donnerai à entendre, mais il me faudra un peu de discours, et à vous un peu de patience.

Voici comment vous procédez quand il y a deux joueurs :

Si deux joueurs, jouant en plusieurs parties, se trouvent en cet état qu'il manque deux parties au premier et trois au second, pour trouver le parti il faut, dites-vous, voir en combien de parties le jeu sera décidé absolument.

Il est aisé de supputer que ce sera en quatre parties ; d'où vous concluez qu'il faut voir combien quatre parties se combinent entre deux joueurs ; et voir combien il y a de combinaisons pour faire gagner le premier, et combien pour le second, et partager l'argent suivant cette proportion. J'eusse eu peine à entendre ce discours-là, si je ne l'eusse su de moi-même auparavant ; aussi vous l'aviez écrit dans cette pensée. Donc, pour voir combien quatre parties se combinent entre deux joueurs, il faut imaginer qu'ils jouent avec un dé à deux faces (puisque'ils ne sont que deux joueurs), comme à croix et pile, et qu'ils jettent quatre de ces dés (parce qu'ils jouent en quatre parties) ; et maintenant il faut voir combien ces dés peuvent avoir d'assiettes différentes. Cela est aisé à supputer : ils peuvent en avoir seize, qui est le second degré de quatre, c'est-à-dire le carré. Car figurons-nous qu'une des faces est marquée *a*, favorable au premier joueur, et l'autre *b*, favorable au second ; donc ces quatre dés peuvent s'asseoir sur une de ces seize assiettes : *a a a a... b b b b*.

<i>a a a a</i>	<i>a b a a</i>	<i>b a a a</i>	<i>b b a a</i>
<i>a a a b</i>	<i>a b a b</i>	<i>b a a b</i>	<i>b b a b</i>
<i>a a b a</i>	<i>a b b a</i>	<i>b a b a</i>	<i>b b b a</i>
<i>a a b b</i>	<i>a b b b</i>	<i>b a b b</i>	<i>b b b b</i>

Et parce qu'il manque deux parties au premier joueur, toutes les faces qui ont deux *a* le font gagner : donc il en a 11 pour lui ; et parce qu'il y manque trois parties au second, toutes les faces où il y a trois *b* le peuvent faire gagner : donc il y en a 5. Donc il faut qu'ils partagent la somme comme 11 à 5.

Voilà votre méthode quand il y a deux joueurs. Sur quoi vous dites que, s'il y en a davantage, il ne sera pas difficile de faire les partis par la même méthode. Sur cela, Monsieur, j'ai à vous dire que ce parti pour deux joueurs, fondé sur les combinaisons, est très juste et très bon ; mais que, s'il y a plus de deux joueurs, il ne sera pas toujours juste, et je vous dirai la raison de cette différence.

Je communiquai votre méthode à nos Messieurs, sur quoi M. de Roberval me fit cette objection :

Que c'est à tort que l'on prend l'art de faire le parti sur la supposition qu'on joue en quatre parties, vu que, quand il manque deux parties à l'un et trois à l'autre, il n'est pas de nécessité que l'on joue quatre parties, pouvant arriver qu'on n'en jouera que deux ou trois, ou, à la vérité, peut-être quatre ; et ainsi qu'il ne voyait pas pourquoi on prétendait de faire le parti juste sur une condition feinte qu'on jouera quatre parties, vu que la condition naturelle du jeu est qu'on ne jouera plus dès que l'un des joueurs aura gagné, et qu'au moins, si cela n'était faux, cela n'était pas démontré, de sorte qu'il avait quelque soupçon que nous avions fait un paralogisme.

Je lui répondis que je ne me fondais pas tant sur cette méthode des combinaisons, laquelle véritablement n'est pas en son lieu en cette occasion, comme sur mon autre méthode universelle, à qui rien n'échappe, et qui porte sa démonstration avec soi, qui trouve le même parti précisément que celle des combinaisons ; et de plus je lui démontrai la vérité du parti entre deux joueurs par les combinaisons en cette sorte :

N'est-il pas vrai que, si deux joueurs, se trouvant en cet état de l'hypothèse qu'il manque deux parties à l'un et trois à l'autre, conviennent maintenant de gré à gré qu'on joue quatre parties complètes, c'est-à-dire qu'on jette les quatre dés à deux faces tous à la fois, n'est-il pas vrai dis-je, que s'ils ont délibéré de jouer les quatre parties, le parti doit être tel que nous avons dit, suivant la multitude des assiettes favorables à chacun ?

Il en demeura d'accord ; et cela en effet est démonstratif ; mais il niait que la même chose subsistât, en ne s'astreignant pas à jouer les quatre parties. Je lui dis donc ainsi :

N'est-il pas clair que les mêmes joueurs, n'étant pas astreints à jouer quatre parties, mais voulant quitter le jeu dès que l'un aurait atteint son nombre, peuvent, sans dommage ni avantage, s'astreindre à jouer les quatre parties entières, et que cette convention ne change en aucune manière leur condition ? Car, si le premier gagne les deux premières parties de quatre, et qu'ainsi il ait gagné, refusera-t-il de jouer encore deux parties, vu que, s'il les gagne, il n'a pas mieux gagné, et s'il les perd, il n'a pas moins gagné, car ces deux que l'autre a gagnées ne lui suffisent pas, puisqu'il lui en faut trois, et ainsi il n'y a pas assez de quatre parties pour faire qu'ils puissent tous deux atteindre le nombre qui leur manque.

Certainement il est aisé de considérer qu'il est absolument égal et indifférent à l'un et à l'autre de jouer en la condition naturelle à leur jeu, qui est de finir dès qu'un aura son compte, ou de jouer les quatre parties entières ; donc, puisque ces deux conditions sont égales et indifférentes, le parti doit être tout pareil en l'une et en l'autre. Or, il est juste quand ils sont obligés de jouer quatre parties, comme je l'ai montré ; donc il est juste aussi en l'autre cas.

Voilà comment je le démontrai : et si vous y prenez garde, cette démonstration est fondée sur l'égalité des deux conditions, vraie et feinte, à l'égard de deux joueurs, et qu'en l'une et l'autre un même gagnera toujours ; et si l'un gagne ou perd en l'une, il gagnera ou perdra en l'autre, et jamais deux n'auront leur compte.

VARIA OPERA
MATHEMATICA
D. PETRI DE FERMAT,
SENATORIS TOLOSANI.

*Accesserunt selectæ quædam ejusdem Epistolæ, vel
ad ipsum à plerisque doctissimis viris Gallicè, Latinè,
vel Italicè, de rebus ad Mathematicas disciplinas,
aut Physicam pertinentibus scriptæ.*



TOLOSÆ,

Apud JOANNEM PECH, Conisionum Fuxensium Typographum, juxta
Collegium PP. Societatis JESU.

M. DC. LXXIX.

Suivons la même pointe pour trois joueurs, et posons qu'il manque une partie au premier, qu'il en manque deux au second et deux au troisième. Pour faire le parti suivant la même méthode des combinaisons, il faut chercher d'abord en combien de parties le jeu sera décidé, comme nous avons fait quand il y avait deux joueurs : ce sera en trois, car ils ne sauraient jouer trois parties sans que la décision soit arrivée nécessairement.

Il faut voir maintenant combien trois parties se combinent entre trois joueurs ; et combien il y en a de favorables à l'un, combien à l'autre, et combien au dernier ; et, suivant cette proportion, distribuer l'argent, de même qu'on a fait en l'hypothèse de deux joueurs.

Pour voir combien il y a de combinaisons en tout, cela est aisé : c'est la troisième puissance de 3, c'est-à-dire son cube 27. Car si on jette trois dés à la fois (puisque'il faut jouer trois parties) qui aient chacun, trois faces (puisque'il y a trois joueurs), l'une marquée *a* favorable au premier, l'autre *b* pour le second, l'autre *c* pour le troisième, il est manifeste que ces trois dés jetés ensemble peuvent s'asseoir sur 27 assiettes différentes, savoir :

<i>a a a</i>	<i>b a a</i>	<i>c a a</i>
<i>a a b</i>	<i>b a b</i>	<i>c a b</i>
<i>a a c</i>	<i>b a c</i>	<i>c a c</i>
<i>a b a</i>	<i>b b a</i>	<i>c b a</i>
<i>a b b</i>	<i>b b b</i>	<i>c b b</i>
<i>a b c</i>	<i>b b c</i>	<i>c b c</i>
<i>a c a</i>	<i>b c a</i>	<i>c c a</i>
<i>a c b</i>	<i>b c b</i>	<i>c c b</i>
<i>a c c</i>	<i>b c c</i>	<i>c c c</i>

Or, il ne manque qu'une partie au premier : donc toutes les assiettes où il y a un *a* sont pour lui ; donc il y en a 19.

Il manque deux parties au second : donc toutes les assiettes où il y a deux *b* sont pour lui ; donc il y en a 7.

Il manque deux parties au troisième : donc toutes les assiettes où il y a deux *c* sont pour lui ; donc il y en a 7.

Si de là on concluait qu'il faudrait donner à chacun selon la proportion de 19, 7, 7, on se tromperait trop grossièrement, et je n'ai garde de croire que vous le fassiez ainsi ; car il y a quelques faces favorables au premier et au second tout ensemble, comme *abb*, car le premier y trouve un *a* qu'il lui faut, et le second deux *b* qui lui manquent ; ainsi *a c c* est pour le premier et le troisième.

Donc il ne faut pas compter ces faces qui sont communes à deux comme valant la somme entière à chacun, mais seulement la moitié. Car, s'il arrivait l'assiette *a c c*, le premier et le troisième auraient même droit à la somme, ayant chacun leur compte ; donc ils partageraient l'argent par la moitié. Mais s'il arrive l'assiette *a b b*, le premier gagne seul. Il faut donc faire la supputation ainsi :

Il y a 13 assiettes qui donnent l'entier au premier, et 6 qui lui donnent la moitié, et 8 qui ne lui donnent rien ; donc, si la somme entière est une pistole, il y a 13 faces qui lui valent chacune une pistole, il y a 6 faces qui lui valent chacune une 1/2 pistole, et 8 qui ne valent rien.

Donc, en cas de parti, il faut multiplier :

13 par une pistole, qui font..13
6 par une demi, qui font.....3
8 par zéro, qui font.....0

Somme ... 27	Somme... 16

et diviser la somme des valeurs, 16, par la somme des assiettes, 27, qui fait la fraction $16/27$ qui est ce qui appartient au premier en cas de parti, savoir 16 pistoles de 27.

Le parti du second et du troisième joueur se trouvera de même :

Il y a 4 assiettes qui lui valent une pistole :

multipliez 4

Il y a 3 assiettes qui lui valent $1/2$ pistole :

multipliez.....1,5

Et 20 assiettes qui ne lui valent rien 0

Somme 27	Somme 5,5

Donc il appartient au second joueur 5,5 pistoles sur 27, et autant au troisième, et ces trois sommes, 5,5 ; 5,5 et 16, étant jointes, font les 27.

Voilà, ce me semble, de quelle manière ; il faudrait faire les partis par les combinaisons suivant votre méthode, si ce n'est que vous ayez quelque autre chose sur ce sujet que je ne puis savoir. Mais, si je ne me trompe, ce parti est mal juste.

La raison en est qu'on suppose une chose fausse, qui est qu'on joue en trois parties infailliblement, au lieu que la condition naturelle de ce jeu-là est qu'on ne joue que jusqu'à ce qu'un des joueurs ait atteint le nombre de parties qui lui manque, auquel cas le jeu cesse.

Ce n'est pas qu'il ne puisse arriver qu'on joue trois parties ; mais il peut arriver aussi qu'on n'en jouera qu'une ou deux, et rien de nécessité.

Mais d'où vient, dira-t-on, qu'il n'est pas permis de faire en cette rencontre la même supposition feinte que quand il y avait deux joueurs ? En voici la raison :

Dans la condition véritable de ces trois joueurs, il n'y en a qu'un qui peut gagner, car la condition est que, dès qu'un a gagné, le jeu cesse. Mais, en la condition feinte, deux peuvent atteindre le nombre de leurs parties : savoir, si le premier en gagne une qui lui manque, et un des autres deux qui lui manquent ; car ils n'auront joué que trois parties : au lieu que, quand il n'y avait que deux joueurs, la condition feinte et la véritable convenaient pour les avantages des joueurs en tout ; et c'est ce qui met l'extrême différence entre la condition feinte et la véritable.

Que si les joueurs, se trouvant en l'état de l'hypothèse, c'est-à-dire s'il manque une partie au premier et deux au second et deux au troisième, veulent maintenant de gré à gré et conviennent de cette condition, qu'on jouera trois parties complètes, et que ceux qui auront atteint le nombre qui leur manque prendront la somme entière s'ils se trouvent seuls qui l'aient atteint, ou, s'il se trouve que deux l'aient atteint, qu'ils la partageront également : en ce cas, le parti se doit faire comme je viens de le donner, que le premier ait 16, le second 5,5 ; le troisième 5,5 de 27 pistoles, et cela porte sa démonstration de soi-même, en supposant cette condition ainsi.

Mais s'ils jouent simplement à condition, non pas qu'on joue nécessairement trois parties, mais seulement jusqu'à ce que l'un d'entre eux ait atteint ses

parties, et qu'alors le jeu cesse sans donner moyen à un autre d'y arriver, alors il appartient au premier 17 pistoles, au second 5, au troisième 5, de 27.
Et cela se trouve par ma méthode générale, qui détermine aussi qu'en la condition précédente il en faut 16 au premier, 5,5 au second, et 5,5 au troisième, sans se servir des combinaisons, car elle va partout et sans obstacle. [...] »

Le 25 septembre suivant, Fermat répondit à Pascal et signala l'erreur faite par Pascal en supposant qu'un cas tel que $a c c$ est également favorable aux deux joueurs a et c . En fait, fait remarquer Fermat ce cas n'est favorable qu'au joueur a , car ce joueur gagne un point avant le joueur c , ce qui met un terme à la partie puisque a n'a besoin que d'un point pour gagner.

Fermat écrit :

« [...] Je prends l'exemple des trois joueurs, au premier desquels il manque une partie, et à chacun des deux autres deux, ce qui est le cas que vous m'opposez. Je n'y trouve que 17 combinaisons pour le premier et 5 pour chacun des deux autres ; car, quand vous dites que la combinaison $a c c$ est bonne pour le premier et pour le troisième, il semble que vous ne vous souveniez plus que tout ce qui se fait après que l'un des joueurs a gagné ne sert plus de rien. Or, cette combinaison ayant fait gagner le premier dès la première partie, qu'importe que le troisième en gagne deux ensuite, puisque, quand il en gagnerait trente, tout cela serait superflu ? »

2.2. Biographie de Fermat (1601-1665)

Pierre Fermat naît à Beaumont-de-Lomagne, dans le Tarn-et-Garonne, le 20 août 1601. Son père Dominique est négociant en cuir (corroyeur) et sa mère, Claire de Long, est originaire d'une illustre famille toulousaine. La famille Fermat est l'une des familles bourgeoises et influentes de Beaumont. Le jeune Fermat fait ses études au monastère franciscain de Grandselve. Il quitte ensuite Beaumont pour aller suivre des études de droit civil, d'abord à Orléans, puis à Toulouse, où il assiste au cours de Raymond de Maran. Outre le droit, ses études lui font acquérir une bonne connaissance des langues classiques (latin et grec), comme contemporaines (espagnol et italien). En 1627, Fermat s'inscrit au barreau de Toulouse. Mais lors d'un séjour à Bordeaux, vers 1629, il s'adonne aux mathématiques avec « ces messieurs de Bordeaux ». Il se familiarise avec les travaux d'Archimède, de Pappus et d'Apollonius, ainsi qu'avec les écrits de François Viète. Il fait alors sa première découverte, celle d'une méthode générale pour traiter les problèmes de maxima et minima. De retour à Toulouse, en 1630, il achète la charge de conseiller au Parlement et de commissaire aux requêtes du Palais. Une partie de son travail consiste en l'examen des pétitions adressées au Roi. Cela lui permet de se familiariser avec les affaires, mais le titulaire d'une telle charge ne peut accéder aux autres grades de la hiérarchie judiciaire qu'en acquérant une autre charge. L'année suivante il obtient son diplôme de bachelier agrégé

en droit civil, ce qui lui permet d'exercer à partir du 14 mai 1631. Le 1^{er} juin de la même année il épouse Louise de Long, une cousine de sa mère, avec qui il aura cinq enfants.

Fermat exercera toute sa vie la profession de magistrat. Il ne se consacra aux mathématiques que durant ses moments de loisir. Au début des années 1630, il décrit la construction d'une parabole passant par quatre points donnés et, en 1636, il montre que les homothéties, déplacements, similitudes et inversions transforment un lieu plan en un lieu plan. Cette même année il rédige un mémoire : *Ad locos planos et solidos isagoge*, dans lequel, indépendamment de Descartes, il fonde la géométrie analytique. Il se rend compte de la puissance de cette nouvelle technique, dont l'utilisation aurait rendu beaucoup plus simple l'établissement de ses résultats antérieurs. Il fait toutefois ce constat sans regret et il écrit à ce propos : « Il y a en effet pour la science un certain intérêt à ne pas dérober à la postérité les travaux encore informes de l'esprit ; l'œuvre d'abord simple et grossière se fortifie et grandit par les nouvelles inventions. Il est même important pour l'étude de pouvoir contempler pleinement les progrès cachés de l'esprit et le développement spontané de l'art. »

Ses travaux de mathématiques, aussi élevés soient-ils, ne le détournent pas de sa carrière de magistrat. En 1637, il achète l'office de Raynaldy, qu'il conservera toute sa vie, et devient conseiller ordinaire du Roi à la Chambre d'Enquêtes. Il y reste quatorze ans. À partir de 1638 il siège également à la Chambre de l'Édit à Castres et se rend donc régulièrement dans cette ville. Il monte dans la hiérarchie et devient un notable. Son nom est d'ailleurs, depuis 1631, précédé d'une particule : Pierre de Fermat. Ses fonctions lui assurent des revenus confortables et, par ailleurs, il possède un domaine d'environ cent quarante hectares de bonnes terres de culture. Il accède en 1652 à l'un des postes les plus élevés du Parlement : la Chambre Criminelle. Mais en 1652, la peste s'abat sur la France et Fermat tombe gravement malade. Son ami Bernard Medon annonce même sa mort à ses collègues. Heureusement, il n'en est rien, Fermat recouvre la santé et reprend son travail. En 1653 il est nommé à la « Grand Chambre » du Parlement de Toulouse.

Ses moments libres sont toutefois toujours aussi bien occupés. À partir de 1636 Fermat entre en correspondance avec les savants parisiens, en particulier le père Mersenne et l'ami de ce dernier Jean de Beaugrand, et étend rapidement le cercle de ses correspondants mathématiciens. Fermat s'intéresse, depuis 1636 aux sections coniques et fera de nombreux apports dans ce domaine, qu'il finalise dans *L'Isagoge ad locos ad superficiem*, en 1643. C'est probablement en 1637 qu'il énonce le « grand théorème de Fermat ». En 1640 il donne, sans démonstration, le « petit théorème de Fermat » qui s'énonce ainsi : Si p est un nombre premier et a un nombre entier, alors $a^p - a$ est divisible par p . Pendant cette même période il travaille également sur

le calcul intégral. Il n'est pas de domaine où sa puissance mathématique ne s'exerce pas, ainsi, toujours dans les années 1640, il passe maître dans l'étude des carrés magiques. La correspondance de Fermat avec les mathématiciens s'estompe dans la période allant de 1643 à 1654. Il y a plusieurs raisons à cela : Son métier de magistrat l'accapare, le soulèvement de La Fronde fait rage, à partir de 1648 (et jusqu'en 1652), et Toulouse n'est pas épargné. Enfin la peste sévit en France en 1652.

La correspondance de Fermat reprend à partir de 1654, en particulier avec Pascal au sujet du calcul des probabilités. Les deux hommes prévoient de se rencontrer, mais leurs santés fragiles empêchent de concrétiser ce projet. En 1656, Fermat entre en contact avec Huygens, avec lequel il correspond sur la théorie des nombres.

Après une vie honnête, bien rangée, remplie de sentiments de justice, Fermat meurt le 12 janvier 1665 à Castres dans sa soixante-cinquième année. Deux jours plus tôt il terminait une affaire en cours dans cette ville. Fermat est enterré à Toulouse.

Fermat n'a rendu public qu'un seul mémoire de son vivant. Cependant, à l'âge mûr, il a cherché des correspondants pour mettre au point et éditer ses travaux. Il a pensé à Carcavi, à Blaise Pascal, à Christian Huygens, mais ces tentatives d'édition ont échoué. À sa mort, c'est son fils Samuel qui se charge de ce travail. Bien qu'érudit et lettré, Samuel Fermat n'est pas mathématicien, et ses éditions s'en ressentent. On lui doit, en particulier, la réédition, en 1670, de l'*Arithmetica* de Diophante (il s'agit de l'édition de Bachet de Méziriac) incluant les remarques de son père et les *Varia Opera mathematica* (1679), mais qui ne contiennent pas l'ensemble des découvertes de son père.

2.3. La naissance de l'analyse combinatoire et le *Traité du triangle arithmétique* de Pascal

2.3.1. L'analyse combinatoire

L'analyse combinatoire est née de l'étude des jeux de hasard et s'est fortement développée sous l'influence du calcul des probabilités. Les premiers travaux publiés furent ceux de Pascal dans son ouvrage *Traité du triangle arithmétique, avec quelques petits traités sur la même matière* qui avait été imprimé (à l'exception des quatre premières pages de titre)¹ en 1654 mais n'avait pas été distribué avant 1665. Pascal en adressa cependant une copie à Fermat avant le 29 août 1654. Le triangle arithmétique était connu avant Pascal. On trouve sur le titre d'un ouvrage d'arithmétique d'Apianus daté de 1527,

1. Le premier tirage comporte une erreur typographique page 42 : la majuscule N du début du texte est imprimée à l'envers.

la première représentation imprimée de ce triangle. On sait également que ce triangle était connu du mathématicien et astronome Omar Khayyam dès le XI^e siècle. Il est probable que le triangle arithmétique parvint en Europe en provenance de Chine via l'Arabie.



Apianus (Petrus) Eyn Newe vnd wolgegründte underweysung aller Kauffmans Rechnung in dreyen büchern, mit schönen Regeln und fragstücken begriffen, FIRST EDITION

Ingolstadt, Georgius Apianus, 9 Aug. 1527

Le mathématicien chinois Yang Hui avait écrit deux traités datés de 1261 et de 1275 qui utilisent des fractions décimales. Il se référait au travail d'un autre mathématicien chinois du XI^e siècle Jia Xian qui avait décrit une méthode d'extension du calcul des racines carrées et cubiques à des racines d'ordre supérieur en utilisant ce que nous appelons aujourd'hui le triangle de Pascal.

De nombreux auteurs avaient étudié les nombres figurés ou géométriques, et quelques problèmes d'analyse combinatoire bien avant Pascal, mais celui-ci développa plus complètement le sujet dans son *Traité*. Nous citerons seulement quelques faits :

TRAITE

DV TRIANGLE
ARITHMETIQUE,
AVEC QUELQUES AUTRES

PETITS TRAITEZ SUR LA

MESME MATIERE.

Par Monsieur PASCAL.



A PARIS,

Chez GUILLAUME DESPREZ, rue saint Jacques,
à Saint Prosper.

M. DC. LXV.

- Jordanus de Nemore, un mathématicien allemand montre vers 1225 dans son manuscrit de *Arithmetica*, qu'il sait que :

$$\sum_k \binom{n}{k} = 2^n \text{ et } \sum_k \binom{n}{k} 2^k = 3^n.$$

- Michael Stiffel publie en 1544, un triangle des nombres figurés.
- En 1556, Tartaglia publie une méthode de calcul des coefficients binomiaux et une table du triangle arithmétique dans son ouvrage *General trattato di numeri et misure*¹.
- En 1636, le père Marin Mersenne insère dans son *Harmonicorum Libri XII*², une table du triangle arithmétique où il donne la valeur $\binom{12}{36} = 1\,251\,677\,700$, calculée à la main.

À la fin du XVII^e siècle ce triangle arithmétique fut le point de départ de trois branches des mathématiques : recherches sur les séries infinies, calcul des différences finies et théorie des probabilités.

On a sans doute associé le nom de Pascal au triangle arithmétique parce que ce triangle fut appelé par Montmort en 1708, dans son ouvrage *Essai d'analyse des jeux de hasards*, « la table combinatoire de M. Pascal » et par de Moivre³, en latin, en 1730, *Triangle arithmétique de Pascal*, comme on peut le voir sur la figure ci-dessous tirée de la page 181 du traité de De Moivre :

Triangulum Arithmeticum PASCALIANUM.

	1,	2,	3,	4,	5,	6,	7,	8,	9,
Ordo 1 ^{us}	1,	1,	1,	1,	1,	1,	1,	1,	1,
2 ^{us}	1,	2,	3,	4,	5,	6,	7,	8,	
3 ^{us}	1,	3,	6,	10,	15,	21,	28,		
4 ^{us}	1,	4,	10,	20,	35,	56,			
5 ^{us}	1,	5,	15,	35,	70,				
6 ^{us}	1,		21,	56,					
7 ^{us}			28,						
8 ^{us}			8,						
9 ^{us}									

1. Tartaglia Niccolo, *La prima (-sesta) parte del general tratto di numeri et misure*, Venice, curtio Troiano de i Nauo, 1556-1560.
2. F. Marini Mersenni, *Harmonicorum Libri XII*..., Lutetiae Parisiorum, Sumptibus G. Baudry, 1636.
3. *Miscellanea analytica de seriebus et quadraturis*, Londini, Tonson and Watts, 1730.

Ce triangle est quelquefois appelé, en Italie, *Triangle de Tartaglia*.

Le triangle de Pascal est construit comme l'indique la figure ci dessous, chaque ligne horizontale étant formée à partir de celle qui est au dessus en donnant à chacun des nombres qui la composent une valeur égale à la somme des deux nombres placés dans la ligne précédente juste au-dessus de lui et à la gauche de ce dernier. Par exemple le quatrième nombre dans la cinquième ligne est $10 = 4 + 6$.

	0	1	2	3	4	5	...
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	5	10	10	5	1	
M							

Les coefficients (notés $\binom{n}{i}$ ou C_n^i) de la ligne n sont ceux du développement du binôme $(a + b)^n$.

$$(a + b)^n = \sum_{p=0}^n \binom{n}{p} a^{n-p} b^p.$$

En posant $a = b = 1$ on obtient $2^n = \sum_{p=0}^n \binom{n}{p}$. La somme des coefficients de la ligne n est donc 2^n .

En posant $a = b = -1$ on obtient $0 = \sum_{p=0}^n \binom{n}{p} (-1)^p = \sum_{p=0}^n (-1)^p$.

Permutations

On appelle permutation l'arrangement de n objets parmi n en tenant compte de l'ordre : $P_n = A_n^n = n!$

Combinaisons de p objets parmi n

Si l'on ne tient pas compte de l'ordre des objets dans le rangement on a :

$$C_n^p = \frac{n!}{p!(n-p)!}$$

Dans le triangle de Pascal C_n^p est placé à l'intersection de la ligne n et de la colonne p .

Les coefficients de la colonne 1 valent 1 car $C_n^1 = 1$

Les coefficients de la diagonale 1 valent 1 car $C_n^n = 1$

Les autres coefficients se calculent par :

$$C_{n+1}^{p+1} = C_n^p + C_n^{p+1} \text{ ou } \binom{n+1}{p+1} = \binom{n}{p} + \binom{n}{p+1}.$$

Cette égalité est appelée formule de Pascal. Elle peut se démontrer par récurrence :

$$(a+b)^n = \sum_{p=0}^n \binom{n}{p} a^{n-p} b^p.$$

On peut montrer que si cette formule est vraie pour n , elle est également vraie pour $n+1$.

$$\begin{aligned} (a+b)^{n+1} &= (a+b)(a+b)^n = (a+b) \sum_{p=0}^n \binom{n}{p} a^{n-p} b^p \\ &= \sum_{p=0}^n \binom{n}{p} a^{n+1-p} b^p + \sum_{p=0}^n \binom{n}{p} a^{n-p} b^{p+1} \\ &= \binom{n}{0} a^{n+1} + \sum_{p=1}^n \binom{n}{p} a^{n+1-p} b^p + \sum_{p=0}^{n-1} \binom{n}{p} a^{n-p} b^{p+1} + \binom{n}{n} b^{n+1} \\ &= \binom{n}{0} a^{n+1} + \sum_{p=1}^n \binom{n}{p} a^{n+1-p} b^p + \sum_{p=1}^n \binom{n}{p-1} a^{n+1-p} b^p + \binom{n}{n} b^{n+1} \\ &= \binom{n}{0} a^{n+1} + \sum_{p=1}^n \left(\binom{n}{p} + \binom{n}{p-1} \right) a^{n+1-p} b^p + \binom{n}{n} b^{n+1} \\ &= \binom{n}{0} a^{n+1} + \sum_{p=1}^n \binom{n+1}{p} a^{n+1-p} b^p + \binom{n}{n} b^{n+1} \\ &= \sum_{p=0}^{n+1} \binom{n+1}{p} a^{n+1-p} b^p \end{aligned}$$

La formule étant vraie pour $n+1$, est donc vraie pour tout n .

Pascal a construit son triangle à partir de l'égalité $\binom{n+1}{p+1} = \binom{n}{p} + \binom{n}{p+1}$ et en posant $\binom{n}{0} = 1$. La disposition de son triangle, dans le texte initial, est figurée plus loin dans le présent texte.

2.3.2. Les propriétés du triangle arithmétique et les probabilités

La seconde partie du *Traité du triangle arithmétique* comporte quatre sections :

- usage du triangle arithmétique pour les ordres numériques,
- usage du triangle arithmétique pour les combinaisons,
- usage du triangle arithmétique pour déterminer les partis qu'on doit faire entre deux joueurs qui jouent en plusieurs parties,
- usage du triangle arithmétique pour trouver les puissances des binômes et apotomes.

La première section traite des ordres numériques.

Il s'agit des nombres entiers dits figurés parce qu'on peut les représenter par des points disposés géométriquement. Les nombres figurés sont d'origine très ancienne. On attribue généralement à Pythagore les premières études de nombres figurés et Diophante a résolu plusieurs problèmes les concernant.

Pascal, dans son *Traité* compose la table de nombres dont la première rangée horizontale est composée d'unités ; le second rang est formé par l'addition des nombres successifs du rang supérieur, ce qui donne les nombres naturels ; la troisième est engendrée par l'addition successive des nombres naturels, ce qui donne les nombres dits triangulaires c'est-à-dire figurés par les points d'un triangle ; la même règle produit les nombres pyramidaux, puis les triangulo-triangulaires, etc.. Pascal montre que la table de ces « ordres numériques » est le triangle arithmétique.

Voici ce qu'écrivait Pascal :

« On a considéré dans l'arithmétique les nombres des différentes progressions ; on a aussi considéré ceux des différentes puissances et des différents degrés ; mais on n'a pas, ce me semble, assez examiné ceux dont je parle, quoiqu'ils soient d'un très grand usage : et même ils n'ont pas de nom ; ainsi j'ai été obligé de leur en donner : et parce que ceux de *progression*, de *degré* et de *puissance* sont déjà employés, je me sers de celui d'*ordres*.

J'appelle donc *Nombres du premier ordre* les simples unités 1, 1, 1, 1, ... etc.

J'appelle donc *Nombres du second ordre*, les naturels qui se forment par l'addition des unités 1, 2, 3, 4, 5, etc.

J'appelle donc *Nombres du troisième ordre*, ceux qui se forment par l'addition des naturels, qu'on appelle triangulaires. 1, 3, 6, 10, etc. C'est-à-dire que le second des triangulaires, savoir 3, égale la somme des deux premiers naturels, qui sont 1, 2 ; ainsi le troisième triangulaire 6 égale la somme des trois premiers naturels, 1, 2, 3, etc.

J'appelle donc *Nombres du quatrième ordre*, ceux qui se forment par l'addition des triangulaires, qu'on appelle pyramidaux, 1, 4, 10, 20, etc. [...]

Et ainsi à l'infini.

Or, si on fait une table de tous les ordres des nombres, où l'on marque à côté les exposants des ordres, et au dessus les racines en cette sorte :

	1	2	3	4	5	etc.
Unités.Ordre 1	1	1	1	1	1	etc.
Naturels.Ordre 2	1	2	3	4	5	etc.
Triangul.Ordre 3	1	3	6	10	15	etc.
Pyramid.Ordre 4	1	4	10	20	35	etc.

On trouvera cette table pareille au Triangle arithmétique. »

Ces nombres sont donc, en fait, les coefficients binomiaux. Pour l'ordre que Pascal appelle k , le nombre de la n ième colonne est C_{n+k-2}^{k-1} .

La deuxième section Usage du triangle arithmétique pour les combinaisons, donne la définition suivante du terme combinaison:

« Lorsque de plusieurs choses on donne le choix d'un certain nombre, toutes les manières d'en prendre autant qu'il est permis entre toutes celles qui sont présentées, s'appellent les *différentes combinaisons*. »

La troisième section est relative au problème des partis.

Reprenons ce que nous avons écrit dans le paragraphe consacré à la correspondance entre Pascal et Fermat :

C'est cette correspondance publiée à Toulouse dans les *Varia Opera* de Fermat en 1679, d'une part, et le chapitre intitulé *l'usage du triangle arithmétique pour déterminer les partis qu'on doit faire entre deux joueurs qui jouent en plusieurs parties*, publié par Pascal dans son *Traité du triangle arithmétique*, d'autre part, qui font que l'on attribue à Pascal, en oubliant Fermat, la paternité du calcul des probabilités.

Le problème des partis, abordé dans le chapitre précédent, peut être, également, formulé comme suit : Deux joueurs A et B jouent à « Pile ou Face » et misent chacun une somme M. Le gagnant du jeu sera le premier qui aura remporté c coups ; il lui sera alors attribué les deux mises. On demande, dans l'hypothèse où le jeu est interrompu alors qu'il reste a coups à gagner pour A et b coups pour B, comment les joueurs doivent se partager équitablement les mises.

Ce même problème avait été étudié à la fin du XV^e siècle par Luca Paccioli¹ puis au XVI^e par Nicola Tartaglia². Paccioli a proposé de partager les mises

proportionnellement aux coups déjà gagnés soit : pour A, $2M \frac{c-b}{2c-a-b}$ et pour B, $2M \frac{c-a}{2c-a-b}$.

Tartaglia indiquait, pour sa part : $M \left(1 + \frac{b-a}{c} \right)$ pour A et $M \left(1 - \frac{b-a}{c} \right)$ pour B.

Voici, à titre d'exemple du style de Pascal dans son traité, quelques lignes relatives au problème des partis.

« Méthode pour faire les partis entre deux joueurs qui jouent en plusieurs parties, par le moyen du triangle arithmétique.

Avant que de donner cette méthode, il faut faire ce Lemme :

Lemme : Si deux joueurs jouent à un jeu de pur hasard, à condition que si le premier gagne, il lui appartiendra une portion quelconque sur la somme qu'ils jouent, exprimée par une fraction, et que s'il perd, il lui appartiendra une moindre portion sur la même somme, exprimée par une autre fraction : s'ils veulent se séparer sans jouer, la condition du parti se trouvera en cette sorte. Soient réduites les deux fractions à même dénomination, si elles n'y sont pas ; soit prise une fraction dont le numérateur soit la somme des deux numérateurs, et le dénominateur double des précédents : cette fraction exprime la portion qui appartient au premier sur la somme qui est au jeu. [...]

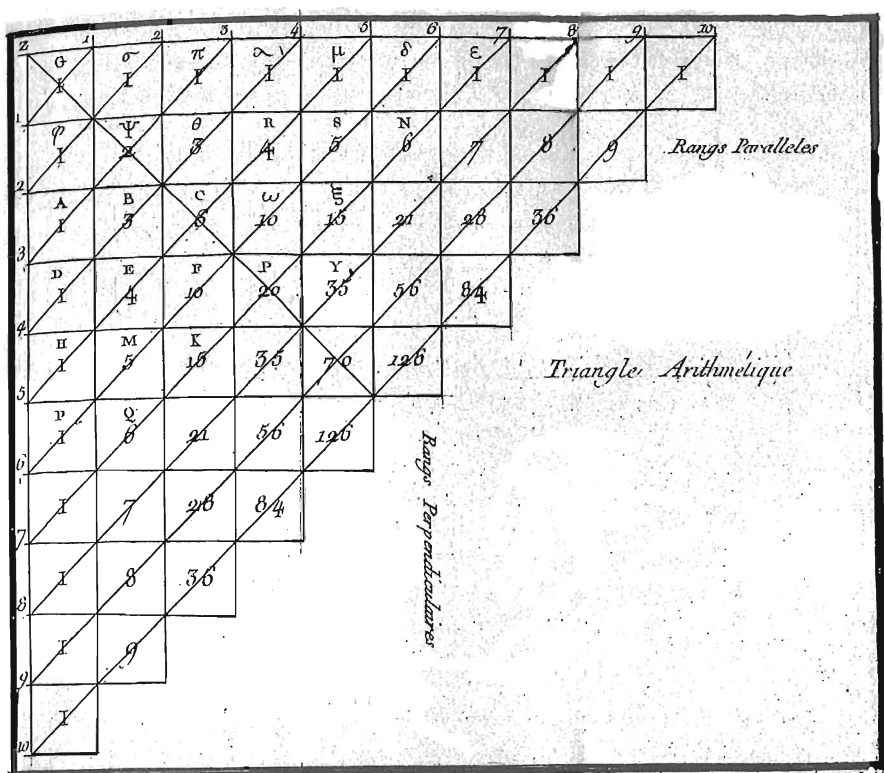
Problème : Étant donnés deux joueurs, à chacun desquels il manque un certain nombre de parties pour achever, trouver par le Triangle Arithmétique le parti qu'il faut faire (s'ils veulent se séparer sans jouer) eu égard aux parties qui manquent à chacun.

Soit prise dans le triangle arithmétique la base dans laquelle il y a autant de cellules qu'il manque de parties aux deux ensembles : ensuite soient prises dans cette base autant de cellules continues à commencer par la première, qu'il manque de parties au premier joueur ; et qu'on prenne la somme de leurs nombres. Donc il reste autant de cellules qu'il manque de parties à l'autre. Qu'on prenne encore la somme de leurs nombres : ces sommes sont l'une à l'autre comme les avantages des joueurs réciproquement : de sorte que si la somme qu'ils jouent est égale à la somme des nombres de toutes les cellules de la base, il en appartiendra à chacun ce qui est contenu en autant de cellules qu'il manque de parties à l'autre ; et s'ils jouent une autre somme, il leur en appartiendra à proportion.

Par exemple, qu'il y ait deux joueurs, au premier desquels il manque deux parties, et à l'autre quatre : il faut trouver le parti.

1. Pacioli Luca, *Somma di aritmetica geometria proporzioni e proporzionalita*, Venice, Paganinus de Paganinis, 1494.
2. *General trattato di numeri et misure di Nicolò Tartaglia*, In Vinegia, per Curtio Troiano dei Navò, 1556.

Soient ajoutés ces deux nombres 2 et 4, et soit leur somme 6 ; soit prise la sixième base du triangle arithmétique $P\delta$, dans laquelle il y a par conséquent six cellules, $P, M, F, \omega, S, \delta$.



Soient prises autant de cellules à commencer par la première P , qu'il manque de parties au premier joueur, c'est-à-dire, les deux premières P et M ; donc il en reste autant que de parties à l'autre, c'est-à-dire quatre, F, ω, S, δ : je dis que l'avantage du premier est à l'avantage du second, comme $F + \omega + S + \delta$ à $P + M$, c'est-à-dire que si la somme qui se joue est égale à $P + M + F + \omega + S + \delta$, il en appartient à celui à qui il manque deux parties la somme des quatre cellules $F + \omega + S + \delta$; et à celui à qui il manque quatre parties, la somme des deux cellules $P + M$; et s'ils jouent une autre somme, il leur en appartient à proportion. [...]

Et s'il manque six parties à l'un, et deux à l'autre, le parti s'en trouvera dans la huitième base dans laquelle les six premières cellules contiennent ce qui appartient à celui à qui il manque deux parties, et les deux autres ce qui appartient à celui à qui il en manque six. Et ainsi à l'infini. »

Pascal introduit dans la solution qu'il développe, ce que nous appelons aujourd'hui l'espérance mathématique, c'est-à-dire la probabilité d'un gain multipliée par la valeur de ce gain. Il écrit :

« Sur quoi la première chose qu'il faut remarquer est que deux joueurs qui jouent en deux parties, dont le premier en a une à point, sont en même condition que deux autres qui jouent en trois parties, dont le premier en a deux, et l'autre une : car il

y a cela de commun que, pour achever, il ne manque qu'une partie au premier, et deux à l'autre : et c'est en cela que consiste la différence des avantages, et qui doit régler les partis ; de sorte qu'il ne faut proprement avoir égard qu'au nombre des parties qui restent à gagner à l'un et à l'autre [...] »

La dernière section *Usage du triangle arithmétique pour trouver les puissances des binômes et apotomes*¹ indique simplement où trouver dans le *triangle arithmétique* les coefficients binomiaux :

$$(a + b)^n = \sum_{p=0}^n \binom{n}{p} a^{n-p} b^p$$

Pascal indique, à la fin de cette section qu'il n'y donne pas de démonstrations parce qu'elles sont évidentes et qu'on peut les trouver dans divers cours dont ceux d'Hérigone.

D'autres travaux sur l'analyse combinatoire furent ensuite publiés par van Schooten en 1657, dans la cinquième partie de son ouvrage déjà cité, puis par Leibniz en 1666.

En 1675, Jean Prestet dans ses *Éléments de mathématiques ou principes généraux de toutes les sciences qui ont les grandeurs pour objet* (Paris, André Pralard, libraire) consacra tout le livre 2 de la seconde partie de son ouvrage à l'analyse combinatoire et cita (déjà) les travaux de van Schooten.

Lorsque John Wallis publia en 1655 son *Arithmetica Infinitorum*, l'importance de l'ouvrage provoqua immédiatement une réaction de Fermat qui adressa un challenge à Wallis et aux autres mathématiciens anglais sous la forme du problème suivant : Trouver un cube tel que la somme $n^3 + n^2 + n + 1$ soit un carré. Wallis publia sa réponse en 1658, dans son *Commercium Epistolicum* puis, en 1685, dans son ouvrage intitulé *Algebra*, il consacra la dernière partie de son chapitre *Discourse of Combinations, Alternations, and Aliquot parts* aux problèmes posés par Fermat.

2.4. Biographie de Pascal (1623-1662)

Blaise Pascal naît le 19 juin 1623 à Clermont. Son père Étienne Pascal, second président de la Cour des aides de Clermont, est passionné de mathématiques (on lui doit par exemple le « limaçon de Pascal »). Sa mère, Antoinette Begon, meurt en 1626. Étienne se retrouve seul avec trois enfants : ses deux filles, Gilberte et Jacqueline, et Blaise le dernier né, qui n'a que trois ans quand sa mère meurt, et est de santé particulièrement fragile. Étienne se charge lui-même de l'instruction de Blaise et de ses filles. Il décide que son fils n'étudiera pas les mathématiques avant l'âge de 15 ans et retire tous les livres de mathématiques de sa maison. Mais Blaise parvient seul, à l'âge de 12 ans,

1. Les apotomes sont des différences de deux quantités dont seules les puissances sont commensurables, par exemple $\sqrt{2} - 1$.

à démontrer la trente-deuxième proposition d'Euclide (la somme des angles d'un triangle est égale à deux droits). Étienne autorise alors son fils à lire « pour ses heures de récréation » *Les Éléments* d'Euclide. Dès l'âge de 14 ans Blaise accompagne son père aux réunions organisées par Mersenne, il peut y rencontrer des savants comme Gassendi, Roberval, Auzout et Desargues. L'année 1639 marque un tournant : alors qu'au mois de juin il présente ses premiers travaux à Mersenne (essentiellement des théorèmes de géométrie projective), en décembre la famille quitte Paris pour Rouen. Étienne Pascal reçoit en effet de Richelieu une mission de commissaire à la levée des impôts en Normandie. Ce changement ne diminue en rien l'ardeur créatrice de Blaise qui en février 1640, à l'âge de 17 ans, publie un *Essai pour les coniques* qui fait l'admiration de ses aînés. Entre 1642 et 1645, il met au point une machine arithmétique, capable d'effectuer les quatre opérations, pour aider son père dans ses comptes professionnels. Cette machine, connue sous le nom de « Pascaline » sera produite à cinquante exemplaires entre 1645 et 1652^{1,2}.

L'année 1646 marque la première conversion de Pascal. Des médecins viennent soigner son père et font lire à Blaise des ouvrages religieux, peut-être même ceux de Jansénius. Cela amène Pascal d'une foi médiocre à une piété ardente. À l'automne de cette même année Blaise Pascal et son père répètent les expériences de Torricelli sur le vide. Ce dernier avait montré que contrairement aux conceptions aristotéliennes suivant lesquelles « la nature à horreur du vide », cette « horreur » disparaissait au-delà d'un certain seuil³. Pascal refait cette expérience avec plusieurs sortes de fluides (les verriers de Rouen réussissent, dit-on, la prouesse de lui confectionner des tubes de verre de 12 m de long, mais cela est peu probable) et montre la relation existant entre hauteur du liquide dans le tube et poids spécifique. Il publie, le 8 octobre 1647, un mémoire de trente-deux pages intitulé *Expériences nouvelles touchant le vide*⁴. Pascal comprend que la hauteur de mercure dans

1. À la base la *Pascaline* fait des additions et des soustractions, mais à partir de ces deux opérations élémentaires, on peut effectuer des multiplications, des divisions et même des extractions de racine carrée. Il existe des machines « en base dix » permettant de faire les « calculs abstraits ». D'autres comportent un nombre inégal de subdivisions et sont destinées à la comptabilité.
2. Pascal est la deuxième personne à inventer une machine à calculer mécanique. La première machine a été inventée et produite par Wilhelm Schickard (1592-1635) en 1624. Toutefois la Pascaline est totalement indépendante de la machine de Schickard, qui du reste aurait été inadaptée à traiter des comptes français du fait de la spécificité de la monnaie française d'alors (il y avait 20 sols dans une livre et 12 deniers dans un sol). Plus tard, en 1671, Leibniz réalisera également une machine à calculer.
3. L'expérience de Torricelli consistait à retourner un tube plein de mercure sur un récipient rempli du même métal ; le mercure s'élève à 27 pouces (~ 73 cm), mais pas au-delà. Le physicien italien fait alors l'hypothèse que l'espace laissé libre au sommet du tube est vide.
4. Peu avant, le 23 septembre 1647, Pascal reçoit la visite de Descartes. La discussion est âpre car ce dernier ne croit pas à l'existence du vide. Descartes, caustique, écrit à Huygens après cette visite (qui n'a duré que deux jours) que « Pascal a trop de vide dans sa tête » !

le tube doit s'abaisser à mesure que l'on prend de l'altitude, car « le poids de l'atmosphère » y devient plus faible (la lettre du 11 juin 1644 de Torricelli à Michelangelo Ricci, avait déjà révélé l'expérience barométrique.) En 1648, le beau-frère de Pascal, Florin Périer, réalise l'expérience jusqu'au sommet du Puy-de-Dôme et confirme la prévision. De son côté, Pascal réalise ensuite l'expérience au sommet de la Tour Saint-Jacques à Paris.

Le 24 septembre Étienne Pascal meurt, ce qui bouleverse la vie de Blaise. Quelque temps après Pascal connaît une période mondaine, durant laquelle il cherche sa voie du côté des sciences et de la vie de société. Il confie à sa sœur Jacqueline qu'au milieu « de ses occupations [...] et parmi toutes les choses qui pouvaient contribuer à lui faire aimer le monde [...], il était de telle sorte sollicité de quitter tout cela, et par une aversion extrême qu'il avait des folies et des amusements du monde, et par le reproche continuel que lui faisait sa conscience, qu'il se trouvait détaché de toutes choses [...], mais que d'ailleurs il était dans un si grand abandonnement du côté de Dieu, qu'il ne sentait aucun attrait de ce côté-là. »

Il rédige entre 1651 et 1653 son *Traité de l'équilibre des liqueurs et de la pesanteur de la masse de l'air* dans lequel il énonce les principes de l'hydrostatique¹. Les compétences de Pascal en hydrostatique sont mises à profit dans l'entreprise d'assèchement des marais poitevins, dont il est sociétaire et conseiller scientifique. Blaise ne délaisse pas pour autant les mathématiques, il reprend en 1653 et 1654 des travaux antérieurs de géométrie projective et finalise son traité sur *La génération des sections coniques*. C'est également en 1654 (Pascal a 31 ans) que son ouvrage *Traité du triangle arithmétique, avec quelques petits traités sur la même matière* est partiellement imprimé.

Pascal est de constitution fragile. Il a une mauvaise santé et souffre depuis sa jeunesse de violentes migraines. Malgré cela son travail scientifique est intense... jusqu'en octobre 1654. À cette époque il est victime d'un accident qui aurait pu lui coûter la vie. Il s'en sort sans blessure, mais cet accident semble l'avoir beaucoup affecté psychologiquement. Un autre événement le marque profondément : « Dieu, dit-il, lui parle le 23 novembre 1654 de dix heures et demie du soir jusqu'à environ minuit et demie ». Cette « nuit de feu » correspond à sa seconde conversion, et son intérêt pour la science fléchit quelque peu². Il se rend alors au monastère de Port Royal des Champs à environ trente kilomètres au sud-ouest de Paris. Il commence à écrire des travaux anonymes sur des sujets religieux, puis, en 1656 et au début de 1657, il publie dix-huit lettres appelées *Les Provinciales*. Ces lettres défendent, en

1. Ce traité ne sera publié qu'en 1663, après sa mort.

2. Il déclarera par exemple dans *Les Pensées* : « Les hommes sont dans une impuissance naturelle de traiter quelque science dans un ordre absolument accompli [...] nous brûlons du désir de trouver une assiette ferme et une dernière base pour y édifier une tour qui s'élève à l'infini, mais tout notre fondement craque et la terre s'ouvre jusqu'aux abîmes. »

fait, la cause de son ami Antoine Arnauld, qui, s'opposant aux Jésuites et soutenant la cause du Jansénisme, est en procès devant la faculté de théologie de Paris. À la fin de l'année 1656 et pendant les années 1657 et 1658, Pascal écrit son œuvre philosophique majeure, *Les Pensées*, qui sera éditée après sa mort par sa sœur. Bien que Pascal soit moins intéressé depuis quatre ans par les questions scientifiques, il réfléchit de nouveau en 1658 à des problèmes de mathématiques pendant les nuits d'insomnies que provoquent le mal qui le ronge. En juin 1658, Pascal, sous le pseudonyme d'Amos Dettonville adresse une lettre circulaire posant six problèmes de mathématiques et offrant deux prix d'un total de 60 pistoles. C'est à Pierre de Carcavy, l'ami de Pascal, que les réponses doivent être soumises avant le 1^{er} octobre de la même année. Le célèbre mathématicien anglais John Wallis envoie une première réponse le 18 août et demande une prolongation du délai. Pascal répond sur un ton fort désagréable par une lettre circulaire du 7 octobre. Il y annonce la clôture du concours, en prétendant qu'il ne peut prolonger le délai car « même ceux qui auraient gagné les prix en se trouvant les premiers entre ceux dont on aurait reçu les solutions au 1^{er} octobre, ne seraient jamais sûrs d'en pouvoir jouir puisqu'ils leur pourraient toujours être contestés par d'autres solutions qui pourraient arriver tous les jours, premières en date, et qui les excluraient sur la foi des bourgmestres et officiers de quelque ville à peine connue du fond de la Moscovie, de la Tartarie, de la Cochinchine et du Japon. » Selon le jugement de Pascal, ni Wallis, ni Lalouvière qui a aussi participé au concours, ne soumettent des réponses convenables et il n'y a pas de prix attribués. Mais après avoir bien étudié les solutions proposées par ses correspondants, qui ignorent son identité, Pascal publie les *Lettres de Détonville*. Wallis rend alors public son *Tractatus duo, prior de Cycloide et corporibus inde genitis, posterior epistolaris in qua agitur de Cissoide et corporibus inde genitis, et de curvarum, tum linearum ...* (Oxoniae, 1659), où il détaille ses propres solutions.

Durant l'été 1662 Pascal est pris de violentes douleurs dans les intestins (dues à une gangrène de l'intestin) et ses migraines deviennent insupportables. Il meurt le 19 août 1662, à trente-neuf ans, en prononçant, dit-on, « Que Dieu ne m'abandonne jamais. »

2.5. Références

- R. Taton, *Blaise Pascal*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 10, p. 330-342, 1970.
- M.S. Mahoney, *Pierre Fermat*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 4, p. 566-576, 1970.
- M.S. Mahoney, *The Mathematical Career of Pierre de Fermat (1601-1665)*, Princeton, 1994.

- W.F. Edwards, *Pascal's arithmetical triangle*, New York, 1987.
- P. Humbert, *L'œuvre scientifique de Pascal*, Paris, 1947.
- A. Koyré, *Études d'histoire de la pensée scientifique*, Paris, 1966.
- Les cahiers de Science & Vie, *Blaise Pascal*, juin 1995.
- *Pascal et la théologie*, Pour la science, n° 16, Les génies de la science, août-nov. 2003.
- S. Hollingdale, *Makers of mathematics*, London, 1989.
- E.T. Bell, *Men of mathematics*, Simon et Schuster, New York, 1986.

Huygens

« Les géomètres du XVII^e siècle qui se sont occupés du calcul des probabilités, ne l'ont employé qu'à déterminer les chances de différents jeux de cette époque ; et ce n'est que dans le siècle suivant qu'il a pris toute son extension, et qu'il est devenu une des principales branches des mathématiques, soit par le nombre et l'utilité de ses applications, soit par le genre d'analyse auquel il a donné naissance¹. »

3.1. La genèse du mémoire *De ratiociniis in ludo aleae*

En 1665, Huygens effectua sa première visite à Paris et y rencontra quelques mathématiciens français. Parmi ces derniers, Gilles Personne de Roberval (1602-1675) et Claude Mylon² (1618-1660) lui indiquèrent les travaux de Pascal et Fermat sur les probabilités et le problème des partis dans les jeux de hasard. Ni les méthodes de Pascal et Fermat, ni les solutions qu'ils avaient développées ne furent révélées à Huygens, mais de retour en Hollande, ce dernier étudia par lui-même les premiers éléments de la théorie des Probabilités. C'est ainsi que fut rédigé le mémoire *De ratiociniis in ludo aleae*.

La correspondance sur le sujet des probabilités entre Pascal et Fermat date de 1654. Les lettres de Pascal furent éditées à Toulouse dans les *Varia Opera* de Fermat en 1679 et reproduites avec, en complément, quelques lettres inédites de Fermat dans les *Œuvres de Blaise Pascal* publiées à La Haye, chez Detune en 1789. Cette correspondance entre Pascal et Fermat n'a donc eu aucune influence sur les travaux de Huygens puisqu'elle ne fut publiée que plus tard.

Le mémoire de Huygens a été inclus dans le livre que Frans van Schooten (1615-1660) publia sous le titre *Exercitationum mathematicorum libri quinque* en 1657 (Leiden, Elsevir) où il occupe les pages 519 à 534, les deux premières comportant une lettre de Huygens à Van Schooten datée du 27 avril 1657.

1. Poisson, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités*, Paris, Bachelier, 1837, p. 1.
2. Le mathématicien Claude Mylon, anima, entre 1654 et 1660 un cercle scientifique qui assura la continuité avec l'académie mathématique du Père Mersenne.



D E
R A T I O C I N I I S
I N
L U D O A L E Æ.

SI lusionum, quas sola fors moderatur, incerti solent esse eventus, attamen in his, quanto quis ad vincendum quàm perdendum propior sit, certam semper habet determinationem. Ut si quis primo jactu unâ tesserâ senarium jacere contendat, incertum quidem an vincet; at quantò verisimilius sit eum perdere quàm vincere, reipsâ definitum est, calculoque subducitur. Ita quoque, si cum aliquo certem hâc ratione, ut ternis lusibus constet victoria, atque ego jam unum lusum vicerim, incertum adhuc uter nostrum prior tertii victor sit evasurus. Verùm quanti exspectatio mea, & contra quanti illius, æstimari debeat, certissimo ratiocinio consequi licet, atque hinc definire, si ludum uti est imperfectum relinquere inter nos convenerit, quantò major portio ejus quod depositum est mihi quàm adversario meo tribuenda esset: vel etiam si quis in locum sortemque meam succedere cupiat, quo pretio me eam ipsi vendere æquum sit. Atque hinc innumeræ quæstiones exoriri possunt inter duos, tres, pluresve collusores. Cumque minimè vulgaris sit hujusmodi supputatio, & sæpe utiliter adhibeatur, breviter hîc quâ ratione aut methodo expedienda sit exponam, ac deinde etiam, quæ ad aleam sive tesseras propriè pertinent, explicabo.

Hoc autem utrobique utar fundamento: nimirum, in aleæ ludo tanti æstimandam esse cujusque sortem seu expectationem ad ali-

V v v

quid

Le papier de Huygens publié par Schooten comporte quatorze propositions relatives aux jeux de dés ou de cartes ou de tirage de boules de diverses couleurs. Le résultat de la troisième proposition est bien connu : Si un joueur a p chances de gagner a et q chances de gagner b , son espérance mathématique est $(pa + qb)/(p + q)$. Ce traité d'Huygens comporte également cinq problèmes posés sans analyse ni solutions et qui seront résolus par Jakob Bernoulli dans son *Ars conjectandi*. Il a également le mérite de dégager, sous le nom d'attente, la notion que nous appelons aujourd'hui espérance mathématique. Il est resté la meilleure analyse du sujet jusqu'aux travaux de Bernoulli, Montmort, et de Moivre.

Le traité est précédé, dans l'édition originale, d'une lettre à van Schooten, où Huygens expose son but :

« [...] On aurait pu dire que je me serais trop occupé d'une chose légère et frivole [...] mais, si on veut examiner à fond ce que nous enseignons, l'on ne tardera pas à reconnaître qu'il ne s'agit point d'un pur amusement, mais des principes d'une belle et très subtile théorie. Je me flatte même que les travaux de ce genre ne paraîtront pas moins importants que ceux de Diophante, quant à la profondeur des recherches ; et que, ne se réduisant pas, comme ceux-ci, à une pure considération des nombres, ils feront encore plus de plaisir. Mais, afin qu'on ne m'attribue pas indûment l'honneur de la première invention, je dois dire que, depuis longtemps, les plus grands géomètres de toute la France se sont occupés de ce genre de calcul ; mais, en résolvant les questions très difficiles qu'ils se proposaient réciproquement, ils cachaient leurs méthodes, de sorte que j'ai été obligé de remonter jusqu'aux premiers éléments de la matière. J'ignore donc encore s'ils partent du même principe que moi ; mais l'expérience m'a souvent fait connaître que nos solutions cadrent parfaitement. On trouvera à la fin de l'ouvrage quelques-uns de ces problèmes. Je les ai donnés sans analyse, non seulement parce que je n'aurais pu être clair sans entrer dans de trop longs détails, mais encore parce que j'ai cru devoir laisser quelque chose pour exercer les lecteurs, s'il s'en trouve. »

3.2. Le mémoire de Huygens

Voici le texte proprement dit de Huygens dans une traduction inspirée de celle donnée par L. Vastel¹, en 1801.

« DU CALCUL DANS LES JEUX DE HASARD

Quoique dans les jeux de pur hasard les événements soient incertains, on peut néanmoins déterminer avec certitude de combien un joueur est plus près de gagner que de perdre. Par exemple, si quelqu'un parie qu'il amènera six points avec un seul dé au premier coup, on ne peut pas dire s'il gagnera ; mais la probabilité qu'il a de gagner ou de perdre n'a rien de douteux, et l'on peut la soumettre au calcul. De même, si je joue avec quelqu'un une partie en trois jeux, et que j'en aie déjà gagné un, on ne

1. *L'art de conjecturer traduit du latin de Jacques Bernoulli par L. G. F. Vastel*, Imp. Le Roy, Caen, 1801.

sait pas encore lequel des deux le premier en aura trois de gagnés ; mais on peut découvrir, par des raisonnements très certains, la valeur de son espérance et celle de la mienne, et déterminer, en conséquence, dans quelle proportion le dépôt ou la mise commune devrait se partager entre nous, si nous convenions de quitter la partie en cet état ; ou pour quel prix je pourrais équitablement vendre mon sort à quelqu'un qui désirerait me remplacer. Il peut s'élever une infinité de questions semblables entre deux, trois, etc. joueurs : mais comme ce genre de calcul n'est nullement vulgaire, et qu'on a souvent l'occasion d'en faire usage, j'en exposerai ici la méthode en peu de mots, et j'expliquerai ensuite ce qui concerne spécialement les dés.

Le principe fondamental de ma méthode est que, dans les jeux de hasard, la valeur du sort ou de l'attente de chaque joueur est précisément ce qu'il faudrait qu'il eût pour parvenir de nouveau, en jouant à jeu égal, à un sort ou à une attente semblable. Par exemple, si quelqu'un cache trois sous dans une main et sept sous dans l'autre, et qu'il m'en donne le choix, je dis que cet avantage est pour moi le même que s'il me donnait cinq sous ; car, avec cinq sous, je puis me retrouver dans le cas d'avoir une attente égale pour trois sous ou pour sept sous, et cela en jouant à jeu égal.

PROPOSITION PREMIÈRE

Si j'attends a ou b , et que l'un puisse m'échoir avec la même facilité que l'autre, mon attente vaut $\frac{a+b}{2}$.

Voici comment on peut, non seulement démontrer, mais encore trouver cette règle.

Soit x la valeur de mon attente. Il faut qu'ayant x , je puisse, en jouant à jeu égal, parvenir de nouveau à un sort semblable. Supposez donc que je joue avec un autre aux conditions suivantes : 1°. que chacun dépose x ; 2°. que celui qui gagnera le dépôt donne a à son adversaire. Ces conditions sont justes et il est clair que je puis, avec la même facilité, avoir a et $2x - a$; savoir, a si je perds le pari, et $2x - a$ si je le gagne : car, dans ce dernier cas, j'ai le dépôt $2x$, dont je dois donner a à mon antagoniste. Si donc $2x - a$ la même valeur que b , j'ai une égale probabilité pour a

et pour b : ainsi je fais $2x - a = b$ et j'ai $x = \frac{a+b}{2}$ pour la valeur de mon attente.

La démonstration en est facile ; car, ayant $\frac{a+b}{2}$, je puis jouer avec un autre qui déposera aussi $\frac{a+b}{2}$, à condition que celui qui gagnera le pari donnera a à celui

qui le perdra. De cette manière, mon sort est égal pour a et pour b car j'ai a si je perds le pari et j'ai b si je le gagne, puisque, dans ce dernier cas, je gagne le dépôt $a + b$, dont je donne a à mon adversaire.

En, nombres. Si je dois avoir 3 ou 7, et avec une égale probabilité, mon attente vaut 5 d'après la proposition précédente ; et il est certain qu'ayant 5 je puis parvenir encore à la même attente : car ayant 5 je puis jouer avec quelqu'un 5 contre 5, à condition que celui qui perdra le pari donnera 3 à l'autre. La loi du jeu sera parfaitement juste, et il est évident que mon sort sera égal pour 3 et pour 7 ; car j'aurai 3 si je perds le pari, et 7 si je le gagne, puisque, dans ce dernier cas, j'aurai 10, dont mon adversaire retirera 3.

PROPOSITION II

Si j'attends a , b ou c , et que chacune de ces choses ou de ces quantités puisse m'échoir avec la même facilité, mon attente vaut $(a + b + c)/3$

Pour trouver cette règle, soit encore x la valeur de mon attente. Il faut donc que, possédant x , je puisse à jeu égal, parvenir à la même attente. Supposons que je joue contre deux adversaires ; chacun de nous trois met x ; je conviens avec le premier qu'il me donnera b si il gagne le jeu et réciproquement, avec le second qu'il me donnera c si il gagne et réciproquement. Il est visible que ce jeu est juste. J'aurai ainsi une chance égale d'avoir b , savoir si le premier joueur gagne, ou c si le deuxième gagne, ou enfin $3x - b - c$ si je gagne moi-même ; car dans ce cas j'obtiens l'enjeu total $3x$, dont je donne b à l'un et c à l'autre. Or, si $3x - b - c$ était égal à a , j'aurais des chances égales d'avoir a , b ou c . Je pose donc $3x - b - c = a$, d'où je tire $x = \frac{a + b + c}{3}$, valeur de ma chance. On trouve, de même, qu'avoir des chances égales

d'obtenir a , b , c , ou d me vaut $\frac{a + b + c + d}{4}$, et ainsi de suite.

PROPOSITION III

Si j'ai p chances d'obtenir a et q d'obtenir b , et si tous les cas peuvent arriver avec une égale facilité, mon attente vaudra : $\frac{pa + qb}{p + q}$.

Pour découvrir cette règle, appelons de nouveau x la valeur de mon attente. Il faut donc qu'ayant x , je puisse parvenir au même sort en jouant un jeu équitable. Pour cela, je prends un nombre de joueurs tel qu'avec moi il y en a $p + q$ en tout, dont chacun met x , de sorte que l'enjeu sera $px + qx$; et il faudra que chacun de nous ait le même espoir de gagner le pari. Je conviendrai avec q joueurs, c'est-à-dire avec chacun d'eux, que si l'un d'eux gagne la partie, il me donnera la somme b , et que si moi je gagne, je lui donnerai la même somme. De même avec les $p - 1$ joueurs qui restent, ou plutôt avec chacun d'eux, je conviendrai que si l'un d'eux gagne la partie, il me donnera la somme a , et que je lui donnerai également la somme a si c'est moi qui gagne. Il est évident que ces conditions sont justes, puisqu'elles ne font tort à aucun joueur. Il est clair que j'ai q chances d'obtenir b , $p - 1$ chances d'obtenir a et une chance, si c'est moi qui gagne, d'avoir $px + qx - bq - ap + a$; en effet, dans ce dernier cas je reçois l'enjeu $px + qx$ dont je suis tenu de donner b à chacun des q joueurs et a à chacun des $p - 1$ autres, ce qui fait en tout $qb + ap - a$. Or, si $px + qx - bq - ap + a$ était égal à a , j'aurais p chances d'avoir a (car j'avais déjà $p - 1$ chances d'obtenir cette somme) et q chances d'avoir b ; je serais donc revenu à mes chances premières. Je fais donc $px + qx - bq - ap + a = a$, et j'ai

$x = \frac{pa + qb}{p + q}$ pour la valeur de mon attente.

PROPOSITION IV

Supposons que je joue avec quelqu'un, à condition que le dépôt doive appartenir à qui aura gagné le premier trois parties, et que j'aie déjà gagné deux parties et lui une. Je demande quelle partie de l'enjeu je dois prétendre au cas où nous voulons interrompre le jeu et partager les mises dans une juste proportion.

Pour arriver à la solution de la question proposée en premier lieu au sujet du partage de l'enjeu entre plusieurs joueurs dont les chances sont inégales, il est nécessaire de commencer par les cas les plus faciles.¹

Premièrement on doit considérer les jeux qui manquent à chacun des joueurs ; car il est certain que si, par exemple, il est convenu entre quelqu'un et moi que le dépôt sera à celui qui le premier aura eu vingt fois l'avantage, que je l'aie eu dix-neuf fois, et mon adversaire dix-huit fois, mon sort l'emportera autant sur le sien que dans l'espèce dont il s'agit, où sur trois jeux j'en ai marqué deux, et mon adversaire un seul ; car, dans l'une et l'autre espèce, il ne me manque qu'un jeu, tandis qu'il en manque deux à mon adversaire.

Pour trouver ensuite combien il nous revient à chacun, il faut examiner ce qui arriverait, si nous achevions la partie. Il est certain que si je gagnais le premier jeu, je complèterais le nombre prescrit, et que j'aurais tout le dépôt que j'appellerai a ; mais si mon adversaire le gagnait, son sort serait égal au mien, puisqu'il manquerait un jeu de part, et d'autre, et chacun aurait a . Or, il est manifeste que j'ai la même facilité pour gagner ou pour perdre le premier jeu, de sorte que mon attente est

égale pour a et pour $\frac{1}{2}a$, ce qui, par la première proposition, me vaut la moitié de a et de $\frac{1}{2}a$, c'est-à-dire, $\frac{3}{4}a$. Il reste à mon adversaire $\frac{1}{4}a$. Ce qu'on aurait pu trouver directement de la même manière. On voit donc que si quelqu'un voulait me remplacer, il faudrait qu'il me donnât $\frac{3}{4}a$ pour me dédommager ; et qu'ainsi

celui qui parierait de marquer un jeu, avant qu'un autre en eût marqué deux, devrait toujours mettre trois contre un.

PROPOSITION V

Supposons qu'il me manque un jeu et qu'il en manque trois à mon adversaire. Il faut faire la distribution.

De nouveau regardons l'état où nous serions si je gagnais la première partie ou si mon adversaire la gagnait. Si je la gagnais, j'aurais l'enjeu a , mais si c'était lui qui la gagnait, il lui en manquerait encore 2 parties contre une qui me manquerait à moi ; nous nous trouverions donc dans l'état décrit dans la proposition précédente,

et il me reviendrait $\frac{3}{4}a$, comme nous l'avons montré. J'ai donc des chances égales d'obtenir a ou $\frac{3}{4}a$, ce qui, d'après la première Proposition, me vaut $\frac{7}{8}a$. Reste $\frac{1}{8}a$

pour l'autre joueur. Ma chance est donc à la sienne comme 7 est à 1.

De même que le présent calcul dépend du calcul précédent, de même le résultat obtenu ici est nécessaire au calcul du cas suivant, où il me manque une partie, tandis qu'il en manque quatre à mon adversaire. On trouve de la même façon qu'il

me revient dans ce cas $\frac{15}{16}$ de l'enjeu, et $\frac{1}{16}$ à lui.

1. Dans l'édition de van Schooten de 1657, les deux premiers paragraphes sont, à la suite d'une erreur, imprimés dans l'ordre inverse.

PROPOSITION VI

Supposons qu'il me manque deux jeux et qu'il en manque trois à mon adversaire. À l'issue de la première partie, ou bien il me manquera encore une partie et trois à l'autre joueur (auquel cas il me reviendra $\frac{7}{8}a$ (par la précédente proposition)), ou bien il manquera deux parties à chacun de nous, auquel cas il me revient $\frac{1}{2}a$, attendu que nos sorts sont devenus égaux. Mais j'ai une chance contre une de gagner la première partie ou de la perdre ; j'ai donc des chances égales d'obtenir $\frac{7}{8}a$ ou $\frac{1}{2}a$, ce qui me vaut $\frac{11}{16}a$ d'après la première Proposition. Onze parties de l'enjeu me sont donc dues contre cinq à mon compagnon de jeu.

PROPOSITION VII

Supposons qu'il me manque encore deux jeux et quatre à mon adversaire. Si je gagne la première partie, j'aurai donc encore une partie à gagner contre quatre ; si je la perds, encore deux contre trois. J'ai donc des chances égales d'avoir $\frac{15}{16}a$ ou $\frac{11}{16}a$, ce qui me vaut autant que $\frac{13}{16}a$ d'après la première Proposition. Il en résulte qu'il est plus avantageux d'avoir deux parties à gagner contre quatre, qu'une contre deux. Car dans ce dernier cas : à savoir le cas d'une partie contre deux, ma part est $\frac{3}{4}a$ d'après la IV^e Proposition, qui est plus petite que $\frac{13}{16}a$.

PROPOSITION VIII

Supposons à présent qu'il y a trois joueurs et qu'il manque un jeu, tant au premier qu'au second mais qu'il en manque deux au troisième. Pour trouver le sort du premier, il faut encore examiner quelle part lui serait due si c'était lui ou l'un des deux autres qui gagnait le jeu suivant. Si c'était lui, il aurait le dépôt que j'appelle a ; si c'était le second, le premier n'aurait rien, car la partie serait finie : mais si c'était le troisième, alors il manquerait encore un jeu à chacun des trois joueurs, et partant il serait dû $\frac{1}{3}a$, tant au premier qu'à chacun des deux autres. Et comme chacun des trois peut, avec une égale facilité, gagner le jeu prochain, le premier a une chance pour a , une pour 0 et une pour $\frac{1}{3}a$, ce qui lui vaut $\frac{4}{9}a$, par la II^e proposition ; le second a également $\frac{4}{9}a$, et il reste au troisième, $\frac{1}{9}a$, part qu'on aurait pu trouver aussi séparément, et d'après laquelle on aurait déterminé celles des deux autres.

PROPOSITION IX

Pour trouver la part de l'un quelconque d'autant de joueurs qu'on voudra, auxquels il manque divers nombres de jeux, il faut examiner combien il serait dû à celui dont on veut trouver la part, si c'était lui, ou l'un des autres consécutivement qui gagnait

le jeu suivant. Si l'on ajoute ensemble ces valeurs et qu'on divise la somme par le nombre des joueurs, le quotient donnera la part cherchée.

Soient trois joueurs A , B et C , et supposons qu'il manque un jeu à A , deux à B et, deux à C : il s'agit de trouver ce que B a le droit de prétendre dans le dépôt que j'appelle q .

Il faut examiner d'abord quel serait le sort de B , si c'était lui, ou A ou C qui gagnait le premier des jeux suivants.

Si c'était A , la partie serait finie, et, par conséquent, il serait dû 0 à B ; si c'était B , il lui manquerait encore un jeu, il en manquerait un à A et deux à C ; et partant, il

serait dû à B , $\frac{4}{9}q$, par la VIII^e proposition : enfin, si c'était C , il manquerait un jeu tant à A qu'à C , et deux à B , et, par conséquent, il serait dû à B $\frac{1}{9}q$, par la même

proposition VIII. Maintenant il faut ajouter ensemble les valeurs qui seraient dues à B dans ces trois hypothèses ; savoir, 0, $\frac{4}{9}q$, $\frac{1}{9}q$ dont la somme est $\frac{5}{9}q$, qui, divisée

par 3, nombre des joueurs, donne $\frac{5}{27}q$ telle est la part cherchée. Il est facile de le démontrer d'après la seconde proposition ; car, puisque B a la même probabilité

pour 0, $\frac{4}{9}q$ et $\frac{1}{9}q$, il a par la II^e proposition $\frac{0 + \frac{4}{9}q + \frac{1}{9}q}{3}$, c'est-à-dire $\frac{5}{27}q$.

Et il est certain que ce diviseur 3 est le nombre des joueurs. Or, pour trouver l'attente de l'un quelconque des joueurs dans chaque hypothèse, c'est-à-dire, en supposant que ce soit lui ou quelqu'un des autres qui gagne le jeu suivant, il faut chercher d'abord la solution des cas les plus simples, et par leur moyen celle des suivants ; car comme on n'aurait pu résoudre le cas qui vient d'être proposé, avant d'avoir soumis au calcul celui de la proposition VIII, où les jeux manquants étaient 1, 1, 2, de même aussi l'on ne peut supputer la part de l'un quelconque des joueurs, dans le cas où les jeux manquants sont 1, 2, 3, qu'on n'ait soumis au calcul celui où les jeux manquants sont 1, 2, 2, comme nous l'avons déjà fait, et de plus celui où les jeux manquants sont 1, 2, 3, qu'on aurait pu calculer également par la proposition VIII. Et c'est ainsi qu'on peut calculer successivement tous les cas compris dans la table suivante, et une infinité d'autres.

TABLE POUR TROIS JOUEURS

Parties qui leur manquent	1. 1. 2	1. 2. 2	1. 1. 3	1. 2. 3
Leurs parts	4. 4. 1	17. 5. 5	13. 13. 1	19. 6. 2
	9	27	27	27

Parties qui leur manquent	1. 1. 4	1. 1. 5	1. 2. 4	1. 2. 5
Leurs parts	40. 40. 1	121. 121. 1	178. 58. 7	542. 179. 8
	81	243	243	729

Parties qui leur manquent	1. 3. 3	1. 3. 4	1. 3. 5	
Leurs parts	65. 8. 8	616. 82. 31	629. 87. 13	
	81	729	729	

Parties qui leur manquent	2. 2. 3	2. 2. 4	2. 2. 5	2. 3. 3	2. 3. 4	2. 3. 5
Leurs parts	34. 34. 13	338. 338. 53	353. 353. 23	133. 55. 55	451. 195. 83	1433. 635. 119
	81	729	729	243	729	2187

Des dés à jouer

En ce qui concerne les dés, on peut poser ces questions : à savoir en combien de fois l'on peut accepter de jeter avec un dé un 6 ou bien un des autres nombres ; de même en combien de fois 2 six avec 2 dés ou 3 six avec 3 dés. Et bien d'autres questions encore. Pour les résoudre, il faut remarquer ce qui suit.

D'abord qu'on peut faire avec un dé six coups différents qui peuvent arriver avec une égale facilité car je suppose que le dé a une figure parfaitement cubique.

Pour deux dés il y a 36 différents coups encore également probables ; car chaque coup de l'un des dés peut se rencontrer avec chacun des six coups de l'autre. Et six fois 6 font 36 coups. Pour trois dés, il y a 216 coups différents car avec chacun des 36 coups des deux peut se rencontrer chacun des 6 coups d'un troisième dé. Et six fois 36 font 216 coups.

Avec trois dés on a	1 3 6 10 15 21 25 27	coups pour	5ou18 4ou17 5ou16 6ou15 7ou14 8ou13 9ou12 10ou11	points

On voit également qu'il y a six fois 216 c'est-à-dire 1296 coups pour quatre dés ; c'est ainsi qu'on peut calculer, de proche en proche, combien il y a de coups pour un nombre quelconque de dés en sextuplant le nombre de coups qu'il y avait pour un dé de moins. Au surplus, il faut remarquer qu'avec deux dés il n'y a qu'un coup

pour 2 ou 12 points, mais qu'il y en a deux pour 3 ou 11 ; car si nous appelons les dés A et B il est clair qu'on peut avoir 1 point en A et 2 en B ou 1 en B et 2 en A, et qu'ainsi il y a deux manières d'avoir 3 points. De même on peut avoir 6 points en A et cinq en B ou 5 en A, ce qui donne deux manières d'avoir 11 points. Il y a trois coups pour 4 points, car on peut avoir 1 point en A et 3 en B ou 3 en A et 1 en B ou 2 en A et 2 en B. Il y a pareillement trois coups pour 10 points ; quatre coups pour 5 ou 9 points, cinq coups pour 6 ou 8 points, six coups pour 7 points.

PROPOSITION X

Trouver en combien de jets on peut parier d'amener six points avec un seul dé.

Qu'un joueur parie d'amener le six au premier jet, on voit qu'il a un cas pour gagner et pour avoir l'argent du jeu, mais qu'il a cinq cas pour perdre et pour ne rien avoir ; car il y a cinq coups de dé contre lui, et il n'y en a qu'un en sa faveur. Soit a le dépôt ou l'argent du jeu. Il a une chance pour a et cinq pour 0, ce qui, par la seconde

proposition, vaut $\frac{1}{6}a$; et il reste pour son adversaire $\frac{5}{6}a$. Ainsi on ne peut parier

que 1 contre 5 d'amener le six au premier jet.

Si un joueur parie d'amener une fois le six en deux jets, son sort se calcule comme suit. S'il amène six au premier jet, il obtient a ; s'il amène un autre point, il lui reste

un seul jet, qui, comme on le vient de voir, vaut $\frac{1}{6}a$. Or, il n'a qu'un cas pour avoir le

six au premier jet, et il en a cinq pour avoir un autre point. Donc, au commencement

du jeu, il a un cas pour a et cinq pour $\frac{1}{6}a$, ce qui, par la seconde proposition, vaut $\frac{11}{36}a$. Ainsi il reste pour son adversaire $\frac{25}{36}a$; de sorte que le sort de l'un, ou la

valeur de son attente, est au sort de l'autre, dans la raison de 11 à 25, plus faible que celle de 1 à 2.

De là l'on trouve, par un calcul semblable, que le sort de celui qui parie d'amener

une fois le six en trois jets, vaut $\frac{91}{216}a$; de sorte qu'il peut parier 91 contre 125,

c'est-à-dire, un peu moins de 3 contre 4.

PROPOSITION XI

Trouver en combien de jets on peut entreprendre d'amener deux six, avec deux dés.

Si quelqu'un entreprend d'amener deux six avec deux dés au premier jet, il est visible qu'il a un cas pour gagner la partie, ou pour avoir a , et qu'il en a 35 pour la perdre, ou pour ne rien avoir, puisqu'il y a 36 coups de dés. Il a donc, par la seconde

proposition $\frac{1}{36}a$.

Celui qui entreprend la même chose en deux jets, obtient a , si le premier jet donne deux six ; s'il ne le donne pas, il lui reste un jet, ce qui, comme on le vient de voir,

lui vaut $\frac{1}{36}a$.

Or, il a un cas pour avoir deux six au premier jet, et 35 pour ne pas l'avoir ; donc, au commencement du jeu, il a un cas pour a et 35 pour $\frac{1}{36}a$, ce qui, par la seconde proposition, vaut $\frac{71}{1296}a$, et il reste, pour son adversaire $\frac{1225}{1296}a$.

On peut, en partant de là, trouver le sort ou la part de celui qui entreprend, la même chose en quatre jets, en omettant le cas du pari en trois jets. En effet, celui qui parie d'amener deux six en quatre jets, obtient a s'il réussit au premier ou au second jet, autrement il lui reste deux jets qui, par ce qui précède, valent $\frac{71}{1296}a$, mais, par

la même raison, il a aussi 71 cas pour avoir deux six au premier ou au second jet, contre 1225 qui ne le donnent pas. Il a donc, au commencement du jeu, 71 cas pour a et 1225 pour $\frac{71}{1296}a$, ce qui vaut, par la deuxième proposition, $\frac{178\ 991}{1\ 679\ 616}a$ et il reste pour son adversaire $\frac{1\ 500\ 635}{1\ 679\ 616}a$. Ainsi leurs sorts sont dans le rapport

de 178 991 à 1 500 925.

D'après cela, on trouve de la même manière l'attente de celui qui parie d'amener deux six en 8 jets, et de là le sort du joueur qui l'entreprend en 16. Le sort de ce dernier et de celui qui entreprend la même chose en 8 jets, servent à déterminer l'attente du joueur qui parie d'amener deux six en 24 jets. Et, comme dans cette opération, on cherche principalement combien il faut de jets pour que le sort du joueur qui parie d'amener deux six devienne égal à celui de son adversaire, on pourra supprimer quelques-uns des derniers chiffres, sans quoi les nombres croîtraient immensément. C'est ainsi que je trouve que celui qui entreprend d'amener cette combinaison en 24 jets, n'est pas encore au pair, et que celui qui l'entreprend en 25 jets commence à avoir l'avantage.

PROPOSITION XII

Trouver avec combien de dés on peut entreprendre d'amener deux six au premier jet.

Cette question est précisément la même que s'il s'agissait de savoir en combien de jets on peut entreprendre d'amener deux fois le six avec un seul dé. Si quelqu'un

l'entreprenait en deux jets, il lui écherrait, comme on l'a vu, $\frac{1}{36}a$. Celui qui se

chargerait de l'amener en trois jets, et qui ne l'aurait pas au premier, aurait encore deux jets, dont il faudrait que chacun donnât le 6, ce que nous avons dit valoir $\frac{1}{36}a$

, mais s'il avait le 6 au premier jet, il suffirait que les deux autres lui donnassent une fois ce nombre de points, ce qui, par la proposition X, vaut $\frac{11}{36}a$: or il est

certain qu'il n'a qu'un cas pour avoir le 6 au premier jet, et qu'il en a cinq pour ne l'avoir pas. Il a donc au commencement du jeu un cas pour $\frac{11}{36}a$, et cinq cas pour

$\frac{1}{36}a$, ce qui, par la proposition II, vaut $\frac{16}{216}a$ ou $\frac{2}{27}a$. Si l'on continue le calcul

graduellement, en supposant toujours un jet de plus, on trouve qu'on peut parier avec avantage d'amener deux fois le 6 en dix jets avec un seul dé, ou d'amener deux 6 avec dix dés au premier jet.

PROPOSITION XIII

Si j'ai un coup à jouer avec deux dés, et que la loi du jeu soit que je gagnerai le pari si j'amène 7 points, que mon adversaire le gagnera si j'en amène 10, et que, dans tout autre cas, nous partagerons également l'argent du jeu, quelle part avons-nous chacun, à y prétendre ?

Puisque sur les 36 cas que donnent deux dés, il y en a 6 pour sept points et 3 pour dix, il en reste 27 où les deux sorts sont égaux, et où il nous appartient à chacun $\frac{1}{2}a$. S'il n'échet aucun de ces 27 cas, j'en aurai 6 pour gagner le pari, c'est-à-dire, pour avoir a , et 3 pour le perdre et ne rien avoir ; ce qui, par la II^e proposition, me vaudrait $\frac{2}{3}a$. J'ai donc, au commencement du jeu, 27 cas pour $\frac{1}{2}a$, et 9 cas pour $\frac{2}{3}a$; ce qui, par la même proposition, équivaut à $\frac{13}{24}a$, et il reste pour mon adversaire $\frac{11}{24}a$.

PROPOSITION XIV

Si jouant avec deux dés, que nous jetons alternativement mon adversaire et moi, je dois gagner dès que j'amènerai 7 points, et lui dès qu'il en amènera 6, et que je lui aie accordé le premier jet, quel sera le rapport de mon sort au sien ?

Soient x mon sort et a le dépôt ; le sort de mon antagoniste sera $a - x$, et il est clair que toutes les fois que les dés lui reviendront, mon sort sera encore égal à x : mais quand c'est à moi à jouer, mon sort a plus de valeur. Soit cette valeur y . Puisque sur 36 cas que donnent deux dés, il y en a 5 qui donnent 6 points et qui font gagner la partie à mon adversaire, et 31 autres où les dés me viennent, j'aurai, avant qu'il joue, 5 cas pour 0 et 31 cas pour y , ce qui, par la III^e proposition, vaut $\frac{31y}{36}$; mais

nous avons supposé d'abord que mon sort était égal à x . On aura donc $\frac{31y}{36} = x$, et partant $y = \frac{36x}{31}$. Nous avons supposé d'ailleurs que quand c'est à moi à jouer, mon sort vaut y . Or, j'ai alors 6 cas pour a , puisqu'il y en a 6 pour 7 points qui me

font gagner la partie ; et j'ai 30 cas pour x , puisqu'il y en a 30 où les dés reviennent à mon adversaire, ce qui, par la III^e proposition vaut $\frac{6a + 30x}{36} = y$. Mais nous avons déjà trouvé $y = \frac{36x}{31}$. On a donc $\frac{36x}{31} = \frac{6a + 30x}{36}$ d'où l'on tire $x = \frac{31a}{61}$, valeur de mon sort, par conséquent, celui de mon adversaire vaut $\frac{30a}{61}$: ainsi, mon sort est au sien comme 31 à 30. »

Après ses quatorze propositions Huygens propose cinq problèmes :

« PROBLÈME I

A et B jouent ensemble avec deux dés, et voici la loi du jeu : A gagnera la partie s'il amène six points, et B s'il en amène sept ; A commence et joue une fois, après quoi B joue deux fois consécutivement, puis A reprend et joue deux fois, et ainsi de suite, jusqu'à ce que l'un ou l'autre ait gagné. On demande quel est le rapport du sort de A au sort de B ? je réponds : comme 10 355 à 12 276.

PROBLÈME II

Trois joueurs A, B et C, prennent 12 jetons, 4 blancs et 8 noirs, et jouent à cette condition : celui qui, les yeux voilés, amènera le premier un jeton blanc gagnera le pari :

A tire le premier, B le second, C le troisième, puis A recommence et ainsi de suite alternativement. On demande en quel rapport seront leurs sorts respectifs.

PROBLÈME III

A parie avec B que sur 40 cartes à jouer, 10 de chaque espèce ou couleur, il en tirera 4, telles qu'il y en ait une de chaque espèce. Et il se trouve que le sort de A est au sort de B, comme 1 000 à 8 139.

PROBLÈME IV

A prend douze jetons, 4 blancs et 8 noirs, comme ci-dessus, et il parie avec B que, les yeux voilés, il en tirera 7, dont trois seront blancs. On demande quel est le rapport du sort de A au sort de B.

PROBLÈME V

A et B prennent chacun 12 écus et jouent avec trois dés aux conditions suivantes : s'il vient 11 points A donnera un écu à B ; mais s'il en vient 14, B donnera un écu à A ; et celui qui le premier aura réuni tous les écus gagnera la partie. Il se trouve que le sort de A est au sort de B 244 140 625 à 282 429 536 481. »

Montmort, Bernoulli et Struyck ont développé des solutions à ces problèmes. Ces solutions sont détaillées dans de nombreux ouvrages.

3.3. La notion d'attente ou d'espérance de gain et celle de probabilité

Le traité de Huygens introduit pour la première fois la notion d'espérance.

On trouve sous la plume de Huygens le terme de *Hazard*, en français, pour désigner un nombre de cas favorables ; on peut en effet lire dans une lettre qu'écrivait Huygens à Carcavi¹ le 6 juillet 1656 :

« Si le nombre de hazards qu'on a pour avoir b soit p , et le nombre des hazards qu'on a pour avoir c est q : cela vaut autant que si l'on avait $\frac{bp + cq}{p + q}$. Par exemple, si j'avais 2 hazards pour avoir $\frac{1}{3}$ de ce qui est mis au jeu et 5 hazards pour en avoir

1. Lettre n° 308, *Œuvres complètes de Ch. Huygens*, tome 1, p. 442, La Haye, Martinus Nijhoff, 1888.

$\frac{1}{2}$, je multiplie $\frac{1}{3}$ par 2 ; et $\frac{1}{2}$ par 5, puis j'ajoute ensemble les produits qui font $\frac{2}{3}$ et $\frac{5}{2}$; la somme est $\frac{19}{6}$, laquelle je divise par 5 + 2 c'est 7, et j'ai $\frac{19}{42}$. Je dis qu'il m'appartient $\frac{19}{42}$ de ce qui est mis au jeu. »

Dans la traduction latine publiée par van Schooten, le mot *expectatio* définit l'attente ou l'espérance de gain.

Il est à noter que la meilleure définition de ce concept se trouve sous la plume de Baruch Spinoza, dans une lettre¹ adressée de Voorburg, à Jean Van der Meer, le 1^{er} octobre 1666 :

« Monsieur,

Dans la campagne solitaire où je vis j'ai réfléchi au problème que vous m'avez posé un jour et j'en ai trouvé la solution simple. La démonstration, qui est générale, repose sur ce principe : le jeu est équitable quand les chances de gain ou de perte, c'est-à-dire l'espérance des deux joueurs, sont égales. Cette égalité doit consister en un rapport entre les chances de gain et l'argent que hasardent les deux adversaires. Si donc les chances sont égales des deux côtés, il faut que les enjeux, les sommes hasardées soient aussi égales. Si les chances sont inégales, l'un devra exposer d'autant plus d'argent que ses chances l'emportent sur celles de l'autre et de la sorte l'espérance sera la même des deux côtés. Si en effet A, jouant avec B, a, supposons, deux chances de gagner et une seulement de perdre, tandis que B n'a qu'une chance de gagner et deux de perdre, on voit clairement que A devra hasarder pour chacune de ses chances de gagner autant que B pour sa chance unique, ainsi l'enjeu de A devra être double de celui de B.

Pour montrer cela plus clairement, supposons que trois joueurs A, B, C jouent entre eux avec des chances égales et que tous trois exposent la même somme. Il est manifeste que, les enjeux étant égaux, chacun ne hasarde que le tiers de l'argent mis au jeu et peut gagner les deux tiers et que, puisqu'il joue contre deux personnes, il n'a qu'une chance de gagner et deux chances de perdre. Si nous supposons que l'un deux, disons C, avant le commencement de la partie veuille se retirer du jeu, il devra recouvrer son enjeu, c'est-à-dire le tiers de la masse totale et B, s'il veut acheter l'espérance de C et prendre sa place, devra mettre au jeu autant que reçoit C. A ne peut s'opposer à cette substitution, car c'est pour lui tout un d'avoir contre lui deux chances appartenant à deux joueurs différents ou de courir le même risque avec un seul adversaire. Cela étant, si, alors qu'un joueur étend la main, l'autre doit choisir entre deux nombres avec cette conviction que, s'il tombe juste, il gagne une certaine somme d'argent et en perd une égale s'il tombe à côté, il est manifeste que l'espérance est la même des deux côtés, pour celui qui met à la devine et pour celui qui doit deviner. De même si l'un des deux joueurs doit choisir entre trois nombres et, en cas qu'il tombe juste, doit gagner une certaine somme, et en perdre la moitié dans le cas contraire. L'espérance sera encore la même pour les deux joueurs, si celui qui étend la main donne à l'autre la faculté de faire deux conjectures successives en convenant que, si l'une d'elles est juste, il gagnera une certaine somme et, si

1. Lettre n° 43, Édition Bruder des *Œuvres de Spinoza*, Leipzig.

aucune ne l'est, il perdra le double. Il y a encore égalité de chance et d'espérance si c'est entre quatre nombres qu'il faut choisir et qu'on ait le droit de faire trois conjectures, devant gagner une certaine somme si l'une d'elles est juste, et perdre le triple si aucune ne l'est. Ou encore s'il y a cinq nombres et que l'on puisse faire quatre tentatives avec chance de gagner un et de perdre quatre. Il suit de là que celui qui étend la main droite, est dans les mêmes conditions que quelqu'un qui, autant de fois qu'il le veut, cherche à deviner un nombre parmi plusieurs et qui à toutes ces tentatives hasarde une fraction de la somme totale mise au jeu dont le numérateur est le nombre de tentatives. Si, par exemple, il y a cinq nombres entre lesquels il faut en désigner un et que l'on ne fasse qu'une tentative, l'enjeu de celui qui cherche à deviner devra être $1/5$ celui des autres $4/5$; s'il y a deux tentatives ces enjeux seront respectivement $2/5$ et $3/5$; s'il y en a trois, $3/5$ et $2/5$; s'il y en a quatre, $4/5$ et $1/5$; s'il y en avait cinq $5/5$ et $0/5$. Et en conséquence il revient au même pour celui qui donne à deviner de risquer $1/6$ du total mis au jeu pour gagner $5/6$, qu'il ait affaire au même joueur faisant cinq tentatives ou à cinq joueurs dont chacun fait une tentative comme c'est le cas dans votre problème. »

3.4. La vie de Huygens (1629-1695)

Christiaan Huygens de Zuylichem naît le 14 avril 1629 à La Haye dans une famille hollandaise influente. Son grand-père Christian a été secrétaire du premier stadhouder des Pays-Bas (le prince Guillaume I d'Orange), puis l'un des principaux conseillers du prince Maurice de Nassau. Quant à son père, Constantijn, il a été le secrétaire du prince Frédéric-Henri, puis du stadhouder Guillaume II d'Orange. La voie de Christian semblait donc toute tracée, mais son don et sa passion pour les sciences font qu'il en sera autrement. Constantijn Huygens, le père de notre physicien consacre ses loisirs à la poésie, à la littérature, à la peinture (il découvre un jeune peintre du nom de... Rembrandt !), à la musique, à la philosophie et aux sciences. René Descartes, qui a vécu une grande partie de sa vie en Hollande, se rendait régulièrement chez les Huygens, et Constantijn entretenait une correspondance assidue avec le Père Mersenne. Le jeune Christian grandit en ce milieu privilégié dans une ambiance intellectuelle intense. La famille Huygens, Constantijn, son épouse Suzanne, née Van Baerle, et leurs cinq enfants, vivent dans un hôtel particulier de La Haye. Constantijn Huygens par ailleurs fait construire une propriété, achevée en 1642, *Hofwijck*¹, située à Voorburg, ville voisine de La Haye. Christian se montre particulièrement doué et habile de ses mains ; ainsi dès l'âge de quatorze ans il construit des instruments. À seize ans il entre à l'Université de Leyde pour y étudier le droit et les mathématiques. Il participe également aux échanges épistolaires avec le Père Mersenne. Il passe ensuite, de 1647 à 1649, deux ans au Collège d'Orange à Breda (fondé par son père) et se montre brillant en mathématiques tout autant que passionné par les techniques. Ses études terminées

1. Cette jolie propriété, entourée d'un magnifique jardin, est aujourd'hui un musée consacré à Constantijn Huygens.

il part faire un voyage, en tant qu'assistant diplomate, au Danemark. Le départ a lieu en octobre 1649. Il espère pouvoir poursuivre jusqu'en Suède pour rencontrer Descartes, mais la rigueur de l'hiver ne le lui permet pas¹. À son retour son père lui accorde une confortable pension qui lui permet de se consacrer uniquement aux mathématiques et aux sciences. Une période d'intense activité débute alors. Il est critique vis-à-vis de la physique cartésienne, ce qui ne l'empêche pas de se montrer en accord avec sa méthode, et travaille dans le petit laboratoire familial où il construit en particulier des instruments d'optique. En 1651, à vingt-deux ans, il écrit un premier essai *Exetasis quadraturae circuli*, dans lequel il montre l'erreur faite par Grégoire de Saint-Vincent dans sa tentative de quadrature du cercle. Il étend ce travail dans une seconde publication datée de 1654 : *De circuli magnitudine inventa*. Il entreprend également, à partir de 1652, l'étude des collisions élastiques ce qui le conduira à sa découverte de la conservation de la quantité de mouvement. C'est également à cette époque que son ouvrage magistral, *De motu corporum ex percussione*, prend forme². Parallèlement, en 1654, Christian Huygens et l'un de ses frères, Constantijn, décident de fabriquer eux-mêmes des lunettes et de polir les lentilles. Au début de 1655, Huygens observe Saturne, avec la lunette de douze pieds qu'il a confectionnée, et découvre le satellite Titan.

En juillet 1655, à vingt-six ans, Christian part avec un autre de ses frères, Lodewijk, pour Paris. Il y rencontre Pierre Gassendi (1592-1655), qui n'avait plus que quelques semaines à vivre, et Roberval (1602-1675). Lors de son séjour, il correspond par ailleurs avec Pascal et Fermat sur le calcul des probabilités. De retour en Hollande, au tout début 1656, il reprend ses observations de Saturne, découvre ses anneaux et étudie les premiers éléments de la théorie des probabilités. Ses travaux dans ce domaine se concrétisent par l'édition en 1657 d'un mémoire sur les jeux de hasard : *De ratiociniis in ludo aleae*. Pendant cette même période (1656-1657) il travaille sur la quadrature de la parabole et met au point l'horloge à pendule. Il démontre que les lames en forme de cycloïde assurent l'isochronisme rigoureux des oscillations et écrit, en janvier 1657, à Van Schooten « J'ai dernièrement trouvé une nouvelle construction d'horloge qui bat avec tant de régularité qu'il y a grande chance que l'on pourrait s'en servir pour déterminer les longitudes si on l'emportait avec soi en mer ». Il présente officiellement son horloge aux autorités le 16 juin 1657. Plus tard (janvier 1675) Huygens introduit également le ressort à spirale, ce qui lui permet de réaliser les premières montres (« Des horloges rendues très justes ensemble et portatives »). L'étude de l'horloge à pendule conique le conduit en 1659 à des travaux sur la force centrifuge. Contrairement à Galilée, qui pensait qu'un corps animé d'un mouvement uniformément circulaire se poursuit sans forces extérieures, Huygens introduit

1. De toute façon la rencontre aurait été improbable, Descartes meurt le 11 février 1650 à Stockholm.
2. Cet ouvrage ne sera toutefois publié qu'à titre posthume, en 1703.

cette notion de force centrifuge et écrit un traité *De Vi Centrifuga*¹. Huygens met également au point sa théorie de la chute des corps grâce à l'emploi systématique du principe de relativité.

En octobre 1660, Huygens part de nouveau pour Paris. Il y rencontre de nombreux savants français : Pascal, Desargues, Auzout, Carcavi, Pierre Petit..., et en particulier Montmor lors des réunions de l'Académie. Le 22 décembre 1660, il est présenté au roi Louis XIV, et lui offre probablement une horloge. Déjà, Huygens se plaît à Paris, il écrit : « Je sais que je ne vivrai jamais plus heureux qu'en cette ville où une communauté d'hommes érudits et une politesse unique au monde m'attachent tous les jours davantage. » Les savants anglais l'invitent ensuite à Londres, où il se rend en avril 1661. Il participe à une réunion de la toute nouvelle *Royal Society* au Gresham College, ce qui lui donne l'occasion de rencontrer Wallis, Wren, Boyle, et Oldenburg. Huygens a la chance d'observer le passage de Mercure devant le Soleil, le 3 mai. Par contre il est déçu par la capitale anglaise, il écrit : « Je ne trouvais pas le séjour à Londres si charmant... la puanteur de la fumée y est insupportable et très malsaine, la ville mal bâtie, des rues étroites et mal pavées, rien que de chétifs bâtiments... fort peu de choses et rien auprès de ce que l'on voit à Paris. »

De juin 1661 à avril 1663, Huygens séjourne de nouveau à La Haye. En mathématiques, il travaille sur les logarithmes hyperboliques, en physique sur la mesure du temps, les tons de la musique et la division tempérée de la gamme, ainsi que sur la machine pneumatique (dont il a vu un exemplaire chez Boyle). Puis il repart pour Paris une troisième fois et y arrive le 3 avril 1663. Il y reste peu de temps. En juin de cette même année, il fait un deuxième séjour à Londres, où il reste jusqu'à la mi-septembre, et y rencontre de nouveau Boyle. C'est lors de ce séjour qu'il est élu membre de la *Royal Society*, le 22 juin 1663. Lorsqu'il revient à La Haye, sa principale préoccupation est de mettre au point ses horloges marines et d'assurer la protection légale de son invention. Huygens aime la France et Paris l'attire. C'est le 21 avril 1666 qu'il arrive de nouveau dans la capitale, cette fois pour un séjour très long. Grand seigneur de la science, il est sollicité par Louis XIV et Colbert pour contribuer au renom de l'Académie des sciences qui va être constituée.

Huygens, dont la santé n'a jamais été très bonne, tombe sérieusement malade en 1670. Se croyant près de mourir, il quitte Paris pour rejoindre la Hollande. Heureusement le mal n'est que passager et il peut revenir dans la capitale française en 1671. Cependant en avril 1672, Louis XIV déclare la guerre à la Hollande et envahit les Provinces Unies. Tous les Hollandais vivant en France reçoivent l'ordre de quitter le pays dans les six mois. Toutefois, Colbert indique à Huygens que cela ne le concerne pas et qu'il peut rester à Paris.

1. *À propos de la force centrifuge*. Le mot centrifuge apparaît ici pour la première fois, il a été forgé par Huygens. Cet ouvrage, écrit en 1659, ne sera publié qu'en 1703.

Au printemps 1672 un étudiant se présente chez Huygens et lui demande de lui donner des leçons de mathématiques. Son nom ? Gottfried Leibniz. Ce dernier, qui connaît déjà bien les mathématiques, souhaite que Huygens le fasse progresser dans ce domaine. Notre physicien, remarquant la brillante intelligence de Leibniz, accepte et prend en main sa véritable formation mathématique, qui va se révéler très efficace. À peu près à la même époque, Huygens a un autre élève qui passera à la postérité, Denis Papin. D'ailleurs notre physicien inspire à ce dernier plusieurs expériences sur le vide, et réalise la première machine à feu.

En 1673, Huygens publie l'un de ses ouvrages fondamentaux, *Horologium oscillatorium sive de motu pendulorum*¹. Trois ans plus tard, Huygens est de nouveau malade et retourne à La Haye. Il y reste deux ans (1676-1678). C'est pendant ce séjour qu'il va travailler sur la double réfraction. À son retour à Paris, en juin 1678, Huygens expose à l'Académie son *Traité de la lumière*. L'ouvrage ne sera toutefois publié que douze ans plus tard. L'auteur y expose sa théorie ondulatoire de la lumière, il explique la réflexion, la réfraction et la double réfraction. La santé d'Huygens, toujours fragile, se dégrade de nouveau en 1679, puis encore en 1681. Il décide alors de revenir en Hollande pour recouvrer la santé. L'année suivante, il se sent rétabli et, vers le mois de novembre, indique qu'il souhaite retourner à Paris. Comme ce n'est pas la bonne saison pour voyager, il obtient de l'Académie un congé de quelques mois lui permettant de poursuivre ses travaux sur ses horloges marines. Finalement accaparé par ses recherches, il passe le premier semestre 1683 dans son pays et demande de pouvoir y rester jusqu'en septembre. La mort de Colbert, survenue le 5 septembre 1683, va modifier le cours de son histoire. François Michel Le Tellier, marquis de Louvois, succède à Colbert au poste de premier ministre, qui se trouve être également le « protecteur » de notre physicien. Huygens lui écrit dès le 16 septembre pour justifier son absence. Six mois s'écoulent sans réponse. Huygens père intervient, sans plus de succès. Début 1685, il semble clair que le retour de Christian Huygens à Paris n'est plus souhaité, il écrit à Paris le 4 janvier : « maintenant que tout est terminé, il me reste seulement à demander la permission de retirer mes meubles et un congé quelque peu honnête. » Cela lui est refusé, car la mort de Colbert a attisé les intrigues à son égard de quelques académiciens jaloux (aux premiers rangs desquels figure La Hire, qui déjà s'était opposé à son entrée à l'Académie). Par ailleurs, les protestants perdaient rapidement leur liberté d'expression et leurs droits jusqu'à la révocation de l'Édit de Nantes, en octobre 1685. En 1681, lorsque Huygens a quitté Paris pour raison de santé, il ne savait pas qu'il ne reviendrait plus jamais en France. Huygens

1. Il adresse un exemplaire de son ouvrage à de nombreux savants, dont, Mariotte, Roberval, Boyle, Hooke, Wallis, Wren et Newton. Ce dernier lit immédiatement le livre et fait savoir, dès le 23 juin 1673, à Oldenburg, par lequel il lui avait transmis, « qu'il était rempli de spéculations astucieuses et utiles, qui sont à l'honneur de l'auteur. »

est amer, mais une autre épreuve l'attend : le 28 mars 1687, son père, dont il a été si proche toute sa vie s'éteint, à quatre-vingt-dix ans passés. À cet âge ce n'était pas inattendu, mais l'événement marqua profondément Christian qui désormais s'habillera en deuil. À la mort de son père, chez qui il vivait à La Haye depuis son retour de France en 1681, il décide d'habiter dans la demeure familiale *Hofwijck*, à Voorburg, endroit qu'il ne quittera plus. Toutefois, lors d'un séjour en Angleterre, Huygens a l'occasion de rencontrer Newton. Le 22 juin 1689 il participe à une réunion de la *Royal Society* au Gresham College. Il y présente un résumé de son *Discours de la cause de la pesanteur* ainsi que de son *Traité d'optique*.

Huygens passe ses dernières années dans la solitude et rédige deux ouvrages, qui seront publiés à titre posthume. Le premier s'intitule *Cosmotheoros*¹, et présente, à un assez large public, les questions soulevées par la cosmologie copernicienne. Le second ouvrage est un volumineux *Traité de dioptrique*² consacré à la théorie des instruments d'optique, domaine sur lequel il travailla toute sa vie.

En mars 1695, Huygens tombe gravement malade. Il écrit son testament qu'il termine par la remarque suivante : « Celui qui a le privilège de mourir sans douleur, doit penser au malheur de ceux à qui l'on refuse l'euthanasie. » Son frère Constantin tient un journal dans lequel il indique l'évolution de la maladie de Christian. Dès le 16 avril on peut y lire : « Christian va mal. Il ne parvient pas à dormir et a peur de devenir fou. Sa chambre est obscurcie. Il ne laisse entrer personne, car ses douleurs augmentent dès qu'il parle. » Ce n'est que trois mois plus tard, après avoir enduré de grandes souffrances, qu'il succombe. Sa maladie intestinale l'emporte le 8 juillet 1695.

Huygens a marqué l'époque intermédiaire entre Galilée et Newton. Contrairement à Descartes, qui pensait qu'il valait mieux faire confiance à notre raison qu'à nos perceptions, Huygens a attaché une grande importance aux confirmations expérimentales. Son œuvre, immense, a été toutefois en partie occultée par les travaux et le renom exceptionnel de Newton.

3.5. Principales publications de Huygens

- « De ratiociniis in ludo aleae », in Van Schooten, *Exercitationum mathematicorum libri quinque*, Leiden, Elsevir, 1657.
- *Systema Saturnium*, Den Haag, Vlacq, 1659.
- *Horologium oscillatorium, sive de motu pendulorum ad horologia aptato demonstrationes geometricae*, Paris, F. Muguet, 1673.

1. Huygens termine ce livre, qui l'occupe pendant un an et demi, en mars 1694, mais il ne sera publié, en latin, qu'en 1698.
 2. *Dioptrica* se trouve dans l'ouvrage *Opera Posthuma*, publié en 1703.

- *Astroscopia compendiaria, tubi optici molimine liberata*, The Hague, Arnold Leers, 1684.
- *Traité de la lumière*, Leiden, Pieter van der Aa, 1690.
- *Cosmotheoros*, The Hague, Adrianus Moetjens, 1698.
- *Opuscula postuma*, Lugduni Batavorum, Boutesteyn, 1703.
- *Opera reliqua*, Amsterdam, Jansson-Waesberg, 1728.
- *Œuvres complètes de Christiaan Huygens publiées par la Société hollandaise des sciences*, 22 volumes, La Haye 1888-1949.

3.6. Références

- H.J.M. Bos, *Christian Huygens*, Dictionary of Scientific Biography, vol. 6, p. 597-613, Charles Scribner's sons, New York, 1970-1980.
- E.T. Bell, *Christian Huygens and the development of science in the seventeenth century*, Edward Arnold & Co, London, 1947.
- *Christian Huygens, Le temps en question*, catalogue réalisé à l'occasion du 350^e anniversaire de sa naissance, Institut Néerlandais, Paris ; Musée Boerhaave, Leyde, 1979.
- C.D. Andriessse, *Christian Huygens*, Paris, Albin Michel, 1998.

Montmort

4.1. L'Essai d'analyse des jeux de hasards

L'ouvrage de Pierre-Remond de Montmort *Essai d'analyse sur les jeux de hasards*¹ est le premier consacré entièrement au calcul des probabilités car il a été publié avant l'*Ars Conjectandi* de Bernoulli, pourtant rédigé avant celui de Montmort. Isaac Todhunter l'analyse dans un chapitre entier de son *History of the theory of probability*². Fontenelle, dans son Éloge de M. de Montmort écrit : « Ces sortes de sujets [les probabilités] n'avaient-ils point été traités ; c'était un vaste pays inculte, où à peine voyait-on cinq ou six pas d'hommes. M. de Montmort s'y engagea avec un courage de Christophe Colomb, et en eut aussi le succès. »

La seconde édition de l'ouvrage a paru en 1713, augmentée d'une importante préface et de la correspondance entre Montmort et Nicolaus (Nicholas I) Bernoulli. C'est cette correspondance échangée entre 1710 et 1712 qui décida, de fait, Nicolaus Bernoulli à publier l'*Ars Conjectandi* de son oncle Jacob.

Le mérite de l'ouvrage de Montmort est d'explicitier les importantes propriétés mathématiques de la théorie probabiliste des jeux de hasard, ouvrant ainsi la voie aux travaux ultérieurs.

Pour le reste, à la page XXXIX de son *Avertissement*, Montmort montre qu'il est moins naïf que Leibniz ou Craig qui croyaient pouvoir démontrer l'existence de Dieu par la théorie des probabilités. Il écrit : « La clarté des Mathématiques et la sainte obscurité de la foi sont des choses trop opposées : je ne crois pas que personne réussisse à en faire jamais l'alliage. »

Montmort indique dans ce même *Avertissement*, les circonstances, qui, selon lui, l'ont amené à entreprendre la rédaction de son livre :

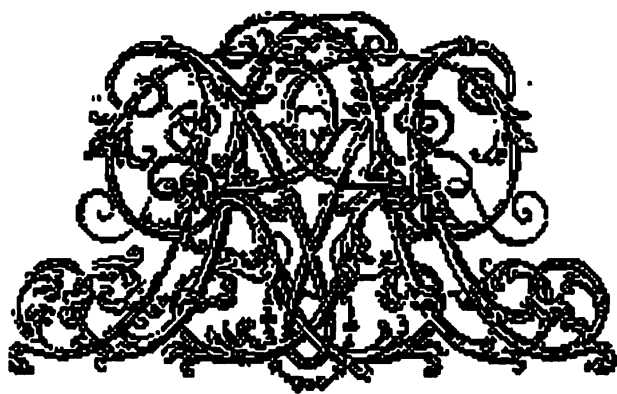
« Plusieurs de mes Amis m'avaient excité, il y a déjà, longtemps à essayer si l'Algèbre ne pourrait point atteindre à déterminer quel est l'avantage du Banquier dans le Jeu du Pharaon. Je n'avais jamais osé entreprendre cette recherche, car je savais que le nombre de tous les divers arrangements possibles de cinquante-deux cartes, surpasse plus de cent mille millions de fois celui des grains de sable que pourrait

1. Imprimé à Paris, chez Quillau en 1708.

2. Cambridge, 1865.

ESSAY
D'ANALYSE
SUR
LES JEUX DE HAZARD.
SECONDE EDITION

Revue & augmentée de plusieurs Lettres.



A PARIS,

Chez JACQUES QUILLAU, Imprimeur-Juré, Libraire
de l'Université, rue Galande.

MDCCXII

AVEC APPROBATION, ET PRIVILEGE DU ROI.

contenir le globe de la terre ; et il ne me paraissait pas possible de démêler, dans un nombre si vaste, les arrangements qui sont avantageux au Banquier, d'avec ceux qui lui sont contraires ou indifférents. Je serais encore dans ce préjugé si les succès de feu M. Bernoulli ne m'eussent invité il y a quelques années à chercher les différents hazards de ce Jeu. Je fus plus heureux que je n'avais osé espérer, car outre la solution générale de ce Problème, j'aperçus les sources qu'il fallait tenir pour en découvrir une infinité de pareils, ou même de beaucoup plus difficiles. »

Son livre est une collection de problèmes d'analyse combinatoire appliquée aux jeux, il est divisé en quatre parties : la première comporte la théorie des combinaisons, la deuxième étudie les jeux de cartes, la troisième les jeux de dés et la dernière divers problèmes de probabilités y compris les cinq proposés par Huygens. Dans la première partie Montmort tente de démontrer le théorème du binôme et discute de la sommation des séries. Dans la seconde édition de son livre Montmort a ajouté sa propre correspondance avec Nicolaus Bernoulli ainsi qu'une lettre de Johann Bernoulli et sa propre réponse.

La première partie, relative aux combinaisons, occupe 72 pages dans la seconde édition. Montmort y explique essentiellement les propriétés du triangle arithmétique de Pascal. Des exemples liés aux jeux de dés et de cartes sont utilisés par Montmort. À la page 32, Montmort, pour tenter de démontrer le théorème du binôme affirme :

« Les coefficients d'un binôme $a + b$, élevé à l'exposant p , sont les mêmes que les bandes perpendiculaires du triangle arithmétique, et ceux-là même qui expriment les diverses combinaisons d'un nombre quelconque de jetons ou de dés qui ont deux faces différentes, l'une appelée a , l'autre appelée b . »

Montmort illustre son propos sur le théorème du binôme en traitant l'exemple de $(a + b)^4$. Avec quatre jetons à deux faces, l'une blanche et l'autre noire, il y a dit Montmort une façon d'obtenir 4 faces noires, trois façons d'avoir 3 noires et une blanche, six façons d'avoir 2 noires et 2 blanches etc. Donc, $(a + b)^4$ doit contenir a^4 et b^4 qui correspond aux résultats de 4 noires ou 4 blanches ; ensuite nous devons avoir le coefficient quatre qui correspond aux occurrences de 3 noires et 1 blanche et aussi de 3 blanches et 1 noire, ensuite on doit avoir le coefficient six qui correspond aux occurrences de 2 faces blanches et 2 noires. Montmort en déduit que les coefficients dans ce binôme doivent être 1, 4, 6, 4, 1.

Il indique ensuite que les coefficients d'un multinôme quelconque de q termes, élevé à la puissance p , sont les mêmes qui expriment les diverses combinaisons d'un nombre quelconque p de dés, qui ont un nombre quelconque de faces désigné par q .

La seconde partie du livre traite des jeux de cartes et, tout d'abord de celui appelé Pharaon.

4.2. La ruine du joueur

Montmort donne, dans la quatrième partie de son *Essai*, la solution de divers problèmes sur le hasard, dont ceux posés par Huygens en 1657. Le cinquième problème de Huygens est le plus intéressant car il est relatif à la ruine de l'un des joueurs. De Moivre et Euler donneront plus tard leurs solutions et Bertrand abordera également le problème en 1888.

Voici l'exposé de l'exercice et la solution de Montmort¹ :

« Proposition XXXVI

Pierre et Paul prennent chacun douze jetons et jouent avec trois dés aux conditions qui suivent. Si les dés amènent onze, Paul donnera un jeton à Pierre. Si les dés amènent quatorze, Pierre donnera un jeton à Paul. Celui des deux qui le premier aura tous les jetons, gagnera. On demande quel est le sort des deux joueurs.

Solution.

Il faut remarquer tout d'abord que trois dés donnent 216 coups, puisque deux dés donnent 36 coups, avec chacun desquels chaque face du troisième dé peut se trouver successivement. On observera ensuite qu'entre ces 216 coups, il y en a 27 qui donnent un jeton à Pierre, savoir six quatre et as qui arrive en six façons, six trois et deux qui arrive en six façons, cinq quatre et deux qui arrive en six façons, cinq cinq et as qui arrive en trois façons, quatre quatre et trois qui arrive en trois façons, trois trois et cinq qui arrive en trois façons, et quinze qui donnent un jeton à Paul, savoir six cinq et trois qui arrive en six façons, cinq cinq et quatre qui arrive en trois façons, six six et deux qui arrive en trois façons, quatre quatre et six qui arrive en trois façons. Cela posé soit A l'argent du jeu, x le sort de Pierre quand il a douze jetons et Paul douze ; y, z, u, t, r, s, q, p, n, m, o, son sort lorsqu'il a gagné à Paul ou un jeton, ou deux jetons, ou trois... ou onze jetons ; k, i, l, h, g, f, e, d, c, b, w, son sort lorsqu'il a perdu contre Paul ou un jeton, ou deux jetons, ou trois... ou onze jetons.

On a $x = (27y + 15k + 174x)/216$, ou ce qui est la même chose $x = (9y + 5k)/14$.

Et de même toutes les autres égalités.

$$\begin{array}{lll} 14y = 9z + 5x & 14p = 9n + 5q & 14h = 9l + 5g \\ 14z = 9u + 5y & 14n = 9m + 5p & 14g = 9h + 5f \\ 14u = 9t + 5z & 14m = 9o + 5n & 14f = 9g + 5e \\ 14t = 9r + 5u & 14o = 9A + 5m & 14e = 9f + 5d \\ 14r = 9s + 5t & 14k = 9x + 5i & 14d = 9e + 5c \\ 14s = 9q + 5r & 14i = 9k + 5l & 14c = 9d + 5b \\ 14q = 9p + 5s & 14l = 9i + 5h & 14b = 9c + 5w \\ & & 14w = 9b + 5x \text{ zéro} \end{array}$$

De toutes ces égalités on tirera

$$y = \frac{31381059609A + 39165289355x}{70546348964}, \text{ et } k = \frac{70497520839}{70548348964}x.$$

Et par conséquent l'on trouvera

$$x = \frac{9y + 5k}{14} = \frac{282429536481A + 352487604195x}{987648885496}$$

1. *Essai d'Analyse...*, 1713, p. 222-223.

en substituant pour y et k leurs valeurs en x , et enfin on aura en réduisant

$$x = \frac{282429536481}{282673677106} A,$$

ce qui exprime le sort de Pierre et

$$A - x = \frac{244140625}{282673677106} A,$$

ce qui représente le sort de Paul.

REMARQUE.

Il est à propos d'observer que la voie analytique n'est peut-être pas ici la meilleure, puisqu'on peut découvrir autrement que les sorts de Pierre et de Paul sont comme les douzièmes puissances des nombres 9 et 5, ainsi qu'il a été observé par Messieurs Bernoulli qui m'en ont averti dans leurs lettres du 17 mars 1710, et 26 février 1711, et depuis par M. Moivre dans son *Traité de Mensura sortis*, qui parut l'année passée. »

4.3. Le jeu de rencontre et les équations de récurrence

Ce problème est étudié pour la première fois par Montmort en 1708 sous le nom de *jeu du Treize* mais il ne donne aucune démonstration de ses résultats dans la première édition de son ouvrage. Dans la seconde édition, il donne, en 1713, deux démonstrations qui lui ont été transmises par Nicolas Bernoulli¹.

L'énoncé du jeu par Montmort peut être résumé comme suit :

Soient A la mise et treize cartes nommées 1, 2, 3, ..., 13. Les treize cartes ayant été battues Pierre les tire une à une successivement. On cherche la chance qu'au moins le numéro d'une carte coïncide avec l'ordre dans lequel elle a été tirée².

La solution met en évidence une équation de récurrence. Voici ce qu'écrit Montmort³ :

« Si l'on nomme S le sort que l'on cherche, le nombre de cartes que Pierre tient étant exprimé par p ; g le sort de Pierre, le nombre de cartes étant $p-1$; d son sort, le nombre de cartes qu'il tient étant $p-2$, on aura $S = \frac{g(p-1)d}{p}$. Cette formule donnera tous les cas, ainsi qu'on les voit résolus dans la table si jointe :

Si $p=1$, on aura $S=A$.

Si $p=2$, on aura $S=A/2$.

Si $p=3$, on aura $S = \frac{2}{3}A = \frac{1}{2}A + \frac{1}{6}A$.

1. *Essai d'Analyse sur les jeux de hazards*, 1713, p. 301-302.

2. Nous donnons, au chapitre 11, l'énoncé du même problème par Euler et le résultat de Lambert.

3. *Essai d'Analyse sur les jeux de hazards*, 1713, p. 134-138.

Si $p = 4$, on aura $S = \frac{5}{8}A = \frac{1}{2}A + \frac{1}{8}A$.

...

Si $p = 13$, on aura $S = \frac{109339663}{172972800}A = \frac{1}{2}A + \frac{22853263}{172972800}A$.

La solution précédente fournit un usage singulier des nombres figurés, car je trouve en examinant la formule, que le sort de Pierre est exprimée par une suite infinie de termes [...]. Il reste, pour le sort de Pierre cette suite très simple

$$\left[\frac{1}{1} - \frac{1}{2} + \frac{1}{1.2.3} - \frac{1}{1.2.3.4} + \frac{1}{1.2.3.4.5} - \dots \text{etc.} \right] A \text{ »}$$

Nous avons déjà défini les nombres figurés que l'on trouve dans le *Triangle arithmétique* de Pascal dans notre chapitre 2. On peut exprimer la formule de récurrence donnée plus haut par Montmort comme suit, en notant S_n « le sort » de Pierre lorsqu'il tient n cartes :

$$nS_n = (n-1)S_{n-1} + S_{n-2} \text{ et } S_n = \left(1 - \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{(-1)^{n+1}}{n!} \right) A$$

4.4. La controverse avec de Moivre

Matthieu Maty, le premier titulaire du poste de bibliothécaire du British Museum, a relaté, dans le *Journal britannique pour les mois de septembre et d'octobre 1755*¹ la controverse entre Montmort et de Moivre en ces termes :

« Peu s'en fallut que l'*Essai d'Analyse de Mr. de Montmort sur les jeux de hazard* qui parut en 1710², ne fit naître une controverse. Mr. Robarts, Mathématicien non moins distingué par son savoir que par sa naissance, ayant lu ce livre proposa à son ami Mr. de Moivre quelques problèmes plus difficiles et plus généraux qu'aucun de ceux qui s'y trouvent. La doctrine des combinaisons et des suites, à laquelle ce dernier s'était depuis longtemps appliqué, lui en fournit les moyens. Le succès lui fit multiplier ses recherches, et lorsqu'il vint à composer la route qu'il avoit suivie avec celle de Mr. de Montmort, il fut surpris de leur diversité, et ne craignit point qu'on lui reprochât d'avoir rien emprunté de lui. La Société Royale en fit le même jugement, et par son ordre le recueil de ses propositions *De mensura sortis* fut imprimé dans les Transactions Philosophiques, et en remplit un cahier. Mr. de Montmort, quoique loué par Mr. de Moivre, ne vit en lui qu'un émule. Il se plaignit à quelques amis, et dans la seconde édition de son livre tâcha d'enlever aux problèmes résolus dans celui de la *mesure du sort* le mérite de l'originalité. Mr. de Moivre ne voulut d'autre juge que Mr. de Montmort lui-même. Ceci fit naître entr'eux un commerce de lettres ; la familiarité et la confiance parurent en être les suites : nos deux Savants se communiquèrent réciproquement leurs inventions sur une matière qu'ils traitaient différemment. Mr. de Montmort passa à Londres en 1715, écrivait-il à Mr. de Moivre, *plutôt pour en voir les Savants que la fameuse éclipse*. Il trouva dans ce dernier un compatriote empressé à lui rendre tous les services de l'amitié, et de retour dans sa patrie il lui en témoigna sa reconnaissance. La seconde édition du livre de Mr. de Moivre plus

1. *Mémoire sur la Vie & sur les Écrits de Mr. De Moivre*.

2. En réalité, 1708.

différente encore que la première d'aucune de celles de Mr. de Montmort lui fut envoyée en 1718, et celui-ci mourut en 1719, sans avoir renouvelé ses premiers reproches. Mais quelques mots de l'éloge de Mr. de Fontenelle semblèrent indiquer dans l'Académicien Français un ressentiment d'autant plus vif qu'il avait été plus longtemps étouffé. Quelque obligeante que fût la manière dont l'illustre Secrétaire de l'Académie des Sciences s'exprimait à l'occasion de Mr. de Moivre, et quoique ce dernier ne manquât pas de lui en faire témoigner sa reconnaissance par leur ami commun Mr. de Varignon, il crut devoir se justifier publiquement de l'odieux soupçon de plagiat dans ses *Mélanges Analytiques*, d'où j'ai tiré ce que je viens de dire¹. »

4.5. La correspondance Montmort-Nicolas Bernoulli

La correspondance entre Nicolas I Bernoulli, l'éditeur de l'*Ars Conjectandi* de son oncle, et Montmort occupe les pages 283-414 de la deuxième édition du livre de Montmort. Dans une lettre datée du 9 septembre 1713 (pages 401-402), Bernoulli propose à Montmort cinq problèmes. Le quatrième est ainsi rédigé :

« A promet de donner un écu à B, si avec un dé ordinaire il amène au premier coup six points, deux écus s'il amène le six au second, trois écus s'il amène ce point au troisième coup, quatre écus s'il l'amène au quatrième et ainsi de suite ; on demande quelle est l'espérance de B. »

Le cinquième problème est celui-ci :

« On demande la même chose si A promet à B de lui donner des écus en cette progression 1, 2, 4, 8, 16, etc. ou 1, 3, 9, 27, etc. [...] au lieu de 1, 2, 3, 4, 5, etc. comme auparavant. »

Ce problème est devenu célèbre sous le nom de paradoxe de Saint-Pétersbourg, car il fut étudié par Daniel Bernoulli, cousin de Nicolas, mais aussi par Buffon, d'Alembert et beaucoup d'autres. Nous l'abordons dans divers chapitres du présent livre.

Par ailleurs, Montmort a souligné la différence entre le problème de la ruine du joueur (voir ci-dessus) et celui de la durée d'un jeu. Dans le problème de la ruine d'un des joueurs, la partie peut durer indéfiniment puisque le gagnant d'un jeu gagne un point et le perdant en perd un. Par contre dans

1. Dans son *Éloge de M. De Montmort (Histoire de l'Académie royale des sciences pour l'année 1719, Paris, 1721, p. 81-93)*, Fontenelle a écrit : « Après le livre de M. de Montmort il en parut un en Angleterre sur la même matière, intitulé *De mensura Sortis*, il est de M. Moivre, fameux géomètre, que la France a droit, puisqu'il est Français, de revendiquer sur l'Angleterre, d'ailleurs fort riche. Je ne dissimulerai point que M. de Montmort fût vivement piqué de cet ouvrage, qui lui parut avoir été entièrement sur le sien, et d'après le sien. Il est vrai qu'il était loué, et n'était-ce pas assez, dira-t-on ? Mais un Seigneur de fief n'en quittera pas pour des louanges celui qui prétend lui devoir foi et hommage des terres qu'il tient de lui. Je parle selon sa prétention, et ne décide nullement s'il était en effet le Seigneur ».

le problème des points la partie durera au plus $a + b - 1$ jeux, $(a + b)$ étant la somme des jetons des deux joueurs, car le gagnant d'un jeu gagne un point et le perdant ne gagne rien. Dans une lettre datée du 15 novembre 1710 adressée à Nicolas Bernoulli, Montmort indique qu'il a résolu le problème de la durée d'un jeu pour des joueurs de forces égales jouant avec le même nombre de jetons¹. Il donne, sans démonstration, un résultat numérique pour la probabilité d'une durée maximale de 26 et 28 jeux pour deux joueurs ayant chacun 6 jetons².

4.6. Biographie de Montmort (1678-1719)

Pierre de Rémond, futur marquis de Montmort naît à Paris le 27 octobre 1678 de François Rémond, écuyer, Sieur de Brévande et de Marguerite Ralle. Après le collège, son père lui fait étudier le droit, car il le destine à une carrière de magistrat. Mais le père était aussi sévère que le fils était rebelle à toute contrainte. Las du droit et de la maison paternelle Pierre décide de partir visiter l'Angleterre, et passe par la Hollande ainsi que par l'Allemagne. Au cours de son voyage il découvre et se passionne pour la *Recherche de la vérité* de Nicolas Malebranche³. Pierre revient en France en 1699 alors qu'il vient tout juste d'avoir vingt et un ans. Son père meurt peu après et Pierre reçoit un héritage considérable qu'il saura gérer sans frivolité. Il se plonge dès lors dans la philosophie et les mathématiques. Il a alors peu de contacts, si ce n'est avec Malebranche qu'il considère à la fois comme son maître, son guide et son ami. Malebranche lui enseigne la philosophie et la physique de Descartes. En 1700, notre futur mathématicien repart en Angleterre, et y rencontre plusieurs savants, dont Newton. De retour en France, il poursuit sa découverte des mathématiques durant trois années. Mais son frère cadet, à qui l'on avait confié un canonicat à Notre Dame de Paris (poste de chanoine, peu exigeant, dit-on), renonce à l'état ecclésiastique et Pierre est sollicité pour le remplacer. Il accepte plus par devoir que par goût : les heures qu'il passe au chœur sont autant de moins de consacrées aux mathématiques. Vers la fin de 1704, Pierre Rémond achète le château de Montmort, situé en Champagne, à 18 km d'Épernay. Ce domaine remonte au Moyen Âge, sans doute à Germond de Montmort en 1042. Le château Renaissance, entouré de son parc et de son jardin potager qu'il domine, est un très beau spécimen de l'architecture du XVI^e siècle.

1. *Essai d'Analyse...*, 1713, p. 303-307.

2. A. Hald a déterminé l'équation qui traduit les résultats de Montmort pour la durée d'un jeu dans son ouvrage *History of probability... before 1750*, Wiley, 1990, p. 350-352.

3. Nicolas Malebranche (1638-1715) est philosophe et théologien. Voulant vivre dans la retraite, il s'enferma dès 1660 dans la congrégation de l'Oratoire. Il a été l'un des plus illustres disciples de Descartes. Par ailleurs mathématicien et physicien, il devient, en 1699, membre de l'Académie des sciences.

À partir de cette époque c'est dans ce château que de Montmort passera la plus grande partie de son existence. Non loin de là, au château de Mareuil, vivait la duchesse d'Angoulême, qui était la veuve du fils illégitime de Charles IX. En lui rendant visite, Pierre de Montmort rencontra mademoiselle de Romicourt, petite nièce et filleule de la duchesse à qui il proposa le mariage en 1706, après avoir renoncé au canonat. Les nouveaux mariés vont mener une vie simple et retirée et Montmort va élaborer son principal ouvrage : *Essai d'analyse sur les jeux de hasard*, qui paraît en 1708. Cet ouvrage est contemporain des recherches de Nicolas Bernoulli (1687-1759), ce neveu de Jacques Bernoulli qui prendra la peine de publier *l'Ars Conjectandi* de son oncle. Nicolas Bernoulli soutient en 1709 sa thèse de doctorat en droit sur le thème de l'art de la conjecture en droit. Il y traite de l'optimisation des rentes viagères, du calcul des usufruits selon les différents âges, etc. La proximité des recherches de Nicolas Bernoulli et de Pierre de Montmort fait naître entre les deux hommes une solide amitié et une saine émulation. Lorsque Nicolas Bernoulli se rend en France, Montmort l'invite dans son château, et Bernoulli y restera trois mois.

Peu de temps après, en 1711, de Moivre présente à la *Royal Society* une étude sur la doctrine du hasard dont la version latine, *De Mensura Sortis*, sera publiée la même année dans les *Transactions*. Montmort est d'abord piqué par cette étude qui lui paraît avoir été entièrement inspirée par son *Essai d'analyse sur les jeux de hasard* paru trois ans plus tôt. Une querelle de priorité, par personnes interposées, commence alors à s'engager, mais heureusement les deux hommes prennent contact directement. À l'occasion d'un voyage à Londres¹, en 1715, de Montmort rencontre de Moivre, tous deux s'expliquent leurs travaux respectifs et finalement tout doute de plagiat disparaît. C'est au cours de ce voyage à Londres, que de Montmort est élu, le 9 juin 1715, membre de la *Royal Society*. Entre-temps, il a fait paraître en 1714 une nouvelle édition de son *Essai d'analyse sur les Jeux de Hasard*. L'ouvrage est considérablement augmenté et comporte la correspondance avec Nicolas Bernoulli. L'année 1716 commence très bien pour lui puisqu'il est admis, en tant qu'associé, à l'Académie des sciences de Paris le 2 février². Mais en septembre 1719, de Montmort quitte son domaine pour un voyage d'affaire à Paris. Il contracte la petite vérole, qui faisait alors beaucoup de ravages, et en meurt le 7 octobre.

-
1. Ce troisième voyage de Montmort à Londres était motivé par l'observation d'une éclipse solaire, qui devait être totale dans la capitale anglaise.
 2. De Montmort ne peut être élu membre à part entière, car il ne réside pas à Paris.

4.7. Publications de Montmort et références

- *Essai d'analyse sur les Jeux de Hasard*, Paris, Quillau, 1708 (2^e éd. 1713-1714).
- *De seriebus infinitis tractatus*, Phil. Trans. Royal Soc., 30, 1720, p. 633-675.

Références

- Fontenelle, *Éloge de M. de Montmort*, Histoire de l'Académie des sciences pour 1719 (1721), p. 83-93.
- I. Hacking, *Montmort*, Dictionary of Scientific Biography, vol. 9, p. 499-500, Charles Scribner's sons, New York, 1970-1980.
- R. Huron, *Un probabiliste disciple de Malebranche : Pierre Rémond de Montmort (1678-1719)*, Annales de la faculté des sciences de Toulouse Sér. 5, S2 (1980), p. 1-31.

Les Bernoulli

5.1. Jacob Bernoulli

5.1.1. La publication posthume de l'*Ars Conjectandi*

L'ouvrage le plus connu de Jacob Bernoulli est son *Ars Conjectandi*, qui pose les fondements de la théorie statistique du calcul des probabilités. Ce livre est laissé inachevé à la mort de Bernoulli, le 16 août 1705. Ce traité est le premier ouvrage conçu, sinon publié, sur la théorie des probabilités, puisqu'il ne paraît qu'en 1713 alors que l'ouvrage de Pierre-Remond de Montmort *Essai d'analyse des jeux de hasards* a été rendu public en 1708.

Bien que Leibniz écrive que c'est à sa demande que Jacob Bernoulli a composé son *Ars Conjectandi* : « Feu M. Bernoulli a cultivé cette matière sur mes exhortations »¹. Bernoulli a, en fait, déjà écrit ce que comporte son traité lorsqu'une correspondance sur les probabilités s'établit entre les deux mathématiciens, comme on peut le vérifier en consultant l'ouvrage du bibliographe de Leibniz, C.I. Gerhardt².

L'*Ars Conjectandi* est donc publié en 1713, huit ans après la mort de son auteur. Une courte préface est ajoutée par Nicolaus Bernoulli, un fils d'un frère de Jacob et de Jean. Cette préface indique que la quatrième partie de l'ouvrage a été laissée inachevée par l'auteur. Le livre contient en outre une annexe, rédigée en Français, intitulée *Lettre à un amy, sur les parties du jeu de paume*, et un traité sur les séries constitué de cinq mémoires écrits par Bernoulli pour servir de sujets de thèse à certains de ses étudiants, dont Nicolaus (également appelé Nicholas I), son neveu. L'éditeur a utilisé d'ailleurs la figure de l'un des mémoires, celui relatif à la thèse de Jacob Hermann (1678-1753) publié en 1696, pour illustrer (face à la page 306) le traité sur les séries dans l'édition originale de l'*Ars Conjectandi*. Cette particularité s'explique d'ailleurs simplement parce que c'est ce même Jacob Hermann qui mit en ordre le manuscrit de Bernoulli, suite à l'intervention de Leibniz qui conseilla ce choix à la veuve de l'auteur. Hermann prit également le soin de faire insérer

1. *Leibnitii Opera Omnia*, studio Ludovici Dutens, Genevae, de Tournes, 1768, vol. VI, 1^{re} partie, page 217.

2. *Leibnizens Mathematische Schriften*, vol. III, 1^{re} partie, Halle, 1855.

JACOBI BERNOULLI,
Profess. Basil. & utriusque Societ. Reg. Scientiar.
Gall. & Pruss. Sodal.
MATHEMATICI CELEBERRIMI,
ARS CONJECTANDI,
OPUS POSTHUMUM.

Accedit

TRACTATUS
DE SERIEBUS INFINITIS,

Et EPISTOLA Gallicè scripta
DE LUDO PILÆ
RETICULARIS.



BASILEÆ,
Impensis **THURNISIORUM, Fratrum.**

cl^o lccc. xiii.

un résumé de l'*Ars Conjectandi* dans les divers éloges consacrés à Bernoulli après sa mort en 1705. L'éloge rédigé, en 1705, et publié dès 1706 par Joseph Saurin¹ (1659-1737) contient l'annonce par Hermann de la loi des grands nombres de Bernoulli. Nous en donnons ci-dessous un extrait :

« [Jacob Bernoulli] achevoit un grand Ouvrage, *De Arte Conjectandi*, et quoiqu'il n'en ait rien paru, nous pouvons en donner une idée sur la foi de M. Herman. Les Règles d'un jeu étant supposées, et deux Joueurs de la même force, on peut, en quelque état que soit une partie, déterminer par l'avantage qu'un des Joueurs a sur l'autre, combien il y a plus à parier qu'il gagnera. Le pari change selon tous les différents états où sera la partie, et quand on veut considérer tous ces changements, on trouve quelquefois des Séries ou suites de Nombres réglées, et même nouvelles et singulières. Si l'on suppose les Joueurs inégaux, on demande quel avantage le plus fort doit accorder à l'autre, ou réciproquement l'un ayant accordé à l'autre un certain avantage, on demande de combien il est plus fort et il est à remarquer que souvent les avantages ou les forces sont incommensurables, de sorte que les deux Joueurs ne peuvent jamais être parfaitement égalés. Les raisonnements que ces sortes de matières demandent sont ordinairement plus déliés, plus fins, composés d'un plus grand nombre de vues qui peuvent échapper, et par conséquent plus sujets à erreurs que les autres raisonnements mathématiques. Par exemple deux Joueurs égaux jouant en 4 parties liées, si l'un en a gagné 3 et l'autre 2, il faut raisonner assez juste pour déterminer précisément que l'on peut parier 3 pour celui qui a les 3 parties et 1 seulement pour celui qui en a 2. Ce cas est des plus simples et on peut juger par là de ceux qui sont infiniment plus compliqués. Quelques grands Mathématiciens et principalement Mrs Paschal & Huguens, ont déjà proposé ou résolu des Problèmes sur cette matière, mais ils n'ont fait que l'effleurer, et M. Bernoulli l'embrassoit dans une plus grande étendue, et l'approfondissoit beaucoup davantage. Il la portoit même jusqu'aux choses Morales et Politiques, et c'est là ce que l'Ouvrage doit avoir de plus neuf et de plus surprenant. Cependant si l'on considère de près les choses de la vie sur lesquelles on a tous les jours à délibérer, on verra que la délibération devoit se réduire, comme les Paris que l'on feroit sur un jeu, à comparer le nombre des cas où arrivera un certain événement au nombre des cas où il n'arrivera pas. Cela fait, on sauroit au juste, et on exprimeroit par des nombres de combien le parti qu'on prendroit seroit le meilleur. Toute la difficulté est qu'il nous échappe beaucoup de cas où l'événement peut arriver, ou ne pas arriver, et plus il y a de ces cas inconnus, plus la connoissance du parti qu'on doit prendre paroît incertaine. La suite de ces idées a conduit M. Bernoulli à cette question, si le nombre des cas inconnus diminuant toujours, la probabilité du parti qu'on doit prendre en augmente nécessairement, de sorte qu'elle vienne à la fin à tel degré de certitude qu'on voudra. Il semble qu'il n'y ait pas de difficulté pour l'affirmative de cette Proposition, cependant M. Bernoulli qui possédoit fort cette matière, assuroit que ce Problème étoit beaucoup plus difficile que celui de la Quadrature du cercle, et certainement il seroit sans comparaison plus utile. Il n'est pas si glorieux à l'Esprit de Géométrie de régner dans la Phisique, que

1. L'*Éloge de M. Bernoulli* rédigé par Joseph Saurin fut publié en 1706 dans le *Journal des Savants*, p. 81 et suiv. (édition de Paris), puis dans les *Mémoires de l'Académie Royale des Sciences*, année 1705, imprimés en 1730, p. 139-150.

dans les choses Morales, si compliquées, si casuelles, si changeantes ; plus une matière lui est opposée, et rebelle, plus il a d'honneur à la dompter. »

On doit à Jacob Bernoulli, dans son traité, la distinction entre le concept de probabilité objective ou statistique, qui est utilisé pour décrire les propriétés de mécanismes aléatoires comme les jeux de hasard, et celui de probabilité subjective utilisé pour mesurer le degré de crédibilité d'une proposition basée sur une certaine évidence qui n'a pas besoin d'être de nature statistique.

L'histoire de la publication de l'*Ars conjectandi* en 1713 est assez complexe. Dans son article *The Significance of Jacob Bernoulli's Ars Conjectandi for the Philosophy of probability today*¹ Glenn Shafer indique que Saurin, Leibniz et Montmort écrivirent à Johann Bernoulli pour lui demander de compléter le livre de son frère et le publier. Celui-ci aurait répondu que les héritiers de Jacob ne le souhaitaient pas. En fait, Nicholas I et Jacob Hermann avaient eu accès au manuscrit de l'*Ars conjectandi*, et Hermann l'avait sans doute mis en ordre mais il partit enseigner à Padoue en 1707 et Nicholas rédigea sa thèse jusqu'en 1709. Finalement, en 1710, écrit Shafer, Montmort a écrit à Johann en proposant de payer le coût de la publication de l'*Ars conjectandi*. L'offre, selon Shafer ne fut pas acceptée. On ne sait exactement le rôle joué par les différents intervenants pour cette édition mais certaines publications éclairent la question.² On est cependant certain que Nicholas I s'est intéressé de près à l'édition du livre de son oncle et était en contact avec l'éditeur. Le 9 septembre 1713, il écrit en effet à Montmort : « Le livre de feu mon oncle vient de sortir de la presse, le libraire m'a dit qu'il en a envoyé un exemplaire par la poste à M. Koenig [...] »

5.1.2. Les quatre parties de l'ouvrage

La première partie de l'*Ars conjectandi* est un commentaire du traité de Huygens *De ratiociniis in aleae ludo*³ que Bernoulli complète et mathématise par des formules.

Le terme *épreuve de Bernoulli* a été adopté depuis la publication de l'*Ars conjectandi* pour définir une expérience aléatoire de paramètre p (réel compris entre 0 et 1) et comportant deux issues, le succès ou l'échec, p représentant la probabilité d'un succès et $q = 1 - p$ représentant la probabilité d'un échec. En effet, dans cette première partie de son livre, Bernoulli indique que la probabilité d'un événement est constante pour un jeu unique et qu'elle est indépendante des résultats de jeux précédents.

1. www.glennshafer.com/assets/downloads/articles/article55.pdf

2. Cf. N. Meusnier, *Nicholas, nouveau exemplaire*, Electronic journal for probability and statistics, vol. 2, n° 1, juin 2006, www.jehps.net

3. Voir notre chapitre consacré à Huygens.

Bernoulli donne une démonstration alternative à celle donnée par Huygens concernant sa très connue troisième proposition, à savoir : Avoir p chances d'obtenir a et q chances d'obtenir b donne une attente (espérance mathématique) égale à $\frac{pa + qb}{p + q}$:

Il formule comme suit cette démonstration¹ :

« Voici une autre démonstration de la même règle : Supposons un nombre de joueurs égal au nombre des cas en général, c'est-à-dire, à $p + q$ de manière qu'il y ait un cas pour chaque joueur ; supposons, par exemple, qu'il y ait $p + q$ boîtes, et qu'on ait caché dans chacune ce que donne l'une des deux espèces de cas, savoir, a dans chacune des boîtes dont le nombre est p , et b dans chacune de celles dont le nombre est q . Si chacun des joueurs reçoit une de ces boîtes, ils en recevront ensemble la totalité, et ils ne pourront manquer d'avoir conjointement tout ce qu'elles contiennent, savoir, $pa + qb$ et comme ils ont tous une attente égale, il faudra diviser ce qu'ils recevront ensemble, par le nombre de joueurs, ou par le nombre des cas : d'où il s'ensuit que l'attente de chacun vaudra $\frac{pa + qb}{p + q}$. On démontrera de la même manière que si j'ai p cas pour a , q cas pour b et r pour c , mon sort sera $\frac{pa + qb + rc}{p + q + r}$.

Corollaire : De là il résulte, que si j'ai p cas pour avoir a , et q cas pour avoir 0, mon attente vaudra $\frac{pa}{p + q}$ etc. »

La seconde partie de l'ouvrage traite de l'analyse combinatoire en neuf chapitres. Le terme *permutation* y est introduit pour la première fois selon Smith².

Bernoulli définit ainsi les permutations et les combinaisons :

« J'appelle permutations, les variations des choses, selon lesquelles la même quantité de choses étant conservée, l'ordre et la position entre ces choses est modifié de diverses façons. »

« Les combinaisons des choses sont les assemblages selon lesquels on ôte quelques-unes des choses qui constituent une multitude déterminée, ces assemblages ne tenant aucun compte de l'ordre ou de la situation des choses. »

Bernoulli démontre diverses règles permettant le calcul des permutations et des combinaisons.

Le chapitre III de cette deuxième partie traite des nombres figurés et donne une table équivalente au Triangle de Pascal.

On trouve également dans cette Partie, la définition des nombres de Bernoulli.

1. Page 8 de l'édition originale de 1713.

2. *History of Mathematics*, vol. 2, p. 528

Les nombres de Bernoulli sont liés aux polynômes de Bernoulli, notés $B_n(x)$, que l'on peut définir par récurrence au moyen des trois conditions :

- $B_0(x) = 1$
- pour tout n : $B_{n+1}(x) = (n+1) B_n(x)$
- $\int_0^1 B_n(t) dt = 0$

On a, de plus : $B_n(x+1) - B_n(x) = n x^{n-1}$. Ces polynômes sont utiles dans des problèmes d'interpolation (approximation de fonctions par des expressions polynomiales). On retrouve ces polynômes dans le développement en série :

$$\frac{te^{xt}}{e^t - 1} = B_0(x) + B_1(x).t + B_2(x).\frac{t^2}{2!} + \dots + B_n(x).\frac{t^n}{n!} + \dots$$

Lorsque $x = 0$, on obtient le développement de $t/(e^t - 1)$ et les nombres de Bernoulli apparaissent en tant que $b_n = B_n(0)$.

Ce même troisième chapitre de la deuxième partie de l'*Ars Conjectandi*, contient, selon Francis Maseres, la première démonstration du théorème du binôme de Newton. La formule du binôme a été publiée pour la première fois, sans démonstration, par John Wallis en 1685. Maseres donne ces indications dans son ouvrage *The doctrine of permutations and combinations*¹. Il affirme que la lecture des seuls trois premiers chapitres de cette seconde partie permet de comprendre la démonstration. Il reproduit donc le texte original de Bernoulli en latin puis donne sa propre traduction en anglais de ces trois chapitres.

La troisième partie de l'ouvrage est relative à l'emploi de la doctrine des combinaisons développée dans la partie précédente.

Si l'on résume le résultat de Bernoulli on peut écrire ce qui suit :

Soit une expérience aléatoire conduisant à deux éventualités : l'une appelée succès de probabilité p , à laquelle on associe la valeur 1 ; l'autre appelée échec de probabilité $q = 1 - p$, à laquelle on associe la valeur 0. On peut définir alors une variable aléatoire X à valeurs dans $\{0,1\}$, dite de Bernoulli.

Répétons n fois une telle épreuve de Bernoulli, ces épreuves successives étant indépendantes, et appelons X_i la variable aléatoire associée à l'expérience d'indice i . Les lois des X_i étant identiques [$\text{Prob}\{X = 1\} = p$ et $\text{Prob}\{X = 0\} = q$], en posant $A = X_1 + X_2 + \dots + X_n$, Bernoulli démontre la loi binomiale :

$$\text{Prob}\{A = k\} = C_n^k p^k q^{n-k}$$

où C_n^k désigne le nombre de combinaisons de k objets parmi n .

L'appellation loi binomiale est due au lien avec le développement de la formule du binôme de Newton $(p + q)^n$: on sait que la somme des probabilités $\text{Prob}\{A = k\}$ pour k variant de 0 à n doit être égale à 1 ; or $(p + q) = 1$ et, précisément, selon la formule de Newton :

$$(p + q)^n = C_n^0 p^0 q^n + C_n^1 p^1 q^{n-1} + C_n^2 p^2 q^{n-2} + \dots + C_n^{n-1} p^{n-1} q^1 + C_n^n p^n q^0$$

Cette troisième partie de l'*Ars Conjectandi* traite vingt-quatre exemples, plus ou moins complexes, de calculs de gains dans divers jeux de hasard.

1. Londres, White, 1795, pages 233-234.

La quatrième partie, restée inachevée, contient les pensées philosophiques de Bernoulli concernant les probabilités ainsi que les applications économiques et morales. Le théorème de Bernoulli (loi faible des grands nombres) est démontré dans cette quatrième partie. Avec les notations précédentes, si k est le nombre de succès obtenu au cours des n expériences, la variable aléatoire A/n , moyenne arithmétique des X_i , prend la valeur k/n . La célèbre loi exprime que plus n est grand, plus il est probable que k/n soit voisin de p [Quel que soit $\epsilon > 0$, lorsque n tend vers l'infini, $\lim Pr\{|k/n - p| < \epsilon\} = 1$].

Cette quatrième partie est intitulée *Usage et application des doctrines précédentes dans les Affaires civiles, morales et économiques*.

Au chapitre II intitulé *De la Science et de la Conjecture* (page 213), Bernoulli donne sa conception de la probabilité :

« Nous disons *savoir* ou *comprendre* ce qui est certain et indubitable ; *conjecturer* seulement ou *opiner*, tout le reste.

Conjecturer quelque chose, est mesurer sa probabilité ; et pour cette raison *l'Art de Conjecturer* ou *Stochastique*, est défini par nous comme l'art de mesurer exactement les probabilités des choses, afin que nous puissions toujours choisir dans nos actions, ce qui est préférable ou plus habile ; la sagesse du Philosophe et la prudence du Politique sont d'ailleurs dirigées en ce sens.

Les probabilités sont exprimées d'après *le nombre*, et d'après *le poids des arguments*, qui prouvent ou indiquent en quelque sorte que quelque chose arrive, arrivera ou est arrivé. D'autre part par *Poids*, j'entends la force de prouver.

Les *Arguments* eux-mêmes sont soit *intrinsèques*, communément artificiels, [...] soit *extrinsèques* et non artificiels, résultant de l'autorité ou du témoignage des hommes. »

Au chapitre III, *Comment les poids des arguments sont estimés en supputant les probabilités des choses*, Bernoulli, distingue trois types d'arguments : ceux qui existent nécessairement et indiquent avec contingence, ceux qui existent avec contingence mais indiquent nécessairement et ceux qui existent et indiquent à la fois avec contingence c'est-à-dire non nécessairement. Ses calculs du poids des arguments observent, de plus, une distinction entre des arguments purs et des arguments mixtes, ces derniers pouvant prouver, selon les cas, une chose ou son contraire.

Le chapitre IV, *La double manière de chercher le nombre de cas. Ce qu'il faut penser de la manière de l'établir par l'expérience*, introduit la distinction entre probabilités définies *a priori* et *a posteriori*. Le premier cas est celui des jeux de hasard pour lesquels les probabilités sont connues par l'exposé du jeu et où l'on suppose généralement l'équiprobabilité comme dans les jeux de pile ou face, ce qui ramène le calcul des probabilités à un problème de dénombrement. Le second cas est celui où la probabilité n'est pas connue *a priori*. L'exemple donné par Bernoulli est celui de l'urne dont on ne connaît pas le

contenu. Dans ce cas c'est l'itération d'un grand nombre d'expériences qui mènera à l'estimation de la probabilité.

La première loi des grands nombres est donc introduite par Bernoulli à partir de la constatation de l'évidence de la méthode de détermination *a priori* des probabilités dans des jeux tels que le jeu de dés, car il suffit de dénombrer les cas possibles et les cas favorables. Mais, dit-il, qu'en est-il lorsqu'il s'agit de prévision du temps ou de jeux d'adresse lorsque les causes sont cachées et l'énumération des cas impossible ? Bernoulli choisit alors la détermination de la probabilité des cas favorables *a posteriori*. Il indique, page 224 que l'on doit considérer qu'un événement particulier doit arriver dans le futur aussi souvent qu'il a été donné d'observer sa réalisation ou sa non-réalisation dans le passé. Bernoulli est en fait le premier à mathématiser cette notion évidente mais vague selon laquelle l'incertitude décroît avec le nombre d'observations. Il ajoute, à la page suivante, que même pour le plus stupide des hommes, l'instinct naturel, et sans aucune instruction, ce qui est remarquable, le conduit à être convaincu que plus on a fait d'observations, moins le danger de s'écarter de son but est grand.

Il donne page 224, la définition de son approche fréquentiste :

« Ce qu'il n'est pas possible d'extraire *a priori*, il sera possible de l'extraire *a posteriori* d'après le résultat des exemples semblables souvent observés ; car il est présumé que ce qui peut ensuite arriver ou ne pas arriver, arrivera ou n'arrivera pas dans un nombre de cas identique au nombre observé précédemment dans des cas semblables. »

Bernoulli a donc entrepris de démontrer que la « certitude morale » peut être approchée autant qu'on le désire en multipliant les observations. Il pose le problème (page 226) :

« Je suppose qu'à votre insu, sont placées dans une urne 3 000 billes blanches et 2 000 billes noires et que pour tenter de déterminer ces nombres de billes vous tirez une bille après l'autre, (en prenant soin de remettre les billes dans l'urne après chaque tirage de telle sorte que leur nombre ne soit pas diminué) et que vous observez combien de fois sortirait le blanc et combien de fois le noir. Il est cherché si vous pouvez faire cela un nombre si élevé de fois qu'il devient dix fois, cent fois, mille fois etc. plus probable, c'est-à-dire plus moralement certain que le nombre de billes blanches et noires tirées sont dans le rapport 3/2 comme elles le sont dans l'urne plutôt que dans un rapport différent ».

Et il ajoute :

« Pour éviter que cela ne soit compris autrement qu'il faut, nous devons noter que le rapport que nous tentons de déterminer par des expériences [...] doit être recherché, non précisément et dans l'indivisible, mais avec une certaine latitude, c'est-à-dire entre deux limites qui peuvent être aussi proches que nous le désirons. Par exemple, dans le cas des billes évoqué plus haut, si nous prenons deux rapports 301/200 et

299/200 ou 3 001/2 000 et 2 999/2 000, etc. qui sont immédiatement l'un supérieur et l'autre inférieur au rapport 3/2, il sera montré que l'augmentation du nombre de tirages mènera à une probabilité plus grande de tomber à l'intérieur des limites spécifiées qu'à l'extérieur de ces limites. »

Bernoulli ajoute :

« Ce problème est celui que je me suis proposé de révéler en ce lieu après lui avoir consacré vingt ans de travaux. D'une part sa nouveauté, d'autre part sa difficulté et sa très grande utilité font que sa solution dépasse en importance et en valeur tous les autres chapitres de la présente doctrine. »

Voici le texte de Bernoulli¹ relatif à ce que nous appelons son théorème :

« *Propos. Princip.* [...] Sit igitur numerus casuum fertilium ad numerum sterilium vel præcise, vel proxime in ratione r/s , adeoque ad numerum omnium in ratione $r/(r+s)$ seu r/t , quam rationem terminent limites $(r+1)/t$ et $(r-1)/t$.

Ostendum est, tot posse capi experimenta, ut datis quotlibet (puta c) vicibus, versimiliter evadut, numerum fertilium observationum intra hos limites quam extra casurum esse, h.e. numerum fertilium ad numerum omnium observationum rationem habiturum nec maiorem quam $(r+1)/t$, nec minorem quam $(r-1)/t$. »

Que nous traduisons comme suit :

« Admettons que le nombre de cas fertiles est exactement ou approximativement dans le rapport r/s , au nombre de cas stériles, si bien que le rapport du nombre de cas fertiles à la totalité des cas est $r/(r+s)$ ou r/t (avec $t = r+s$) qui est compris entre les limites $(r+1)/t$ et $(r-1)/t$. Il est possible d'effectuer un nombre d'essais assez élevé pour qu'il soit plus probable, dans le rapport de n'importe quel nombre donné (disons c), que le nombre de cas fertiles soit compris entre les limites $(r+1)/t$ et $(r-1)/t$, que le cas contraire. »

En termes modernes nous dirions : Si une épreuve de Bernoulli, ayant une probabilité r/t , est répétée un grand nombre de fois, on peut choisir un nombre c aussi grand que l'on veut pour que la probabilité que la fréquence de réalisation des cas favorables soit comprise entre les limites $(r+1)/t$ et $(r-1)/t$, soit c fois plus élevée que la probabilité du cas inverse, c'est-à-dire plus grande que $c/(c+1)$.

5.1.3. La démonstration du théorème de Bernoulli par son auteur

Le chapitre V de la quatrième partie de l'*Ars Conjectandi* (pages 228 à 237) donne une démonstration très complexe du théorème.

Cette démonstration est basée sur 5 lemmes dont nous donnons les énoncés sans démonstrations (qui peuvent être trouvées par exemple dans l'article de Sheynin référencé plus bas).

1. Page 236 de l'édition originale de 1713.

$\frac{mn-r}{r+1}$ & $ms + \frac{mrs-rs}{r+1}$, omnes simul termini inter utrumque limitem T & A comprehensi multo pluribus quàm c vicibus superabunt omnes simul terminos extra limites ab utraque parte protensos. Roperia igitur est finita potestas, quæ optatam habeat proprietatem. Q. E. F.

Propos. Princip. Sequitur tandem Propositio ipsa, cujus gratia hæc omnia dicta sunt, sed cujus nunc demonstrationem sola Lemmatum præmissorum applicatio ad præsens institutum absolvet. Ut circumlocutionis tedium vitem, vocabo casus illos, quibus eventus quidam contingere potest, *fecundos seu fertiles*; & *steriles* illos, quibus idem eventus potest non contingere: nec non experimenta *facienda* sive *fertilia* illa, quibus aliquis casuum fertilium evenire deprehenditur; & *infacienda* sive *sterilia*, quibus sterilius aliquis contingere observatur. Sit igitur numerus casuum fertilium ad numerum sterilius vel præcisè vel proximè in ratione $\frac{r}{1}$, adeoque ad numerum omnium in ratione $\frac{r}{r+1}$, seu $\frac{r}{1}$, quam rationem terminent limites $\frac{r+1}{1}$ & $\frac{r-1}{1}$. Ostendendum est, tot posse capi experimenta, ut datis quotlibet (puta c) vicibus verisimilius evadat, numerum fertilium observationum intra hos limites quàm extra casurum esse, h. e. numerum fertilium ad numerum omnium observationum rationem habiturum nec majorem quàm $\frac{r+1}{1}$, nec minorem quàm $\frac{r-1}{1}$.

« **LEMME 1.** Étant donné un ensemble de nombres naturels $0, 1, 2, 3, \dots, r-1, r, r+1, \dots, r+s$. continué de telle sorte que le dernier terme soit un multiple de $r+s$, cet ensemble peut s'écrire :

$0, 1, 2, 3, \dots, nr-n, \dots, nr, \dots, nr+n, \dots, nr+ns$.

Si on fait croître n , le nombre des termes compris entre nr et $nr-n$ ou $nr+n$, de même que ceux qui sont compris entre ces derniers et 0 ou $nr+n$, augmente, mais aussi grand que soit n , le nombre de termes plus grands que $nr+n$ ne sera pas plus de $s-1$ fois plus grand que le nombre de termes compris entre nr et $nr+n$. De même le nombre de termes inférieurs à $nr-n$ ne sera pas supérieur à $r-1$ fois le nombre de termes compris entre $nr-n$ et nr .

LEMME 2. Toute puissance entière d'un binôme $r+s$ est développée en un nombre de termes égal à l'exposant + 1.

LEMME 3. Le développement du binôme $r+s$ avec un exposant qui est, du moins, un multiple entier de $r+s=t$, par exemple $n(r+s)=nt$, comporte un terme M qui est le plus grand si le nombre de termes qui le précèdent est au nombre de termes qui le suivent dans le rapport s à r , ou ce qui revient au même si les exposants des lettres r et s dans ce terme sont dans le rapport des grandeurs r et s elles-mêmes. Le terme le plus proche de chaque côté de M est plus grand que le terme le plus distant du même côté. Cependant ce même terme M est dans un rapport moindre à un terme plus proche que le rapport de ce terme plus proche à un autre plus distant, sous réserve que le nombre de termes intermédiaires soit le même dans les deux cas.

[Bernoulli adopte donc une loi binomiale pour la variable N du nombre d'échecs, avec

les notations $p = \frac{r}{t}$, probabilité de succès, appelé cas fertile, $q = \frac{s}{t}$, probabilité

d'échec, appelé cas stérile et nt nombre d'épreuves. On a :

$$P(N=k) = \frac{nt.(nt-1).(nt-2)....(nt-k-1)}{1.2.3....k} p^{nt-k} q^k ;$$

avec des notations modernes le maximum M est le terme le plus grand du développe-

ment du binôme $(r+s)^{nt} = \sum_{k=0}^{nt} \binom{nt}{k} . r^{nt-k} s^k$, soit $\binom{nt}{ns} . r^{nt-ms} s^{ns} = \frac{(nt)!}{(nr)!(ns)!}$.]

LEMME 4. Dans le développement d'un binôme d'exposant nt , n peut être pris assez grand pour que le rapport du terme le plus grand M à chacun des deux termes L et Λ distants de M de n termes sur la droite et sur la gauche, soit aussi grand qu'on le désire.

LEMME 5. Dans le développement d'un binôme d'exposant nt , n peut être pris assez grand pour que le rapport de la somme de tous les termes depuis le plus grand, M, jusqu'aux termes L et Λ à la somme des termes restants, soit aussi grand qu'on le désire. »

Suite à l'énoncé du théorème donné plus haut, Bernoulli observe :

« À partir d'une probabilité donnée r/t d'un cas fertile, et d'une probabilité s/t d'un cas stérile, avec $r+s=t$, en nt essais, le nombre de cas possibles avec tous les cas

fertiles, ou tous les cas fertiles moins 1, ou tous les cas fertiles moins 2 etc. sont $r^{nt}, \binom{nt}{1} r^{nt-1} s, \binom{nt}{2} r^{nt-2} s^2, \dots$ etc., c'est-à-dire les termes du développement du binôme $r + s$ élevé à la puissance nt . D'après les lemmes précédents, le nombre d'essais avec nr cas fertiles et ns cas stériles correspond au terme maximal M , puisque selon le lemme III, ns termes le précèdent et nr termes le suivent. De plus, le nombre d'essais avec $nr - n$ ou $nr + n$ sorties fertiles, les autres étant stériles, est la somme des termes entre les deux limites L et Λ (n termes de part et d'autre de M) définies dans le lemme IV. En conséquence, le nombre total de cas dans lesquels il n'y a pas plus de $nr + n$ et pas moins de $nr - n$ observations fertiles, est exprimé par la somme des termes situés entre les limites L et Λ . Le nombre total des autres cas dans lesquels il arrive plus ou moins d'observations fertiles est exprimé par la somme des termes au-delà des limites L ou Λ . L'exposant du binôme peut être choisi assez grand pour que la somme des termes compris entre les limites L et Λ soit plus de c fois plus grande que la somme de tous les termes restant à l'extérieur de l'intervalle L et Λ (d'après les lemmes IV et V), il s'ensuit que le nombre d'observations peut être choisi assez grand pour que le nombre d'essais pour lesquels le rapport du nombre de cas fertiles au nombre total de cas ne dépasse pas les bornes $(nr + n)/nt$ et $(nr - n)/nt$ ou $(r + 1)/t$ et $(r - 1)/t$, soit plus grand que c fois le nombre de cas restants. Donc, il devient c fois plus probable que le rapport du nombre de cas fertiles au nombre de cas total soit compris entre les limites $(r + 1)/t$ et $(r - 1)/t$ plutôt qu'à l'extérieur de cet intervalle. Quod erat demonstrandum. »

Les dernières pages (237 à 239) sont consacrées à un calcul, qui complète le théorème de la limite, afin de donner une valeur minimale du nombre d'épreuves nécessaires pour que la probabilité que la fréquence de réalisation h des cas favorables soit comprise entre deux limites spécifiées. L'exemple de Bernoulli est une épreuve ayant une probabilité de 0,6 ; une valeur du nombre c (défini plus haut) de 1 000 et une fourchette pour h autour de 0,6 de 0,02. Il obtient :

$$P(0,58 \leq h \leq 0,62) > \frac{1000}{1001} \text{ pour } 25\,550 \leq n.$$

En fait, Bernoulli donne ici la première détermination d'un intervalle de confiance et, avant ce calcul, il avait souligné¹ que le rapport entre le nombre de cas favorables (*fertiles*) et défavorables (*stériles*) déterminé par l'itération de l'expérience ne doit pas être accepté comme précis et absolu (*non praecise et in indivisibili acceptam*) mais comme compris entre des limites que l'on peut rapprocher autant que l'on veut.

1. Page 226 de l'édition originale de 1713.

5.1.4. La démonstration du théorème de Bernoulli donnée par Poisson

Nous avons indiqué que la démonstration donnée par Bernoulli de son théorème était complexe¹. S.-D. Poisson² en a donné la démonstration suivante :

« Le théorème [...] est dû à Jacques Bernoulli, qui en avait médité la démonstration pendant vingt années. Celle qu'il a donnée se déduit de la formule du binôme au moyen des propositions suivantes.

Soient, à chaque épreuve, p et q les chances données des deux événements contraires E et F ;

soient aussi g, h, k , des nombres entiers, tels que l'on ait

$$p = g/k, q = h/k, g + h = k, p + q = 1 ;$$

désignons par m, n, μ , d'autres nombres entiers, liés à g, h, k , par les équations

$$m = gk, n = hk, \mu = m + n = (g + h)k,$$

de manière que les chances p et q soient entre elles comme les nombres m et n , que l'on pourra rendre aussi grands qu'on voudra en augmentant convenablement g, h, k , sans changer leur rapport. Cela posé :

1°. Dans le développement de $(p + q)^\mu$, le terme le plus grand sera celui qui répond au produit $p^m q^n$ et comme ce terme est la probabilité de l'arrivée de m fois E et de n fois F (selon nos considérations précédentes) il s'ensuit que cet événement composé, c'est-à-dire, l'arrivée des événements en raison directe de leurs chances respectives, est le plus probable de tous les événements composés qui peuvent avoir lieu dans un nombre quelconque m d'épreuves.

2°. Si ce nombre μ est très grand, le rapport du plus grand terme du développement de $(p + q)^\mu$ à la somme de tous les termes, ou à l'unité, sera une très petite fraction, qui diminuera indéfiniment à mesure que μ augmentera encore davantage ; par conséquent, dans une longue série d'épreuves, l'événement composé le plus probable, le sera cependant très peu, et de moins en moins à mesure que les épreuves seront plus longtemps prolongées.

3°. Mais si l'on considère dans le développement de $(p + q)^\mu$, son plus grand terme, les l termes qui le suivent et les l termes qui le précèdent, et si l'on désigne par λ la somme de ces $2l + 1$ termes consécutifs, on pourra toujours, sans changer ni p ni q , prendre μ assez grand pour que la fraction λ diffère de l'unité, d'autant peu qu'on voudra ; et à mesure que μ augmentera encore davantage, λ approchera de plus en plus d'être égal à un. On conclut de là que dans une longue série d'épreuves, il y a toujours une grande probabilité λ que l'événement E arrivera un nombre de fois compris entre les limites $m \pm l$ et F un nombre de fois compris entre $m - l$ et $m + l$; de sorte que sans changer l'intervalle $2l$ des limites de ces deux nombres, on pourra

1. On trouvera, par exemple, dans *Histoires de probabilités et de statistiques*, Paris, Ellipses, 2004, au chapitre 3 une analyse de la démonstration de Bernoulli par Michel Henry. Voir également : *Jakob Bernoulli, On the law of large numbers. The Art of Conjecturing; Part Four*, Translated in English by Oscar Sheynin, Berlin, 2005 ; ISBN 3-938417-14-5.
2. Siméon-Denis Poisson, *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*, Paris, 1837, chap. 51 page 135.

rendre le nombre μ des épreuves assez grand pour que la probabilité λ soit aussi approchante qu'on voudra de la certitude. Si l'on prend les rapports de ces limites au nombre μ des épreuves, que l'on ait égard aux équations précédentes, et qu'on fasse

$$\frac{l}{\mu} = \delta, p \pm \delta = p', q \pm \delta = q'.$$

ces rapports seront p' et q' ; et comme la fraction δ diminuera indéfiniment à mesure que μ augmentera, il s'ensuit que ces rapports, variables avec μ , approcheront aussi indéfiniment, et avec une très grande probabilité, des chances p et q de E et F; ce qui est l'énoncé du beau théorème de Jacques Bernoulli. »

5.1.5. Commentaires

L'article *Probabilité* de l'encyclopédie de d'Alembert et Diderot souligne le fait que le livre de Bernoulli est le plus important des livres alors parus qui « servent à déterminer la vraisemblance des événements, et les degrés par lesquels nous parvenons à la certitude morale. »

Commentant les travaux de Jacob Bernoulli, Charles Gouraud, dans son *Histoire du calcul des probabilités* (1848) écrit :

« Montmort donna en 1713 une édition bien plus correcte de son Essai, à la suite de laquelle il eut le bon goût de publier la lettre de Johann Bernoulli et toute sa correspondance avec son neveu. Cette nouvelle publication fut reçue avec plaisir; elle eût été sans doute mieux accueillie encore, si cette année-là même Nicolas Bernoulli n'avait tiré enfin des papiers de son oncle, cet *Ars conjectandi* qui changeait la face du Calcul des probabilités et dont l'originalité de génie, captiva, dès qu'il parut, l'attention universelle.

On a vu à combien de titres il méritait de l'obtenir. La doctrine philosophique qu'il donnait pour base à une théorie dénuée jusque-là de ses premiers fondements, le principe dont il dotait une analyse réduite auparavant à quelques règles élémentaires, les perspectives qu'il ouvrait enfin à un calcul resserré soixante ans dans les limites les plus restreintes, tout cela était bien fait pour frapper les esprits et pour donner à la découverte de Pascal et de Fermat la popularité qu'elle avait si longtemps et si inutilement implorée. Un cri unanime s'éleva dans le monde de la science et annonça à la postérité que l'Analyse des hasards allait entrer dans une ère nouvelle et commençait définitivement sa fortune. Cette fortune commença comme on devait s'y attendre par un élan général des esprits vers l'étude des applications.

On avait admiré la constitution philosophique octroyée par Bernoulli au nouveau calcul; on avait surtout, et avec raison, fait le plus grand cas du principe original qu'il avait si heureusement approprié à son usage; mais ce qui devait séduire et ce qui charma en effet toutes les imaginations, ce fut la conquête de ce monde inconnu que l'illustre géomètre avait désigné aux futures entreprises de l'algèbre; ce qui arrêta d'abord et prit toute l'attention, ce furent les derniers mots de ce testament intellectuel où Bernoulli semblant léguer à son siècle la suite et l'achèvement de ses travaux, avait montré du geste au Calcul des probabilités la riche carrière de l'économie civile et politique... »

Les travaux de Bernoulli sur les sommes des puissances des nombres entiers et sur la loi faible des grands nombres qui établit que si une épreuve est répétée un grand nombre de fois, il devient improbable que la moyenne des résultats des n premières épreuves s'écarte sensiblement de l'espérance du résultat d'une épreuve quand n augmente indéfiniment, seront repris par De Moivre, Poisson et Laplace, comme nous l'indiquerons plus loin.

Il faut enfin signaler que les Œuvres de Jacob Bernoulli ont été éditées en deux volumes, à Genève en 1744, par Cramer, sous la direction de Nicholas I sous le titre *Jacobi Bernoulli Basiliensis Opera*.

5.2. Daniel Bernoulli

5.2.1. *Specimen theoriae novae de mensura sortis*

Le concept d'espérance morale a été introduit par Daniel Bernoulli dans un mémoire intitulé *Specimen theoriae novae de mensura sortis* (Exposition d'une nouvelle théorie de la mesure du risque) qui fut publié¹ en 1738.

Il y indique que tous les « géomètres » qui se sont intéressés à « la mesure des sorts » ont considéré que l'espérance d'un gain s'obtient « par la multiplication de la *valeur* de chaque lot par le nombre de cas où son obtention est certaine, sommation de ces produits et division par le nombre total de cas, ces cas étant, entr'eux également possibles. »

Cette définition de l'espérance mathématique appelée, Daniel Bernoulli tente de préciser ce qu'il entend par *espérance morale* :

« [Je considère] le cas d'un homme très pauvre attendant d'un tirage au sort avec une égale probabilité, soit rien, soit vingt mille ducats ; doit-il estimer son sort dix mille ducats et prend-il une mauvaise décision en vendant son droit neuf mille ducats ? Il ne me le semble pas, bien que je crois qu'un homme riche irait contre son intérêt si, pouvant acquérir ce droit à ce prix, il s'y refusait. Si je ne me trompe pas, il est évident qu'on ne peut imposer le même étalon à tous les hommes pour mesurer le sort, ni, par conséquent, utiliser la règle [indiquée plus haut]. Mais si on est désireux d'examiner attentivement le problème, on s'aperçoit qu'on peut donner une définition du mot *valeur* utilisé dans cette règle, de façon que tout ce qui en découle s'impose sans réserve à tous, à savoir celle de *valeur* estimée non pas à partir du prix de l'objet mais à partir de l'utilité (*emolumentum*) que chacun peut en tirer. Le prix estimé à partir de l'objet lui-même est le même pour tous, *L'utilité* dépend de la condition de chacun. Il n'y a pas de doute que mille ducats offrent plus d'avantages pour un pauvre que pour un riche, bien que le montant soit le même pour l'un comme pour l'autre. »

1. *Commentarii Academiae Scientiarum Petropolitanae*, V, p. 175-192, (1730-1731), 1738. Une traduction en Anglais de L. Sommer a été publiée dans *Econometrica*, vol. 22, n° 1, 1954.

Laplace a défini, lui aussi, à plusieurs reprises, ce concept d'espérance morale en ces termes :

« La probabilité des événements sert à déterminer l'espérance ou la crainte des personnes intéressées à leur existence, et c'est sous ce point de vue que la science des hasards est une des plus utiles de la vie civile. Ce mot *espérance* a différentes acceptions : il exprime ordinairement l'état de l'esprit humain lorsqu'il doit lui arriver un bien quelconque dans certaines suppositions qui ne sont que vraisemblables. Dans la théorie des chances, l'espérance est le produit de la somme espérée par la probabilité de l'obtenir. Pour distinguer les deux acceptions de ce terme, j'appellerai la première *espérance morale*, et la seconde, *espérance mathématique*. »

Daniel Bernoulli estime que si une personne possède une somme d'argent x et qu'il reçoit un incrément dx , la valeur relative de cet incrément est proportionnelle à dx et inversement proportionnelle à x .

Laplace appelle x la fortune physique et y la fortune morale dont l'incrément, selon ce qui précède est

$$dy = k \frac{dx}{x} \text{ où } k \text{ est une constante.}$$

Donc $y = k \log x + \log h$

(En latin, Daniel Bernoulli appelle « y » *emolumentum*)

Daniel Bernoulli, après quelques calculs énonce la règle suivante concernant son concept d'espérance morale :

« Chaque gain doit être ajouté à la fortune initiale, puis cette somme doit être élevée à une puissance égale au nombre de façons dont ce gain peut être obtenu ; ensuite tous les termes de ce type correspondant aux divers gains doivent être multipliés entr'eux. On prend ensuite la racine de degré égal au nombre total de cas possibles du produit obtenu. Enfin on soustrait du résultat la fortune initiale. Ce qui reste indique le sort de la proposition. Ce principe est essentiel pour la mesure de la valeur des propositions comportant un risque. »

Laplace¹, qui reprend la théorie de Daniel Bernoulli, explicite cette règle comme suit :

« Supposons maintenant que la fortune physique d'un individu soit a , et qu'il lui survienne l'expectative l'un des accroissements α, β , etc., ces quantités pouvant être nulles ou même négatives, ce qui change les accroissements en diminutions. Représentons par $p, q, r, \dots$ les probabilités respectives de ces accroissements, la somme de ces probabilités étant supposée égale à l'unité. Les fortunes morales correspondantes de l'individu pourront être

$$k \log(a + \alpha) + \log h, k \log(a + \beta) + \log h, \text{ etc.}$$

1. *Théorie analytique des probabilités*, 1812, p. 432-445. On trouvera également dans S.-F. Lacroix, *Traité élémentaire du calcul des probabilités*, Paris, 1816, un chapitre intitulé « Espérance morale », qui reprend les calculs de Daniel Bernoulli et de Laplace.

En multipliant ces fortunes respectivement par leurs probabilités $p, q, r, \dots$, la somme de leurs produits sera la fortune morale de l'individu en vertu de son expectative ; en nommant donc Y cette fortune, on aura

$$Y = kp \log(a + \alpha) + kq \log(a + \beta) + \dots + \log b$$

Soit X la fortune physique qui correspond à cette fortune morale, on aura

$$Y = k \log X + \log b$$

La comparaison de ces deux valeurs de Y donne

$$X = (a + \alpha)^p (a + \beta)^q \dots$$

Si l'on retranche la fortune primitive a de celle valeur de X , la différence sera l'accroissement de la fortune physique qui procurerait à l'individu le même avantage moral qui résulte pour lui de son expectative. Cette différence est donc l'expression de cet avantage. »

Laplace expose ensuite les exemples de Daniel Bernoulli pour illustrer son propos.

5.2.2. Le paradoxe de Saint-Pétersbourg

Nous avons indiqué au chapitre consacré au livre de Montmort que la correspondance entre ce dernier et Nicolas Bernoulli, était incluse à la fin de la deuxième édition de l'ouvrage. Dans une lettre datée du 9 septembre 1713 (pages 401-402), Bernoulli proposait à Montmort cinq problèmes. Le quatrième était ainsi rédigé :

« A promet de donner un écu à B, si avec un dé ordinaire il amène au premier coup six points, deux écus s'il amène le six au second, trois écus s'il amène ce point au troisième coup, quatre écus s'il l'amène au quatrième et ainsi de suite ; on demande quelle est l'espérance de B. »

Le cinquième problème était celui-ci :

« On demande la même chose si A promet à B de lui donner des écus en cette progression 1, 2, 4, 8, 16, etc. ou 1, 3, 9, 27, etc. [...] au lieu de 1, 2, 3, 4, 5, etc. comme auparavant. »

Ce problème est devenu célèbre sous le nom de paradoxe de Saint-Pétersbourg, car il fut étudié par Daniel Bernoulli, cousin de Nicolas, mais aussi par Buffon, d'Alembert et beaucoup d'autres, car il soulevait un problème d'espérance mathématique face, en quelque sorte, à une espérance morale. Le mémoire de Daniel Bernoulli était intitulé *Specimen theoriae novae de mensura sortis* (Exposition d'une nouvelle théorie de la mesure du risque) et il fut publié en 1738.

Daniel Bernoulli modifiait comme suit l'énoncé du problème :

« Pierre jette en l'air une pièce et refait de même jusqu'à ce que, à la chute de la pièce, face apparaisse pour la première fois : si ceci se produit au premier jet, il doit donner à Paul un ducat ; si c'est au second, deux ducats ; au troisième, quatre ; au

quatrième, huit et ainsi de suite en doublant à chaque jet le nombre de ducats. On demande quelle est l'espérance de Paul. »

Il est évident que l'espérance mathématique de Paul est infinie, car la probabilité de l'événement qui correspond à l'apparition de face, seulement au $n^{\text{ième}}$ coup est :

$$\frac{1}{2^{n-1}} \cdot \frac{1}{2} = \frac{1}{2^n}$$

L'espérance E de gain de Paul est donc la somme :

$$E = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2^2} + 4 \cdot \frac{1}{2^3} + \dots + 2^{n-1} \cdot \frac{1}{2^n} + \dots$$

Tous les termes de la somme valent 1/2 et l'espérance de Paul est infinie. Il faudrait donc miser une infinité de ducats pour que le jeu soit équitable, ce qui est bien sûr impossible. En conséquence tout joueur sain d'esprit refusera de jouer à ce jeu car la mise initiale est trop élevée, en principe même infinie. Ce comportement d'apparence irrationnelle est tout simplement l'aversion au risque. Il a été formalisé par la notion de fonction d'utilité par Daniel Bernoulli. Ce dernier mentionne d'ailleurs que son cousin Nicolas souligne dans sa lettre que :

« bien que le sort espéré de Paul soit infini, personne sain d'esprit ne vendrait volontiers son espoir de gain pour vingt ducats. »

La raison de la différence entre le calcul mathématique et « l'estime vulgaire », selon Daniel Bernoulli qui cite Gabriel Cramer (1704-1752) vient de ce que dans un cas on estime l'argent à proportion de sa quantité, alors que les hommes de bon sens l'estiment à proportion de l'usage qu'ils peuvent en faire. Bernoulli suppose alors qu'on limite le gain à 2^{24} ducats, c'est-à-dire que si le gain du jeu intervient après le 25^e coup la somme gagnée sera encore fixée à 2^{24} ducats. L'espérance mathématique sera alors

$$1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2^2} + 4 \cdot \frac{1}{2^3} + \dots + 2^{24} \cdot \frac{1}{2^{25}} + 2^{24} \cdot \frac{1}{2^{26}} + 2^{24} \cdot \frac{1}{2^{27}} + \dots =$$

$$\frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots \text{ jusqu'à 24 termes, } + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots \text{ etc.} = 12 + 1 = 13.$$

Dans ce cas l'espérance est réduite à 13 ducats.

L'explication du paradoxe proposée consiste à faire la supposition de bon sens que le joueur Pierre n'est pas infiniment riche, et va donc cesser de payer au-delà d'une certaine somme, ici 2^{24} ducats et que le jeu cessant au 24^e coup, la mise de 13 ducats sera équitable. Du côté de Paul, au-delà de la somme de 2^{24} ducats, l'utilité marginale du gain devient décroissante.

On notera que l'on trouve une très grande similitude dans la définition de « l'espérance morale » qui vient d'être évoquée, chez Cramer, Daniel Bernoulli et Buffon. Ce dernier s'étend longuement sur le paradoxe de Saint-Pétersbourg et l'espérance morale dans son *Essai d'arithmétique morale*¹. Il écrit :

« Cette question m'a été proposée pour la première fois par feu M. Cramer, célèbre professeur de mathématiques à Genève, dans un voyage que je fis en cette ville, en l'année 1730 ; il me dit qu'elle avait été proposée précédemment par M. Nicolas Bernoulli à M. de Montmort, comme en effet on la trouve pages 402 et 407 de l'Analyse des jeux de hasard, de cet auteur : je rêvai quelque temps à cette question sans en trouver le nœud ; je ne voyais pas qu'il fût possible d'accorder le calcul mathématique avec le bon sens, sans y faire entrer quelques considérations morales ; et ayant fait part de mes idées à M. Cramer, il me dit que j'avais raison et qu'il avait aussi résolu cette question par des voies semblables ; il me montra ensuite sa solution à peu près telle qu'on l'a imprimée depuis dans les Mémoires de l'Académie de Pétersbourg, en 1738, à la suite d'un Mémoire excellent de M. Daniel Bernoulli, sur *la mesure du sort*, où j'ai vu que la plupart des idées de M. Daniel Bernoulli s'accordent avec les miennes, ce qui m'a fait grand plaisir. »

Buffon aborde ensuite le problème de la différence entre l'espérance mathématique et l'espérance morale, en citant une lettre qu'il avait adressée le 3 octobre 1730 à Gabriel Cramer :

« La raison de cette contrariété entre le calcul mathématique et le bon sens me semble consister dans le peu de proportion qu'il y a entre l'argent et l'avantage qui en résulte. Un mathématicien, dans son calcul, n'estime l'argent que par sa quantité, c'est-à-dire par sa valeur numérique : mais l'homme moral doit l'estimer autrement, et uniquement par les avantages ou le plaisir qu'il peut procurer ; il est certain qu'il doit se conduire dans cette vue, et n'estimer l'argent qu'à proportion des avantages qui en résultent, et non pas relativement à la quantité qui, passé de certaines bornes, ne pourrait nullement augmenter son bonheur : il ne serait, par exemple, guère plus heureux avec mille millions qu'il le serait avec cent, ni avec cent mille millions plus qu'avec mille millions : ainsi, passé de certaines bornes, il aurait très grand tort de hasarder son argent. Si, par exemple, dix mille écus étaient tout son bien, il aurait un tort infini de les hasarder ; et plus ces dix mille écus seront un objet par rapport à lui, plus il aura de tort. Je crois donc que son tort serait infini tant que ces dix mille écus feront une partie de son nécessaire, c'est-à-dire tant que ces dix mille écus seront absolument nécessaires pour vivre comme il a été élevé et comme il a toujours vécu. Si ces dix mille écus sont de son superflu, son tort diminue ; et plus ils seront une petite partie de son superflu, plus son tort diminuera : mais il ne sera jamais nul, à moins qu'il ne puisse regarder cette partie de son superflu comme indifférente, ou bien qu'il ne regarde la somme espérée comme nécessaire pour réussir dans un dessein qui lui donnera, à proportion, autant de plaisir que cette même somme est plus grande que celle qu'il hasarde, et c'est sur cette façon d'envisager un bonheur

1. Buffon G., « Essai d'arithmétique morale », in *Histoire naturelle, générale et particulière, servant de suite à l'Histoire naturelle de l'Homme*, Supplément, tome quatrième, Paris, Imprimerie Royale, 1777, p. 46-148.

à venir qu'on ne peut point donner de règles ; il y a des gens pour qui l'espérance elle-même est un plaisir plus grand que ceux qu'ils pourraient se procurer par la jouissance de leur mise. Pour raisonner donc plus certainement sur toutes ces choses il faudrait établir quelques principes :

Je dirais, par exemple, que le nécessaire est égal à la somme qu'on est obligé de dépenser pour continuer à vivre comme on a toujours vécu : le nécessaire d'un roi sera, par exemple, dix millions de rente (car un roi qui aurait moins serait un roi pauvre) ; le nécessaire d'un homme de condition serait dix mille livres de rente (car un homme de condition qui aurait moins serait un pauvre seigneur) ; le nécessaire d'un paysan sera cinq cents livres, parce qu'à moins d'être dans la misère il ne peut moins dépenser pour vivre et nourrir sa famille. Je supposerais que le nécessaire ne peut nous procurer des plaisirs nouveaux, ou, pour parler plus exactement, je compterais pour rien les plaisirs ou avantages que nous avons toujours eus, et d'après cela je définirais le superflu ce qui pourrait nous procurer d'autres plaisirs ou des avantages nouveaux : je dirais, de plus, que la perte du nécessaire se fait ressentir infiniment, qu'ainsi elle ne peut être compensée par aucune espérance ; qu'au contraire le sentiment de la perte du superflu est borné, et que, par conséquent, il peut être compensé. Je crois qu'on sent soi-même cette vérité lorsqu'on joue, car la perte, pour peu qu'elle soit considérable, nous fait toujours plus de peine qu'un gain égal ne nous fait de plaisir, et cela sans qu'on puisse y faire entrer l'amour-propre mortifié, puisque je suppose le jeu d'entier et pur hasard. Je dirais aussi que la quantité de l'argent dans le nécessaire est proportionnelle à ce qui nous en revient, mais que, dans le superflu, cette proportion commence à diminuer, et diminue d'autant plus que le superflu devient plus grand. »

Joseph Bertrand¹ a émis une critique féroce des propos de Daniel Bernoulli relatifs au paradoxe de Saint-Pétersbourg :

« La réponse la plus singulière faite au prétendu paradoxe est celle de Daniel Bernoulli [...].

Cent millions, suivant Daniel Bernoulli, ajoutés à une fortune déjà acquise de cent millions, ne suffisent pas pour la doubler. Quels avantages nouveaux peuvent-ils procurer ?

Il substitue, en conséquence, à l'espérance mathématique l'espérance morale, dans le calcul de laquelle une fortune dépend non du nombre d'écus dont elle se compose, mais des satisfactions qu'elle procure.

Le problème étant ainsi posé, Bernoulli a l'audace de le résoudre. La solution est simple. Un accroissement dx ajouté à une fortune x vaut $\frac{dx}{x}$. Celui dont la fortune était a et devient b gagne un avantage mesuré par

$$\int_a^b \frac{dx}{x} = \log b - \log a .$$

Jamais compte n'a été ni ne sera réglé de la sorte ; mais grâce à d'illustres approbations, la théorie de l'espérance morale n'a pas moins contribué à la célébrité de Daniel Bernoulli que ses admirables travaux de Physique.

Buffon a accru par son éloquence l'importance, je veux dire le retentissement de l'idée de Bernoulli.

1. *Calcul des probabilités*, Gauthier-Villars, 1888, p. 63-67. *

La théorie de l'espérance morale est devenue classique, jamais le mot ne put être plus exactement employé : on l'étudie, on l'enseigne, on la développe dans des livres justement célèbres. Le succès s'arrête là, on n'en a jamais fait et n'en pourra faire aucun usage.

L'importance d'une somme d'argent diminue avec la fortune de celui qui la reçoit. "L'avare, dit Buffon, est comme le mathématicien : tous deux estiment l'argent par sa quantité numérique. L'homme sensé n'en considère ni la masse ni le nombre. Il n'y voit que les avantages qu'il peut en tirer. Il raisonne mieux que le mathématicien. L'écu que le pauvre a mis à part pour payer un impôt de nécessité et l'écu qui complète les sacs d'un financier n'ont pour l'avare et le mathématicien que la même valeur. Celui-ci les comptera par unités égales, l'autre se les appropriera avec un plaisir égal, au lieu que l'homme sensé comptera l'écu du pauvre pour un louis et l'écu du financier pour un liard."

Un commentateur, qui n'est pas sans mérites, Quetelet, a ajouté pour faire mieux comprendre une théorie qu'il expose et approuve :

"Ainsi mille francs pour celui qui ne possède que deux mille francs, ont la même importance que cinq cent mille francs pour celui qui possède un million."

Si, par ce fier dédain de la fortune lorsque le nécessaire est assuré, on échappe au reproche d'avarice, c'est pour en mériter un plus grave ; Celui qui possédant un million en acquiert un second changera fort peu, pas du tout peut-être, les habitudes de sa vie.

Est-ce là, pour qui n'est pas avare, le seul fruit de la richesse ?

Si l'homme sensé dont parle Buffon n'est pas un cynique égoïste, il pourra, sans thésauriser, faire bon usage des millions qu'on lui suppose ; On pourra les doubler, les décupler et les doubler encore, sans ralentir la progression constante du bien qu'il peut faire ; n'a-t-il pas une famille à enrichir, des misères à soulager de grandes œuvres à créer ou à faire naître ? Il évitera, s'il est sage, de jouer au jeu, même à des conditions équitables ; mais s'il ne porte pas, si riche qu'il soit, cent mille francs par jour à la roulette, la crainte de la perte l'arrêtera beaucoup plus que le mépris du gain. »

Le lecteur de ce chapitre pensera, peut-être, qu'une théorie de la relativité restreinte ou généralisée de l'espérance, qu'elle soit mathématique ou morale, reste à écrire.

Mais il est important de souligner que Joseph Bertrand avait, juste avant les lignes que nous avons reproduites, justifié la théorie classique de l'espérance mathématique en supposant qu'une machine puisse jeter cent mille pièces par seconde, voire un milliard de pièces et enregistrer les résultats. N'importe quel PC peut, aujourd'hui, être programmé pour jouer à ce « Saint-Petersbourg Game » avec un tel nombre de parties et donner, en quelque sorte, raison à Joseph Bertrand. En 1973, une étude de certaines méthodes de « résolution » du paradoxe de Saint-Petersbourg et une simulation sur ordinateur¹ ont été faites.

1. J. Lemaire, *Annales Inst. Henri Poincaré*, vol. IX, n° 2, 1973, p. 205-214.

Une excellente justification de l'espérance mathématique et de la règle des partis de Pascal a été publiée en 1997 par S. Jallais et P. Pradier.¹

Nous approuvons leur analyse et citons quelques phrases de leur conclusion :

« La préhistoire de la théorie de l'utilité espérée, telle qu'on la raconte dans les manuels et parfois dans les articles, semble donc erronée. La "règle des partis" va bien au-delà d'une simple règle de calcul (d'espérance). Les considérations de justice qui guident Pascal (puis Huygens) lui permettent de construire une véritable théorie normative des prix (le "juste" prix étant celui qui assure l'équité des jeux). C'est dire que, contrairement à [Daniel] Bernoulli, Pascal n'a cure de mesurer les "perspectives" ou les "risques" tels qu'ils sont "estimés par chacun", ni de s'interroger sur ce qu'il est judicieux ou non de choisir. De la même façon, la théorie bernoullienne ne peut être tenue ni pour une théorie des prix, puisqu'elle considère des valeurs subjectives (sans prévoir leur coordination), ni pour une théorie conforme à la justice puisqu'elle se fonde sur des comportements observés ou, tout au plus, "raisonnables". Point donc de concurrence entre ces deux théories. »

5.2.3. L'inclinaison des plans des orbites des planètes

En 1735, un mémoire de Daniel Bernoulli parut, portant le titre *Recherches physiques et astronomiques sur le problème proposé pour la seconde fois par l'Académie Royale des sciences de Paris : Quelle est la cause physique de l'inclinaison des plans des orbites des planètes par rapport au plan de l'équateur de la révolution du soleil autour de son axe ; Et d'où vient que les inclinaisons de ces orbites sont différentes entre elles.*²

La question avait été posée en 1732 pour la première fois par l'Académie mais aucun des mémoires candidats ne fut primé. La question fut proposée à nouveau en 1734 et cette fois le prix fut décerné conjointement à Daniel Bernoulli et à son père Jean.

On savait, bien avant Bernoulli que la plupart des planètes du système solaire ont une orbite qui se trouve très proche du plan écliptique qui est le plan géométrique qui contient l'orbite de la Terre autour du Soleil.

Pour résoudre le problème, Daniel décide tout d'abord de démontrer qu'il est impossible que les orbites soient toutes proches du plan écliptique par hasard. Il trouve que l'inclinaison mutuelle la plus grande de toutes les planètes prises deux à deux est celle de Mercure par rapport à l'écliptique, c'est-à-dire $6^{\circ}54'$. Considérant l'hypothèse de positions des orbites sur la sphère céleste toutes dues au hasard, il cherche la probabilité qu'elles tombent toutes dans

1. S. Jallais et P. Pradier, *L'erreur de Daniel Bernoulli ou Pascal incompris*, <http://picha.univ-paris1.fr/1997es.pdf>.

2. *Recueil des pièces qui ont remporté les Prix de l'Académie Royale des sciences*, 3, p. 93-122, 1735. Le même mémoire rédigé en latin occupe les pages 125-144.

la zone maximale de $6^{\circ}54'$. La surface de la sphère étant $4\pi r^2$ et la surface d'une zone sphérique de hauteur h égale à $2\pi rh$, la probabilité d'une orbite située dans la zone est

$$\frac{h}{2r} = \frac{1}{2} \sin 6^{\circ}54' = 0,060 \approx \frac{1}{17}.$$

Bernoulli en déduit que la probabilité que les six planètes aient des inclinaisons toutes incluses dans cette même zone, par hasard, est $\frac{1}{17^5}$, donc très faible. Discutant ensuite de variantes dans ses hypothèses, Bernoulli indique que la probabilité est comprise entre 12^{-5} et 17^{-5} .

On ne sait d'ailleurs pas pourquoi Bernoulli a pris la valeur $1/17$ au lieu de $2/17$, c'est-à-dire a pris en compte la sphère au lieu de l'hémisphère.

Sur le plan physique, Bernoulli a tenté ensuite d'expliquer dans son mémoire que les inclinaisons initiales étaient sans doute quelconques mais qu'une *atmosphère* ressemblant à notre air, tournant autour du soleil a fini par imposer une inclinaison croissante aux orbites des planètes vers l'équateur du soleil.

5.2.4. L'inoculation, moyen de prévenir la petite Vérole

Le mémoire de D. Bernoulli intitulé *Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir* a paru en 1766¹.

Le virus de la de variole appelée jadis petite vérole a causé, au cours des siècles, d'effroyables pandémies responsables de millions de morts. On notera que Louis XV est mort de cette maladie en 1774.

Dès le XI^e siècle, les Chinois pratiquaient une technique appelée *variolisation* ou *inoculation* : On mettait en contact la personne à immuniser avec le contenu de la substance, espérée peu virulente, contenue dans les vésicules d'un malade.

Le résultat restait cependant aléatoire et risqué, le taux de mortalité pouvait atteindre 1 ou 2 %. La technique a été importée en occident au début du XVIII^e siècle, par l'épouse de l'ambassadeur de Grande-Bretagne en Turquie où le docteur E. Timoni (circa 1670-1718) l'avait expérimentée puis exposée en 1713 dans les *Philosophical transactions* de la *Royal Society*. Elle a été introduite en France par le docteur Théodore Tronchin (1709-1781) en 1756.

En fait, la découverte de la vaccination par Edward Jenner (1749-1823) n'interviendra qu'en 1796. Depuis cette découverte, la méthode dite de l'inoculation par la substance contenue dans les vésicules d'un malade a été abandonnée

1. *Histoire de l'académie Royale ses sciences pour l'année 1760*, Mémoires, p. 1-45, 1766.

et le mémoire de Bernoulli ne présente plus qu'un intérêt historique car, comme nous allons le voir, en 1760, Daniel Bernoulli démontra que, malgré les risques, la généralisation de cette pratique permettrait de gagner deux à trois ans d'espérance de vie à la naissance. Elle suscita cependant l'hostilité de nombreux médecins.

Bernoulli pose : x l'âge exprimé en année, ξ le nombre de personnes qui survivent à cet âge à partir d'une population initiale ; s le nombre de ces survivants qui n'ont pas eu la petite vérole.

Il suppose, de plus, que la maladie attaque la fraction $1/n$ des personnes qui n'ont pas eu la maladie auparavant, et que $1/m$ de cette fraction en décède.

Il obtient la formule $s = \frac{m\xi}{(m-1)e^{\frac{x}{n}} + 1}$.

Bernoulli applique cette formule, en supposant $m = n = 8$, à la table de mortalité de Breslau, publiée par Halley¹. Il obtient donc une table où apparaît le nombre s de survivants n'ayant pas eu la petite vérole en fonction de l'âge.

Il suppose ensuite que la variole a disparu, et calcule, dans cette hypothèse, le nombre de survivants en fonction de l'âge. En supposant, de plus, que l'inoculation est efficace à 100 % et qu'elle n'entraîne que très peu d'accidents mortels, il peut calculer, à l'aide des résultats précédents, que cette pratique permet de gagner deux à trois ans d'espérance de vie à la naissance.

D'Alembert a critiqué cette approche du problème par Bernoulli, et on trouvera un commentaire à ce propos dans le chapitre consacré à d'Alembert.

5.2.5. La théorie des erreurs et la méthode du maximum de vraisemblance de Daniel Bernoulli

Le mémoire consacré par Daniel Bernoulli aux erreurs d'observations est intitulé *Dijudicatio maxime probabilis plurium observationum discrepantium atque verisimillima inductio inde formanda* et a été publié² en 1778.

Après Simpson et Lagrange, dont Daniel Bernoulli semble avoir ignoré les travaux, on trouve dans ce mémoire une méthode de *choix le plus probable entre des observations divergentes et de formation de l'induction la plus vraisemblable*. Bernoulli insiste sur le fait, qu'ignorant ce qu'il est advenu durant les observations, on est forcé de chercher non ce qui est le plus vrai, mais le plus vraisemblable, non la certitude mais la probabilité la plus élevée (« non

1. Halley E., *An estimate of the degrees of mortality of mankind ; drawn from curious tables of the births and funerals at the city of Breslaw ; with an attempt to ascertain the prices of annuities upon lives*, Philosophical Transactions, 17, n° 196, p. 596-610, 654-656, 1693.
2. *Acta Acad...*, Petrop. pro 1777, pars prior, p. 3-23, 1778.

verissimum sed verisimillimum, non certum sed probabilissimum »). Il admet cependant que son raisonnement est plus *métaphysique* que *mathématique*.

L'auteur indique que la méthode habituelle qui consiste à prendre la moyenne entre des observations discordantes suppose que toutes les observations ont le même poids. Il considère cette hypothèse fautive et pense que les petites erreurs sont plus probables que les erreurs importantes.

Soit e une erreur, Bernoulli propose de mesurer la probabilité de l'erreur par $\sqrt{(r^2 - e^2)}$, où r est une constante égale à la demi-amplitude de la densité qu'il considère comme semi-circulaire. Il écrit ensuite que le meilleur résultat pour un ensemble de n observations sera celui pour lequel le produit des probabilités des erreurs sera maximum. Si les observations ont été $a_1, a_2, \dots, a_n$, et si la vraie valeur est appelée x , on a

$$\sqrt{(r^2 - (x - a_1)^2)} \cdot \sqrt{(r^2 - (x - a_2)^2)} \dots \sqrt{(r^2 - (x - a_n)^2)} = \prod_{i=1}^n \sqrt{r^2 - (x - a_i)^2} = L$$

qui doit être maximum. On pourrait différentier ce produit L par rapport à x et l'égaliser à zéro. Mais si le produit est maximum, son carré l'est aussi. Il est donc préférable de différentier L^2 et d'écrire $\frac{dL^2}{dx} = 0$. On en tire $\bar{x}$. Pour $n = 1$, Bernoulli trouve $\bar{x} = a_1$, pour $n = 2$ il trouve la moyenne arithmétique $\bar{x} = \frac{a_1 + a_2}{2}$. Pour $n = 3$, l'équation à résoudre est du cinquième degré.

Cette technique est celle du *maximum de vraisemblance*. Cependant Bernoulli, au lieu de prendre la loi normale pour la distribution, a pris une loi semi-circulaire.

Euler a ajouté quelques commentaires sur ce papier de Bernoulli à la suite dudit papier dans les Mémoires de l'Académie de Saint-Petersbourg. On trouvera dans Kendall¹ un commentaire et une traduction des mémoires de Bernoulli et d'Euler. Diverses autres analyses ont été publiées.²

1. Kendall M. G., *Daniel Bernoulli on maximum likelihood*, Biometrika, 48 (1961), 1-18.
2. Cf. Todhunter I., *A History of the mathematical theory of probability*, Cambridge, 1865, p. 236-238 ; Sheynin O. B., AHES, 9, p. 45-56 ; Pearson K., *The history of statistics*, 1978, p. 267-269 ; Hald A., *A history of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 83-87.

5.3. Nicholas I Bernoulli

5.3.1. La thèse de Nicholas I Bernoulli : De usu arte conjectandi in jure¹

Nicholas I avait obtenu sa maîtrise en mathématiques en 1704 en présentant la cinquième et dernière partie du travail de son oncle Jacob sur les séries². La troisième partie de ce même ouvrage avait par ailleurs été présentée en 1696 par Jacob Hermann qui travailla avec Nicholas à la mise en ordre des papiers de Jacob après le décès de ce dernier survenu en 1705.

En 1709, à l'âge de 21 ans, Nicholas obtint, à Bâle, son doctorat en jurisprudence avec une thèse intitulée *Dissertatio inauguralis mathematico-juridica de Usu artis conjectandi in jure*.

Ce travail est en fait une continuation des considérations incluses par son oncle Jacob dans son dernier chapitre de l'*Ars conjectandi* resté inachevé. Nicholas l'indique d'ailleurs, lui-même, dans un résumé de sa thèse qu'il a publié en 1711 dans la revue *Acta Eruditorum*³.

D'Alembert, dans le premier tome de l'Encyclopédie⁴, publié en 1751, cite la thèse de Nicholas en ces termes :

« Lorsque M. Nicolas Bernoulli, neveu des célèbres Jacques et Jean Bernoulli, soutint à Bâle en 1709 sa thèse de docteur en Droit ; comme il était grand Géomètre, aussi bien que jurisconsulte, il ne pût s'empêcher de choisir une matière qui admit de la Géométrie. Il prit donc pour sujet de thèse, *De usu artis conjectandi in jure*, c'est-à-dire de l'application du calcul des probabilités aux matières de Jurisprudence ; et le troisième chapitre de cette thèse traite du temps où un absent doit être réputé pour mort. Selon lui, il doit être censé tel, lorsqu'il y a deux fois plus à parier qu'il est mort que vivant [...] ».

La thèse comporte neuf chapitres traitant, après quelques rappels sur le calcul des probabilités, de la durée de la vie humaine, du problème de l'absent évoqué plus haut, des rentes et assurances, des jeux de hasard et de la fiabilité des témoignages. Un chapitre est consacré à la loi romaine dite du *quart Falcidien*, édictée en l'année romaine 714, qui permettait à un testateur de disposer des trois quarts de ses biens, le reste étant cependant réservé, de droit, à son héritier.

1. Nicolaus Bernoulli, *Dissertatio inauguralis mathematico-juridica de Usu artis conjectandi in jure*, Basileae, 1709.
2. Les cinq maîtrises relatives à l'ouvrage de Jacob Bernoulli sur les séries, ont été présentées sous la présidence de ce dernier par J.J. Fritz (1689) ; H. Beck (1692) ; J. Hermann (1696) ; N. Harscher (1698) et Nicholas Bernoulli en 1704.
3. *Acta Eruditorum*, Supplementa, t. IV, sect. IV, p. 159-170, Leipzig, 1711.
4. Page 40.

DISSERTATIO INAUGURALIS
MATHEMATICO-JURIDICA

DE

USU ARTIS
CONJECTANDI
IN JURE,

Quam

DIVINA JUVANTE GRATIA

Auctoritate & Jussu

*Magnifici & Amplissimi Jctorum Ordinis
in Academia Patria*

pro

GRADU DOCTORATUS

In Utroque Jure legitime consequendo

Ad Diem 14 Junii A. C. M DCC IX.

L. H. Q. S.

Publice defendet

M. NICOLAUS BERNOULLI,
Basiliensis.



BASILEÆ,

Typis JOHANNIS CONRADI à MECHEL,

Les chapitres 2 à 6 discutent des probabilités de survie pour une ou deux personnes, du calcul des annuités basé sur la table de Graunt et des assurances maritimes et sur la vie humaine.

Au chapitre 9 de sa thèse Nicholas a un raisonnement étrange lorsqu'il juge la fiabilité des témoignages. Son oncle, dans son chapitre III de la IV^e partie, intitulé *Des genres divers d'arguments et l'estimation de leur poids pour supputer la probabilité des choses*, distinguait trois types d'arguments : ceux qui existent nécessairement et indiquent avec contingence, ceux qui existent avec contingence mais indiquent nécessairement et ceux qui existent et indiquent à la fois avec contingence c'est-à-dire non nécessairement. Ses calculs du poids des arguments observaient de plus une distinction entre des arguments purs et des arguments mixtes, ces derniers pouvant prouver, selon les cas, une chose ou son contraire. Nicholas Bernoulli semble ne pas avoir pris en compte ce chapitre de *l'Ars conjectandi*. Prenant pour exemple le cas d'un crime pour lequel il existe dix indices, chacun de ces indices indiquant que la probabilité que l'accusé soit innocent est deux fois plus grande que celle qu'il soit coupable, il déclare que l'innocence a pour probabilité $\left(\frac{2}{3}\right)^{10} = \frac{1\,024}{59\,049}$ et qu'en conséquence il est moralement certain que l'accusé est coupable.

N. Meusnier¹ fait remarquer qu'avec le même raisonnement, Nicholas Bernoulli aurait pu affirmer que la probabilité de la culpabilité était de $\left(\frac{1}{3}\right)^{10} = \frac{1}{59\,049}$, et qu'en conséquence il est moralement certain que l'accusé est innocent.

5.3.2. La contribution de Nicholas Bernoulli² au théorème de Jakob Bernoulli

Cette contribution est incluse dans une lettre datée de 23 janvier 1713 et adressée à Montmort qui l'a publiée dans la deuxième édition de son livre en 1713, pages 388-394. Le point de départ de cette lettre est de discuter du rapport des naissances des garçons et des filles suite au mémoire d'Arbuthnot paru en 1712 (cf. notre chapitre 18). Le but de Nicholas est de montrer que la régularité d'un phénomène peut être attribuée au hasard et non à une intention divine comme le prétendait Arbuthnot. Il écrit :

« Je vous envoie le Catalogue des Enfants de chaque sexe nés à Londres depuis 1629 jusqu'à 1710, avec mes démonstrations de ce que je vous ai écrit touchant l'argument par lequel on veut prouver que c'est un miracle que les nombres des enfants de chaque sexe nés à Londres ne se sont pas plus éloignés les uns des autres pendant 82 ans de

1. N. Meusnier, *Nicholas, neveu exemplaire*, Electronic journal for probability and statistics, vol. 2, n° 1, juin 2006, www.jehps.net.
2. Cf. Hald A., *Nicholas Bernoulli's Theorem*, International Statistical Review / Revue Internationale de Statistique, vol. 52, n° 1 (Apr., 1984), p. 93-99.

suite, et que par le hazard il serait impossible que pendant un si long temps ils fussent toujours renfermés entre des limites aussi petites que celles qu'on a observées dans le Catalogue de 82 ans. Je prétends qu'il n'y a aucun sujet de s'étonner, et qu'il y a une grande probabilité pour que les nombres des mâles et des femelles tombent entre des limites encore plus petites que celles qu'on a observées. »

Dans le cours de sa démonstration Nicholas affine le théorème de convergence en probabilité de Jacob Bernoulli.

Ce théorème peut, avec des notations modernes, être abordé comme suit :

Soient n épreuves indépendantes, chacune ayant une probabilité p pour la réalisation d'un certain événement et appelons s_n le nombre de succès distribué selon la formule du binôme, on cherche les propriétés de la fréquence relative de succès $h_n = \frac{s_n}{n}$.

Les notations de Bernoulli sont les suivantes : Un essai a un nombre total de réalisations également possibles $t = r + s$ dont r sont favorables, si bien que la probabilité est $p = \frac{r}{r+s}$. La formulation par Jakob Bernoulli de son théorème est la suivante :

« Admettons que le nombre de cas fertiles est exactement ou approximativement dans le rapport r/s , au nombre de cas stériles, si bien que le rapport du nombre de cas fertiles à la totalité des cas est $r/(r+s)$ ou r/t (avec $t = r + s$) qui est compris entre les limites $(r+1)/t$ et $(r-1)/t$. Il est possible d'effectuer un nombre d'essais assez élevé pour qu'il soit plus probable, dans le rapport de n'importe quel nombre donné (disons c), que le nombre de cas fertiles soit compris entre les limites $(r+1)/t$ et $(r-1)/t$, que le cas contraire. »

En termes modernes nous dirions : Si une épreuve de Bernoulli, ayant une probabilité r/t , est répétée un grand nombre de fois, on peut choisir un nombre c aussi grand que l'on veut pour que la probabilité que la fréquence de réalisation des cas favorables soit comprise entre les limites $(r+1)/t$ et $(r-1)/t$, soit c fois plus élevée que la probabilité du cas inverse, c'est-à-dire plus grande que $c/(c+1)$.

Si on appelle n le nombre d'épreuves indépendantes, chacune ayant une probabilité $p = r/t$ pour la réalisation d'un certain événement et s_n le nombre de succès distribué selon la formule du binôme, on a, la fréquence relative de succès étant notée $h_n = \frac{s_n}{n}$:

$$\frac{r-1}{t} \leq h_n \leq \frac{r+1}{t}, \text{ ou bien } |h_n - p| \leq \frac{1}{t},$$

Avec les notations de Bernoulli $p = \frac{r}{r+s}$ et $t = r + s$, pour tout nombre réel c positif, le théorème de Bernoulli s'écrit

$\Pr\left[\left|b_n - p\right| \leq \frac{1}{t}\right] > \frac{c}{c+1}$, pour n assez grand, c'est-à-dire que la fréquence relative $b_n = \frac{s_n}{n}$ converge en probabilité vers p lorsque n tend vers l'infini.

Posons $n = kt$, la démonstration de Bernoulli établit que $\Pr\left[\left|b_n - p\right| \leq \frac{1}{t}\right] > \frac{c}{c+1}$ si $k \geq k(r, s, c)$ où $k(r, s, c)$ est le plus petit entier positif satisfaisant l'inégalité

$$k(r, s, t) \geq \frac{m(r + s + 1) - s}{r + 1},$$

et m est le plus petit entier positif satisfaisant l'inégalité

$$m \geq \frac{\ln[c(s-1)]}{\ln[(r+1)/r]}.$$

Le neveu de Jakob, Nicholas a, dans la lettre datée du 23 janvier 1713 adressée à Montmort, amélioré l'approximation de la distribution binomiale en modifiant les inégalités données par son oncle dans la démonstration de son théorème.

Il a établi que le théorème de Jacob Bernoulli est valable pour

$$k(r, s, c) \geq \frac{\ln(c+1)}{\ln[(r+1)/r]} \frac{r+s+1}{r+1} - \frac{s}{r+1},$$

ce qui améliore l'approximation du nombre d'essais minimal pour obtenir une fourchette de probabilités donnée.

L'exemple donné par Jacob était le suivant :

Une épreuve ayant une probabilité de 0,6 ; une valeur du nombre c (définie plus haut) de 1 000 et une fourchette pour b autour de 0,6 de 0,02. Il obtient :

$$P(0,58 \leq b \leq 0,62) > \frac{1000}{1001} \text{ dès que } 25\,550 \leq n.$$

Avec l'approximation de Nicholas I, on obtient la même fourchette pour $n \geq 17\,350$.

Nicholas I avait indiqué en 1711 dans les *Acta Eruditorum*¹ qu'il espérait que le livre de son oncle Jacob serait édité prochainement. Dans cette lettre à Montmort du 23 janvier 1713, il annonce que l'impression du livre de son oncle Jacob est en cours :

1. Nicolai Bernoulli, [...] *Specimina Artis conjectandi, ad questiones Juris applicatae, Acta Eruditorum*, Supplementa, t. IV, sect. IV, 1711, p. 159-170.

« Je me souviens que feu mon Oncle a démontré [...] dans son *Traité De Arte conjectandi* qui s'imprime à présent à Bâle, [...] que si l'on veut découvrir par les expériences souvent répétées le nombre des cas par lesquels un certain événement peut arriver ou non, on peut augmenter les observations en telle manière qu'enfin la probabilité que nous ayons découvert le vrai rapport qu'il y a entre les nombres des cas, soit plus grande qu'une probabilité donnée. »

Nicholas I Bernoulli a effectué de nombreux autres travaux. Certains d'entre eux sont cités dans notre chapitre 4.

A. Hald¹ a listé et commenté d'une façon exhaustive toutes les contributions de Nicholas Bernoulli à la théorie des probabilités.

5.4. Biographie de Jacob Bernoulli (1657-1705)

Jacob (ou Jacques I) Bernoulli fait partie de la première génération de la dynastie des Bernoulli². Il naît à Bâle le 27 décembre 1657. Son père Nicolas (senior) Bernoulli (1623-1708) est un commerçant et magistrat de la ville de Bâle et sa mère, Margaretha Schönauer, est issue d'une importante famille de banquiers. Destiné par son père à la philosophie et à théologie, Jacques obtient respectivement des diplômes dans ces deux disciplines en 1671 et 1676 à l'Université de Bâle. Mais, en parallèle, il aborde, par goût, l'apprentissage de l'astronomie et des mathématiques. En 1676, il part pour Genève où il commence à enseigner puis voyage durant deux ans en France où il prend connaissance des travaux des successeurs de Descartes. En 1681, il se rend aux Pays-Bas où il rencontre divers mathématiciens dont Hudde (1628-1704), puis en Grande Bretagne où il rencontre Vossius, Hooke (1635-1703) et Boyle (1627-1691). De ces divers voyages, va naître une correspondance suivie de Jacques avec de nombreux scientifiques européens. En 1682, il publie un mémoire sur les comètes³ dont la théorie est inexacte, et l'année suivante un autre sur la gravité (en fait une théorie de l'éther)⁴. De retour en Suisse, il enseigne à partir de 1683 à l'Université de Bâle la mécanique des solides et des liquides et il commence à publier ses recherches en mathématiques et en physique dans le journal *Acta Eruditorum*, qui est créé en 1682 à Leipzig et où Leibniz exerce une très grande influence, et dans le *Journal des Savans* qui paraît à Paris depuis 1665. Il initie également aux mathématiques son jeune frère Jean⁵, né le 27 juillet 1667, dixième enfant de la famille. De son côté, il étudie les travaux de Descartes et les commentaires y afférents de Schooten (1615-1660), ainsi que les traités publiés par Wallis et Barrow (1630-1677)

1. Hald A., *History of Probability and Statistics and their applications before 1750*, Wiley, 1990, chap. 21.
2. Quelques détails sur cette dynastie de mathématiciens sont donnés au § 5.7.
3. *Conamen novi Systematis Cometarum, pro motu eorum sub calculum revocando et apparitionibus praedicendis*.
4. *Dissertatio de Gravitate Aetheris*.
5. Jean Bernoulli est également sous le nom de Jean I^{er} ou Johann.

par lesquels il aborde le calcul infinitésimal qu'il approfondit dès la parution du papier de Leibniz *Nova Methodus pro Maximis et Minimis* dans les *Acta Eruditorum* en 1684, et des travaux de von Tschirnhaus (1651-1708).

Cette même année, Jacques Bernoulli épouse Juditha Stupanus, fille d'un riche pharmacien, dont il aura deux enfants, qui, à l'inverse de nombreux membres de la famille Bernoulli (en trois générations, il y eut huit mathématiciens), ne deviendront pas des scientifiques. Son fils Nicolas (ou Nikolaus) deviendra artiste peintre.

En 1685, Bernoulli aborde les relations de la logique et de l'analyse dans un mémoire intitulé *Parallelismus ratiocinii logici et algebraici*. Il s'agit du débat de thèse qu'il préside, le 9 septembre 1685, pour l'obtention par son frère Jean du diplôme de Maître ès arts.

Jacques est nommé professeur à l'Université de Bâle en 1687. Il occupera ce poste jusqu'à sa mort. Son frère Jean s'intéresse aux mêmes problèmes que lui et peu à peu ils deviennent rivaux¹. À la mort de Jacques en 1705, Jean ne fera, d'ailleurs, aucun effort pour publier l'*Ars conjectandi* de son frère aîné.

Jacques est au sommet de son art entre 1689 et 1692. Il publie en tant qu'auteur et président, entre 1689 et 1704, une série de thèses (pour le diplôme de maîtrise) sur les séries infinies². La cinquième et dernière de la série permet d'ailleurs à son neveu Nicolas I (Nicolaus I) (1687-1759), fils de Nicolas (1662-1716), d'obtenir son diplôme de Maître ès Arts en 1704. C'est ce même neveu qui prendra la peine de publier l'*Ars Conjectandi* de son oncle. Il obtiendra d'ailleurs un doctorat en 1709, en présentant un mémoire sur l'application des probabilités à certaines questions juridiques. Cette suite de dissertations rédigée sur quinze ans comporte soixante propositions numérotées où Bernoulli veut présenter tout ce qui est connu sur les séries sans grand souci d'originalité. On y trouve ainsi des emprunts à Leibniz, Barrow (1630-1677), James Gregory (1638-1675), Wallis, Newton et même Fatio de Duillier (1664-1753). Le théorème du binôme de Newton est présenté comme un théorème déjà connu. L'inégalité, dite de Bernoulli : $(1+x)^n > 1+nx$ avec $x > -1, \neq 0$ et $n > 1$, entier, y est démontrée.

En mai 1690, dans les *Acta Eruditorum* (p. 217), Bernoulli montre sa maîtrise en analysant la solution donnée par Huygens en 1687 et Leibniz en 1689 au problème de la descente uniforme d'un corps dans un champ de gravitation (tautochrone)³. Le problème de descente uniforme a été posé en 1687 ;

1. La dynastie des Bernoulli a connu de nombreux épisodes de rivalités, de jalousies et d'amertumes, particulièrement Jean qui est également jaloux de son fils Daniel.
2. *Positiones arithmeticae de seriebus infinitis, earumque summa infinita quas auctore praeside Jacobo Bernoulli* (elles seront republiées en 1713, à la suite de l'*Ars conjectandi*).
3. Son mémoire a pour titre *De inventione lineae descensus a corpore gravi percurrentae uniformiter, sic ut temporibus aequales altitudines emetiatur*.

Jacques Bernoulli y répond en posant celui de la chaînette. En 1691, il traite le problème de la spirale parabolique et, un an plus tard, celui de la spirale logarithmique¹ (*spira mirabilis*, la spirale miraculeuse).

Bernoulli a fait graver une spirale logarithmique sur sa tombe dans la cathédrale de Bâle, avec l'épigraphe : *eadem mutata resurgo*. Cependant, le graveur a tracé, par erreur, une spirale d'Archimède.

En 1692, il découvre le *Theorema Aureus*, sa formule pour les rayons de courbure qu'il publie dans les *Acta Eruditorum* de 1694 (p. 261), sous le titre *Curvatura laminae elasticae*. C'est la première application non triviale des dérivées d'ordre supérieur à un. En 1691 et 1692, Bernoulli donne des cours particuliers au Marquis de l'Hôpital, qui se fera aider également par Jean, et publie son *Analyse des infiniment petits pour l'intelligence des lignes courbes* en 1696. En 1694, dans les *Acta Eruditorum*, Jacques définit les coordonnées polaires en introduisant sa lemniscate².

En 1695, il commente la géométrie de Descartes³. Son frère Jean est, de son côté, nommé professeur à l'université de Groningue en 1695. L'année suivante, en juin, Jean propose le problème de la brachistochrone⁴. Le problème est résolu par Leibniz, les deux frères Bernoulli et Newton. Lorsque Newton reçoit l'énoncé du problème, il le résout en quelques heures et fait en sorte que sa solution paraisse sous forme anonyme dans les *Philosophical Transactions of the Royal Society* (1697). Jean, en la lisant déclare « On reconnaît la patte du Lion ». La solution de Leibniz est, quant à elle, basée sur une technique d'approximation des courbes par des lignes polygonales. La solution de Jacques, fondée sur le principe de Huygens (principe du temps minimal), est plus générale mais aussi plus laborieuse que celle de son frère Jean. L'Hôpital propose lui aussi une solution mais il a été aidé par Jean. Seul Jacques note que le problème ouvre un nouveau champ de recherches : le calcul des variations, qui sera développé plus tard, bien que Newton ait déjà énoncé en 1687 son problème aérodynamique⁵.

1. La spirale logarithmique peut être définie comme la courbe dont l'angle tangentiel polaire reste constant (non droit) ou comme la courbe dont la courbure est inversement proportionnelle à l'abscisse curviligne.
2. La lemniscate est une courbe ayant pour équation en coordonnées polaires $r^2 = a^2 \cos(2\theta)$ [ou $(x^2 + y^2)^2 = a^2(x^2 - y^2)$ en coordonnées cartésiennes]. C'est la podaire (lieu des projections d'un point donné sur les tangentes à une courbe donnée) de l'hyperbole. Elle est le lieu des points M du plan dont le produit des distances à deux points fixes F et F', distants de 2a, est constant et égal à a^2 ($MF \times MF' = a^2$).
3. Son mémoire s'intitule *Notae et animadversiones tumultuariæ in geometriam Cartesii*.
4. Étant donné deux points A et B dans le plan vertical, déterminer le chemin reliant A et B suivant lequel un corps entraîné par son poids effectuera le trajet entre A et B en un temps minimal.
5. Dans le paragraphe des *Principia*, livre II, consacré au mouvement des corps dans les milieux résistants.

De son côté, Jean Bernoulli se consacre à la diffusion du calcul infinitésimal en Europe. Il devient le correspondant privilégié de Leibniz et de nombreux autres scientifiques européens. Très orgueilleux, il ne supporte pas l'idée d'être le second après son frère auquel il lance toutes sortes de défis mathématiques. Les deux frères échangent des propos peu aimables, voire agressifs, par le biais des journaux scientifiques. Jean aura même, plus tard, des problèmes de propriété intellectuelle avec son fils Daniel dont il copie certains travaux. Lorsque Daniel remporte un prix de l'Académie des sciences de Paris qu'il convoitait lui-même, Jean n'hésite pas à se brouiller avec lui. La responsabilité des mauvaises relations entre Jacques et Jean est sans doute une responsabilité partagée, Jacques ne se privant pas, par ailleurs, de critiquer ses collègues de l'Université. Quelquefois, chacun des deux frères sous-estimait ou ne comprenait pas le travail de l'autre.

Dans les divers domaines des mathématiques, sauf en théorie des nombres, Jacques est un pionnier. En 1701, il rédige et fait défendre par son étudiant J.-J. Episcopius une thèse sur le problème des isopérimètres¹ qui a été posé publiquement en 1697, sous le titre *Analysin magni problematis isoperimetrici*. En 1703, il publie un travail sur les centres d'oscillation, motivé par le problème du pendule, sous le titre *Démonstration générale du centre de balancement ou d'oscillation tirée de la nature du levier*². L'équilibre des leviers correspond, en fait, à l'équilibre des moments.

Jacob Bernoulli est mort le 16 août 1705 à Bâle.

5.5. Principales publications de Jacob Bernoulli

Jacobi Bernoulli, Basileensis Opera, Genevae sumptibus haeredum Cramer & fratrum Philibert, 1744. [Cet ouvrage, en deux volumes, réunit toutes les Œuvres du mathématicien].

5.6. Références relatives à la biographie de Jacob Bernoulli

- J.E. Hofmann, *Dictionary of Scientific Biography*, Charles Scribner's sons, New York, vol. 2, p. 46-50, 1970.
- E.T. Bell, *Men of mathematics*, Simon et Schuster, New York, 1986.
- H. Bernhard, *The Bernoulli family*, in H Wussing and W Arnold, *Biographien bedeutender Mathematiker*, Berlin, 1983.
- J.O. Fleckenstein, *Johann und Jacob Bernoulli*, Basel, 1949.

1. Recherche parmi toutes les courbes de longueur donnée de celle qui limite une aire maximale.
2. Le mémoire paraît dans l'*Histoire de l'Académie des Sciences de Paris* (1703, p. 78).

5. 7. Biographie de Daniel Bernoulli (1700-1782)

La famille Bernoulli, de religion calviniste, est originaire d'Anvers mais elle s'est réfugiée à Bâle, en Suisse, lorsque le Roi d'Espagne a envoyé, en 1567, aux Pays-Bas, une armée pour y rétablir son autorité et accessoirement persécuter les protestants. Les Bernoulli forment une dynastie de mathématiciens unique dans l'histoire, puisqu'elle en comptera huit répartis sur trois générations successives. Jusqu'à Nicolas senior (1623-1708) inclus, les Bernoulli étaient une famille de riches marchands. Deux des trois fils de Nicolas senior sont devenus mathématiciens : Jacob (1654-1705)¹ (alias Jacques I) ; et Jean Bernoulli² (1667-1748) (alias Jean I). Le troisième Nicolas (1662-1716) est devenu peintre. Daniel Bernoulli est le fils de Jean I Bernoulli, il appartient ainsi à la deuxième génération de la dynastie. Daniel naît le 8 février 1700 à Groningue, en Hollande, où son père est professeur de mathématiques depuis 1695. Daniel a deux frères : Nicolas II³ (1695-1726) son aîné et Jean II (1710-1790), son puîné.

En 1705, Jean quitte Groningue avec sa famille pour s'installer à Bâle où il succède à son frère Jacques, qui vient de mourir, à la chaire de mathématiques de l'université de la ville. Curieusement, mais peut-être pour des raisons héréditaires, Jean pousse son fils Daniel à faire des études de commerce. Ce dernier est d'abord inscrit, dès l'âge de treize ans, à l'université de Bâle pour y faire des études de philosophie et de logique. En 1715, à quinze ans, il obtient sa licence et l'année suivante sa maîtrise. Mais Daniel, qui apprend en parallèle les mathématiques depuis l'âge de onze ans auprès de son frère Nicolas (II), souhaite poursuivre dans cette voie. Toutefois son père, après avoir tenté vainement de lui faire faire un apprentissage de commerce, le dissuade de faire des mathématiques (parce que cela ne rapporte pas d'argent) et le persuade de faire des études de médecine. Daniel y consent et poursuit ses études successivement à Bâle (1717), Heidelberg (1718), Strasbourg (1719) et de nouveau Bâle (1720) où il prépare son doctorat. Cette fois, c'est auprès de son père qu'il va compléter ses connaissances en mathématiques et c'est également auprès de lui qu'il étudie les théories de son père sur l'énergie cinétique. Il applique ses compétences inhabituelles pour un futur médecin à sa thèse qui porte sur la mécanique de la respiration et obtient son doctorat en médecine.

1. Voir biographie de Jacques I Bernoulli au § 5.4.
2. Jean Bernoulli, à la suite de son frère Jacques, enseigne et diffuse le calcul différentiel et intégral, notamment en France (1691). Il décompose, simultanément à Leibniz avec lequel il est en relation épistolaire, les fractions rationnelles en éléments simples. En mécanique, on lui doit le principe des déplacements virtuels, qui est à l'origine de la mécanique analytique de J.-L. Lagrange, et l'introduction du symbole g pour désigner l'accélération de la pesanteur.
3. Nicolas I (1687-1759) est le fils de Nicolas (le peintre) et le neveu de Jacob (Jacques I).

Daniel souhaite faire une carrière académique, il postule à la chaire d'anatomie et de botanique, mais n'est pas retenu. Il tente alors sa chance pour une chaire de logique qui vient de se libérer, mais de nouveau cela se solde par un échec. Il décide alors d'aller à Venise pratiquer la médecine et perfectionner ses connaissances dans ce domaine à l'université de Padoue voisine. Malheureusement Daniel Bernoulli tombe malade et ne peut pas exécuter son plan, par contre il réalise ses premiers travaux en mathématiques et publie, en 1724, un livre intitulé *Exercices mathématiques*. Notons que la première partie de ce livre (qui en contient quatre) décrit le jeu du faro¹, ce qui montre que dès cette époque Daniel Bernoulli s'intéressait aux probabilités. En 1725, il revient à Bâle où il apprend qu'il a remporté un prix de l'Académie des sciences de Paris pour la construction d'un sablier horaire naval insensible aux mouvements des bateaux. De plus son livre de mathématiques l'a fait connaître des milieux académiques et on lui offre une chaire de mathématiques à l'université de Saint-Petersbourg. Une offre similaire est également faite à son frère Nicolas II, de sorte qu'ils se rendent ensemble à la fin de 1725 dans cette ville. Malheureusement Nicolas meurt de fièvre huit mois plus tard. Daniel, qui reste seul dans cette ville au climat hostile est désespéré et en fait part à son père Jean. Ce dernier lui propose de lui envoyer son meilleur étudiant, Léonard Euler² pour travailler avec lui. C'est ainsi qu'Euler arrive à Saint-Petersbourg en mai 1727 et occupe un poste à la section médicale de l'Académie. Il ne retournera jamais en Suisse. Daniel Bernoulli restera à Saint-Petersbourg jusqu'en 1733. Finalement ce séjour correspondra à l'une des périodes les plus fructueuses de sa vie : théorie des vibrations harmoniques émises par les instruments à cordes, hydrodynamique³ et probabilités. Ses travaux sur les probabilités sont liés à l'économie politique. Daniel Bernoulli fait l'hypothèse que la valeur morale de l'augmentation de la richesse d'une personne est inversement proportionnelle au montant de cette richesse. Il assigne alors des probabilités aux divers moyens qu'une personne a de gagner de l'argent et déduit une prévision de l'augmentation de cette espérance morale. Il applique ensuite certaines de ses conclusions à des problèmes d'assurance. Malgré le plaisir qu'il a à travailler avec Euler, Daniel Bernoulli ne s'adapte pas à la vie de Saint-Petersbourg. En 1731 il postule à un poste de professeur de botanique à Bâle, mais, de nouveau, il n'est pas retenu.

1. Jeu de cartes qui est à l'origine du poker.
2. Voir par exemple : *Trente livres de mathématiques qui ont changé le monde*, J.-J. Samuëli, J.-C. Boudenot, Éd. Ellipses, 2006 (chapitre 15 « Euler et l'analyse »).
3. C'est à Saint-Petersbourg que Daniel Bernoulli écrit son fameux ouvrage *Hydrodynamica*. Il laisse une première version de son livre à l'imprimeur, mais son livre ne sera finalement publié qu'en 1738. Les modifications qu'il apporte à son ouvrage entre 1734 et 1738 portent d'ailleurs plus sur la forme que sur le fond. L'« *Hydrodynamica* » englobe l'hydrostatique et l'hydraulique. « Ma théorie est nouvelle, précise Daniel Bernoulli, parce qu'elle considère à la fois la pression et le mouvement des fluides ». Il s'agit d'un traité remarquable qui a à peine vieilli aujourd'hui. Daniel Bernoulli y pose les premiers principes de la théorie cinétique des gaz ($PV = 1/3 Nm \langle v^2 \rangle$) et introduit la fameuse « équation de Bernoulli » de l'hydrodynamique des fluides parfaits ($v^2/2 + P/\rho + gz = \text{cst}$).

La dynastie des Bernoulli a connu de nombreux épisodes de rivalités, de jalousies et d'amertumes. Deux anecdotes illustrent bien ce point. Daniel Bernoulli soumet en 1734 un mémoire pour le grand prix de l'Académie des sciences de Paris, mais son père Jean présente également un mémoire pour l'obtention du même prix ! Finalement le Grand Prix est attribué conjointement au père et au fils. Le père est furieux à l'idée que son fils soit placé au même rang que lui et le chasse de sa maison de Bâle. Une situation encore moins digne se produit quelques années plus tard. Comme nous l'avons déjà signalé, Daniel Bernoulli publie son grand ouvrage *Hydrodynamica* en 1738. L'année suivante son père Jean publie un autre ouvrage, *Hydraulica*, largement inspiré par les travaux de son fils. Mais Jean essaye de faire croire que c'est l'ouvrage de son fils qui est fondé sur le sien en antichronisant la date de parution d'*Hydraulica* à 1732 au lieu de la vraie date qui est 1739. Ainsi non seulement le père a essayé de s'attribuer les mérites de son fils, mais encore il discrédite les travaux de ce dernier en laissant croire que ceux-ci ne sont pas originaux ! Heureusement Daniel n'a pas le même caractère que son père et malgré l'épisode malheureux de 1734 il note sur le frontispice de son livre *Hydrodynamica* le nom de l'auteur : « Daniel, fils de Jean ».

Daniel Bernoulli quitte Saint-Petersbourg pour Bâle en 1733 et y enseigne la botanique, ce qui ne le passionne guère. En 1743 on lui offre un autre poste qui est un peu plus à son goût, l'enseignement de la physiologie. Ce n'est finalement qu'en 1750, à l'âge de 50 ans, qu'il devient titulaire de la chaire de physique de l'université de Bâle, chaire qu'il occupera vingt-six ans, jusqu'en 1776. C'est un excellent professeur, s'attachant à faire des expériences pour illustrer ses cours, chose rare à cette époque. Ses travaux en physique mathématique, dont il est l'un des pionniers, sont récompensés par dix Grands Prix de l'Académie des sciences de Paris¹.

Daniel Bernoulli devient membre associé étranger de l'Académie des sciences de Paris en 1748. Beaucoup d'autres sociétés savantes l'ont accueilli en leur sein ; Saint-Petersbourg, Berlin, Londres, Bologne, Zürich, Berne, Turin, Mannheim. Toutefois Daniel Bernoulli raconte que le plus grand honneur qu'on lui ait fait l'a été de façon tout à fait fortuite. Il voyageait un jour avec un homme cultivé qui ne le connaissait pas et celui-ci lui demanda son nom, notre homme répond naturellement « Je suis Daniel Bernoulli ». Croyant à une plaisanterie son interlocuteur lui répondit « Et moi je suis Isaac Newton ».

1. Il obtient le Grand Prix en 1740 (conjointement à Euler) pour un travail sur la théorie des marées de Newton ; en 1743 et 1747 ce sont ses travaux sur le magnétisme qui sont récompensés ; en 1747 un autre Grand Prix lui est également attribué pour sa méthode proposée pour la mesure du temps en mer (indispensable pour la détermination de la longitude en mer) ; en 1751 c'est son essai sur les courants océaniques qui est couronné ; en 1753 il s'agit d'un rapport sur l'effet des différentes forces sur un bateau ; en 1757 il présente une méthode permettant de réduire le tangage des bateaux en haute mer.

Daniel Bernoulli, dont toute la vie a été consacrée à la science, s'est éteint le 17 mars 1782 à Bâle.

5.8. Principales publications de Daniel Bernoulli relatives aux probabilités

- *Recherches physiques et astronomiques sur le problème proposé pour la seconde fois par l'Académie Royale des Sciences de Paris : Quelle est la cause physique de l'inclinaison des plans des orbites des planètes par rapport au plan de l'équateur de la révolution du soleil autour de son axe ; Et d'où vient que les inclinaisons de ces orbites sont différentes entre elles*, Recueil des pièces qui ont remporté les prix de l'Académie Royale des Sciences, 1735, 3, 93-122.
- *Specimen theoriae novae de mensura sortis*, Comment. Acad. Sci. Imp. Petrop., 5, 175-192 (1730-1731), 1738.
- *Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir*, Hist. Mém. Acad. R. Sci., Paris, 1-45 (1760) 1766.
- *Disquisitiones analyticae de novo problemate conjecturali*, Novi Comment. Acad. Sci. Imp. Petrop., 14, 3-25 (1769) 1770.
- *Mensura sortis ad fortuitam successionem rerum naturaliter contingentium applicata*, Novi Comment. Acad. Sci. Imp. Petrop., 14, 26-45 (1769) ; 15, 3-28 (1770).
- *Dijudicatio maxime probabilis plurium observationem discrepantium atque verisimillima inductio inde formanda*, Acta Acad. Sci. Imp. Petrop., 1, 3-23 (1777) 1778.

5.9. Références relatives à la biographie de Daniel Bernoulli

- H. Straub, *Dictionary of Scientific Biography*, Charles Scribner's sons, New York, vol. 2, p. 36-46, 1970.
- E.T. Bell, *Men of mathematics*, Simon et Schuster, New York, 1986.

5.10. Biographie de Nicholas I Bernoulli (1687-1759)

Nicholas I (ou Niklaus ou Nicolas I) Bernoulli est né le 21 octobre 1687 à Bâle. C'est le neveu de Jacques (Jacob) Bernoulli et de Jean (Johann) Bernoulli. Dès sa jeunesse il a appris les mathématiques avec ses oncles. Jacques Bernoulli a supervisé la maîtrise ès arts (sur les séries) de Nicolas à l'université de Bâle, maîtrise qu'il obtient en 1704, à 17 ans. Il prépare alors un doctorat en

droit sur l'application des probabilités aux questions juridiques et l'obtient le 14 juin 1709. Comme cela était assez courant à cette époque, Nicolas, réalise, en 1712, un voyage en Europe passant par la Hollande, l'Angleterre (où il rencontre Newton et de Moivre) et la France. En France, où il passe deux mois, il rencontre Pierre Rémond de Montmort¹, les deux mathématiciens deviennent des amis proches et auront une riche correspondance. En 1716, Nicolas est nommé titulaire de la chaire de mathématiques de l'Université de Padoue, chaire jadis occupée par Galilée (entre 1592 et 1610). Son travail porte alors sur la géométrie et les équations différentielles. Il quitte l'Italie en 1722 pour revenir à Bâle où il devient professeur de logique à l'université de la ville. Toujours intéressé par les questions juridiques, il quitte en 1731 la chaire de logique pour occuper celle consacrée au droit. Il est par ailleurs, à quatre reprises, recteur de l'université. Nicolas n'est pas un mathématicien très prolifique, ses principales contributions se trouvent cachées dans sa volumineuse correspondance, en particulier celle avec Montmort, Leibniz et Euler. On trouve en particulier une longue discussion sur le problème dit de Saint-Petersbourg dans les lettres qu'il échange avec Montmort. Nicolas participe à la publication de l'œuvre majeure de son oncle, *Ars Conjectandi*, qui est resté inachevée à la mort de ce dernier, le 16 août 1705. Finalement cet ouvrage, paraîtra en 1713. Il publiera ensuite les œuvres complètes de son oncle. Dans ses travaux sur les équations différentielles, Nicolas soutient Leibniz et l'aide dans sa controverse avec Newton (en faisant remarquer, en particulier, le traitement incomplet de Newton sur les dérivées d'ordre supérieur). Nicolas a été apprécié de ses pairs, il est nommé membre de l'Académie de Berlin en 1713, *Fellow* de la *Royal Society* l'année suivante et membre de l'Académie de Bologne en 1724. Il meurt à Bâle le 29 novembre 1759.

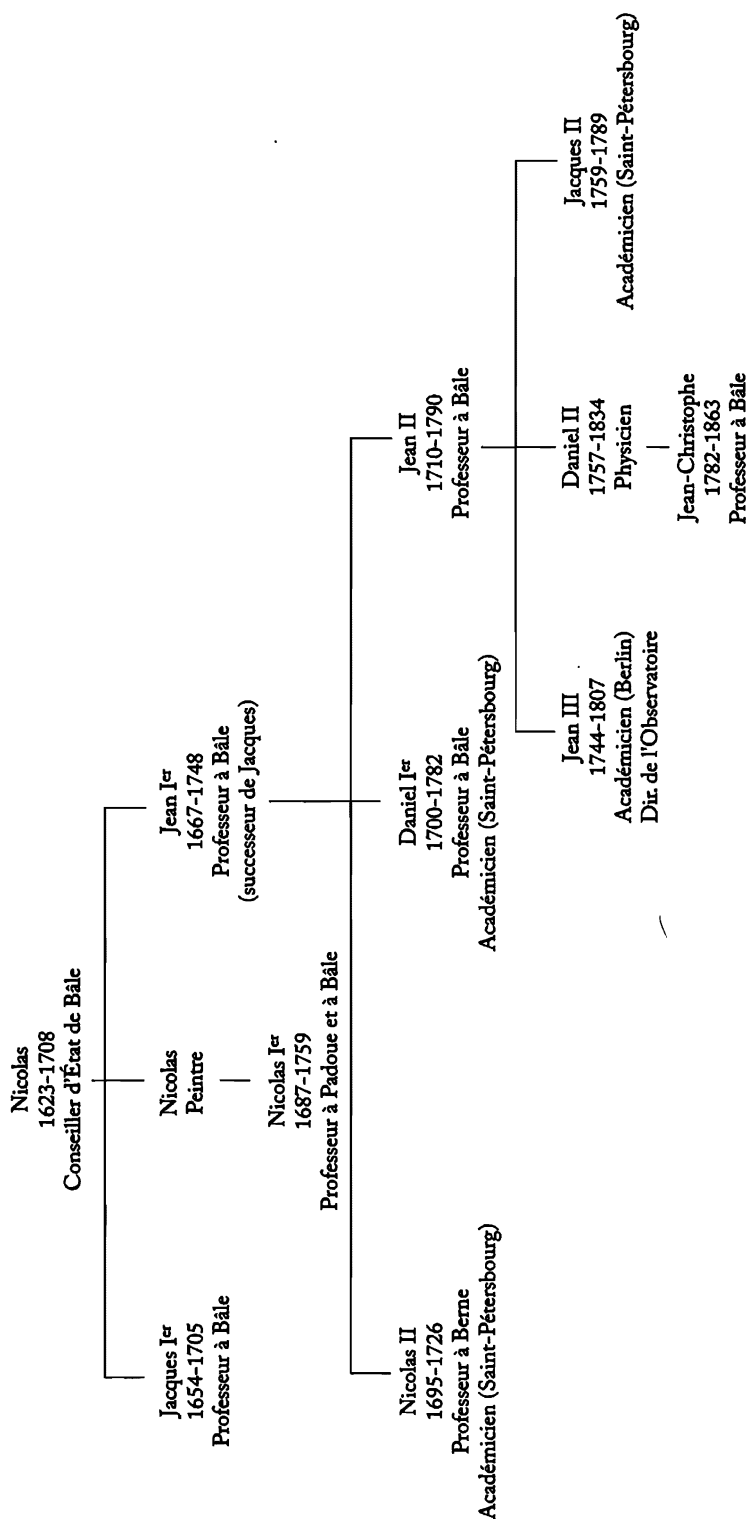
5.11. Ouvrages de Nicholas Bernoulli relatifs aux probabilités

- *De usu Artis Conjectandi in jure*, Nicolas Bernoulli, Conradus, Basilea, 1709.
- *Specimina artis conjectandi, ad quaestiones juris applicatae*, Nicolas Bernoulli, Acta Eruditorum, Supplementa, Tom. IV, Section IV, 159-170, 1711.

5.12. Références relatives à la biographie de Nicholas Bernoulli

- J. O. Fleckenstein, *Nikolaus I Bernoulli*, Dictionary of Scientific Biography, vol. 2, p. 56-57, Charles Scribner's sons, New York, 1970.

1. Voir biographie de Montmort, chapitre 4.



De Moivre

6.1. Le mémoire *De Mensura sortis*

Abraham de Moivre publie, en 1711, un mémoire intitulé *De Mensura sortis, seu de Probilitate Eventuum in Ludis a Casu Fortuito Pendentibus*¹. C'est, selon Karl Pearson², à la demande de Francis Robartes (1650-1718), *third Earl of Radnor*, que De Moivre avait étudié divers problèmes de probabilités en relation avec des jeux de hasard dont il publia les solutions dans le numéro entier des *Philosophical Transactions* pour le premier trimestre 1711.

En 1718, De Moivre développe ce texte initial en un livre intitulé *The Doctrine of Chances : or, a method of calculating the Probabilities of Events in Play*. Une seconde édition paraît en 1738 et une troisième, posthume, en 1756.

Le mémoire *De mensura sortis* de 1711 consiste en un exposé de vingt-six problèmes de probabilités dans les jeux de hasard que de Moivre traite par des méthodes différentes de celles publiées par Montmort trois ans plus tôt. Une courte controverse entre Montmort et de Moivre s'ensuit. Beaucoup de résultats importants sont publiés pour la première fois dans *De Mensura sortis*, bien que certains soient contenus dans l'*Ars Conjectandi* de Bernoulli qui ne paraîtra que deux ans plus tard.



S *p* fit numerus casuum quibus eventus aliquis contingere possit, & *q* numerus casuum quibus possit non-contingere; tam contingentia quam non-contingentia eventus suum habent probabilitatis gradum: Quod si casus omnes quibus eventus contingere vel non-contingere potest, sint æque faciles; probabilitas contingentia, erit ad probabilitatem non-contingentia ut *p* ad *q*.

Les premières phrases du *De Mensura sortis* donnent l'une des toutes premières définitions explicites de la probabilité :

1. In numéro 239 des *Philosophical Transactions* de Londres.
2. *The history of Statistics in the 17th and the 18th centuries*, edited by E.S. Pearson ; London, Griffin, 1978.

« Si p est le nombre de cas par lesquels un événement quelconque peut arriver et q le nombre de cas par lequel il peut ne pas arriver ; alors l'arrivée ou la non-arrivée ont leur degré de probabilité : Car si tous les cas par lesquels l'événement peut arriver ou ne pas arriver sont d'une égale facilité, la probabilité d'arrivée est à la probabilité de non-arrivée dans le rapport de p à q . »

6.2. The Doctrine of Chances et les *Miscellanea analytica de seriebus et quadraturis*

The Doctrine of Chances traite cinquante-trois problèmes dans sa première édition et vingt et un de plus dans sa troisième. On notera que de Moivre présente dans son traité une théorie des séries récurrentes qu'il applique au problème de la durée d'un jeu.

6.2.1. L'introduction des fonctions génératrices

De Moivre a proposé un lemme pour répondre au problème suivant¹ : Trouver le nombre de chances d'obtenir un nombre de points spécifié avec un nombre de dés quelconque, ayant tous le même nombre de faces.

« Soit $p + 1$, écrit de Moivre, le nombre de points spécifié, n le nombre de dés, f le nombre de faces de chacun d'eux. Soient $p - f = q$, $q - f = r$, $r - f = s$, $s - f = t$, etc. et le nombre de chances demandé sera

$$\begin{aligned}
 & + \frac{p}{1} \cdot \frac{p-1}{2} \cdot \frac{p-2}{3}, \text{etc.} \\
 & - \frac{q}{1} \cdot \frac{q-1}{2} \cdot \frac{q-2}{3}, \text{etc.} \times \frac{n}{1} \\
 & + \frac{r}{1} \cdot \frac{r-1}{2} \cdot \frac{r-2}{3}, \text{etc.} \times \frac{n}{1} \cdot \frac{n-1}{2} \\
 & - \frac{s}{1} \cdot \frac{s-1}{2} \cdot \frac{s-2}{3}, \text{etc.} \times \frac{n}{1} \cdot \frac{n-1}{2} \cdot \frac{n-2}{3} \\
 & \text{etc}
 \end{aligned}$$

La série devant être continuée jusqu'à ce que certains des facteurs dans chaque produit deviennent nuls ou négatifs.

N.B. Autant de facteurs doivent être pris dans chacun des produits $\frac{p}{1} \cdot \frac{p-1}{2} \cdot \frac{p-2}{3}, \text{etc.}$, $\frac{q}{1} \cdot \frac{q-1}{2} \cdot \frac{q-2}{3}, \text{etc.}$, qu'il y a d'unités dans $n - 1$.

Ainsi par exemple si l'on demande combien de chances il y a de faire 16 points avec 4 dés, alors en faisant $p + 1 = 16$, nous avons $p = 15$ et le nombre de chances demandé est :

$$+ \frac{15}{1} \cdot \frac{14}{2} \cdot \frac{13}{3} = 455$$

1. *The Doctrine of Chances*, 1738, pages 35-36.

(213)

(Numb. 329.)

PHILOSOPHICAL TRANSACTIONS.

For the Months of January, February, and March, 1711.

D E
MENSURA SORTIS,
SEU, DE

Probabilitate Eventuum in
Ludis a Casu Fortuito
Pendentibus.

Autore Abr. De Moivre, R. S. S.

$$\begin{aligned}
& -\frac{9}{1} \cdot \frac{8}{2} \cdot \frac{7}{3} \times \frac{4}{1} = -336 \\
& +\frac{3}{1} \cdot \frac{2}{2} \cdot \frac{1}{3} \times \frac{4}{1} \times \frac{3}{2} = +6
\end{aligned}$$

Total $455 - 336 + 6 = 125$. Le nombre de chances est cent vingt-cinq. »

Pour démontrer son lemme de Moivre considère un dé avec t faces marquées 1, t^2 faces marquées 2, et ainsi de suite jusqu'à t^f faces marquées f , si bien que le nombre total des marques (le nombre total de chances, écrit de Moivre) sur les faces égale $t + t^2 + \dots + t^f$. Selon la règle de multiplication, le nombre de chances pour les divers résultats d'un jet avec n dés est donné par les termes du produit $(t + t^2 + \dots + t^f)^n$. Ce que l'on recherche étant la somme des points, il faut additionner tous les produits avec le même exposant, disons s , pour trouver le nombre de chances d'avoir la somme s . On a donc à trouver le coefficient de t^s , ce que de Moivre obtient en écrivant

$$\begin{aligned}
(t + t^2 + \dots + t^f)^n &= t^n (1 - t^f)^n (1 - t)^{-n} \\
&= t^n \sum_{i=0}^n (-1)^i C_n^i t^{if} \sum_{j=0}^{\infty} C_{n+j-1}^j t^j.
\end{aligned}$$

Le coefficient de t^s est obtenu en prenant

$$\begin{aligned}
if + j &= s - n \text{ avec } i = 0, 1, \dots, [(s - n)/f]. \\
C_{n+j-1}^j &= C_{s-if-1}^{s-if-n} = C_{s-if-1}^{n-1}
\end{aligned}$$

Cette méthode a été reprise par Simpson et Lagrange pour calculer la distribution de la somme d'erreurs indépendantes, chacune ayant une distribution de probabilité discrète.

La méthode de De Moivre est en fait la méthode des fonctions génératrices. Si on associe à une série de deux nombres réels $p_0, p_1 \dots$ la série des puissances $g(t) = \sum p_x t^x$ et si cette série des puissances converge pour un intervalle de t , alors la fonction $g(t)$ est appelée la fonction génératrice de la série de nombres réels $p_0, p_1 \dots$. Si p_x correspond à la probabilité qu'une variable aléatoire prenne la valeur x et si $\sum p_x = 1$, alors $g(t)$ converge pour $|t| \leq 1$.

6.2.2. Les séries récurrentes

[On appelle *série récurrente* toute série dont chaque terme est la somme d'un nombre fixe des termes précédents multipliés respectivement par des coefficients constants. La relation qui existe entre le terme général V_n et les termes précédents $V_{n-1}, V_{n-2}, \dots$ est la *loi* de la série. Lagrange a nommé *équation génératrice* de cette série l'équation en x qu'on obtient en remplaçant, dans la loi, $V_n, V_{n-1}, V_{n-2}, \dots$ respectivement par $x^n, x^{n-1}, x^{n-2}, \dots$ et en supprimant

ensuite la plus haute puissance de x qui se trouve en facteur commun dans tous les termes.]

De Moivre a été le premier à publier à publier ses résultats sur les séries récurrentes en 1718, mais sans les démontrer. Il avait, cependant pris la précaution de déposer ses démonstrations à la *Royal Society*.

On trouve dans les pages 193-202 des *Doctrine of Chances* de 1738, l'exposé de De Moivre. Il écrit :

« J'appelle série récurrente une série qui est continuée de telle sorte qu'ayant pris un nombre quelconque de ses termes, chaque terme subséquent est obtenu à partir d'une relation sur un même nombre de termes précédents, selon une loi constante, comme la série suivante

$$1 + 2x + 3x^2 + 10x^3 + 34x^4 + 97x^5, \text{etc.}$$

dans laquelle les termes, représentés respectivement par les capitales A, B, C, D, etc. seront :

$$D = 3Cx - 2Bx^2 + 5Ax^3$$

$$E = 3Dx - 2Cx^2 + 5Bx^3$$

$$F = 3Ex - 2Dx^2 + 5Cx^3$$

etc.

Maintenant, les quantités $3x - 2x^2 + 5x^3$, prises ensemble et connectées à leurs propres signes, est ce que j'appelle l'Index ou l'Échelle de la relation ; et quelquefois les simples coefficients $3 - 2 + 5$ sont appelés l'Échelle de la relation. »

De Moivre donne ensuite neuf propositions, sans démonstrations.¹

Déjà, à la page 128 de *The Doctrine of chances* de 1718, de Moivre écrivait :

« Si dans une série quelconque, les termes A, B, C, D, E, F, etc sont continuellement décroissants et sont liés les uns aux autres de telle sorte qu'ils ont avec le même nombre de termes précédents une relation identique, relation toujours exprimable avec le même Index ; je dis que la somme de tous les termes de cette série, *ad infinitum* peut toujours être obtenue. »

Si on appelle ordre d'une série récurrente la puissance la plus élevée de x dans ce qui a été défini plus haut comme l'Index, on peut décrire comme suit la démarche de De Moivre pour trouver la somme d'une série récurrente :

On définit une séquence récurrente $\{r_n\}$, avec $n = 0, 1, \dots$, d'ordre k , au moyen de deux ensembles de k nombres $r_0, r_1, \dots, r_{k-1}$ et $a_1, a_2, \dots, a_k$ où $a_k \neq 0$, et la relation de récurrence

$$r_n = \sum_{i=1}^k (-1)^{i-1} a_i r_{n-i} \quad \text{avec } n \geq k$$

les coefficients $\{(-1)^{i-1} a_i\}$ étant l'Échelle de la relation définie plus haut.

1. Cf. I. Todhunter, *History of the theory of probability*, 1865, p. 179-183 et A. Hald, *History of probability... before 1750*, Wiley, 1990, § 23.1.

Le polynôme

$$A(x) = \sum_{i=1}^k (-1)^i a_i x^i \quad \text{avec } a_0 = 1$$

est appelé l'Échelle différentielle. C'est la fonction génératrice pour les coefficients $\{(-1)^{i-1} a_i\}$.

De Moivre montre que pour une série récurrente $R(x) = \sum_{n=0}^{\infty} r_n x^n$, on a

$$R(x) = \frac{1}{A(x)} \sum_{n=0}^{k-1} x^n \sum_{i=0}^n (-1)^i a_i r_{n-i}.$$

Cette équation signifie que $R(x)$ peut être trouvé comme le rapport de deux polynômes de degrés $k-1$ et k , respectivement, dont les coefficients dépendent des $2k$ nombres définissant la séquence de récurrence.

On sait que, réciproquement, toute fraction algébrique rationnelle en x , réduite en série, donne lieu à une suite de termes dont chacun est égal à la somme algébrique d'un même nombre de termes précédents, multipliés respectivement par certaines quantités constantes, pour toute l'étendue de la série. L'ensemble des constantes par lesquelles on doit multiplier un certain nombre de termes précédents pour former un terme quelconque est appelé l'échelle de relation de la série.

6.2.3. Les *Miscellanea analytica de seriebus et quadraturis*

En 1730, de Moivre a rédigé un livre de mathématiques intitulé *Miscellanea analytica de seriebus et quadraturis*¹ qui comporte la formule de trigonométrie à laquelle son nom est associé :

$$(\cos \theta + \sqrt{-1} \sin \theta)^n = e^{n\theta\sqrt{-1}}$$

Le livre contient également la première tentative d'approximation de la somme des termes du binôme $(a+b)^n$ développé en série, qu'il complète par un papier, daté du 13 novembre 1733, dont le titre est explicite : *Approximatio ad summam terminorum binomii $(a+b)^n$ in seriem expansi*.² Il s'agit de la première publication de ce que nous appelons aujourd'hui la *Loi normale*, résultat qui est devenu l'outil le plus puissant de la théorie des probabilités pour les deux siècles qui ont suivi. De Moivre a traduit lui-même en anglais son mémoire et l'a reproduit dans la seconde (1738) puis la troisième édition (1756) de son ouvrage *Doctrine of Chances*.

1. *Miscellanea analytica de seriebus et quadraturis*, Londini, Tonson and Watts, 1730.

2. Ce mémoire a été imprimé en quelques exemplaires pour les amis de De Moivre. On en connaît trois copies (Prussische Staatsbibliothek de Berlin, University College de Londres et Bibliothèque Macclesfield, n° 1424 du catalogue).

Si Jakob Bernoulli avait établi dans la quatrième partie de son *Ars Conjectandi*, ce qu'il est convenu d'appeler la loi faible des grands nombres, les travaux de De Moivre abordent le calcul de la probabilité d'obtenir exactement m événements favorables en n essais. De Moivre établit que cette probabilité est le terme de rang m dans le développement en série du binôme $(a + b)^n$ soit $C_n^m \cdot a^m b^{n-m}$, où a est la probabilité d'obtenir m événements favorables et $b = 1 - a$. Il en déduit (page 237 de la seconde édition de *Doctrine of Chances* de 1738) que la probabilité d'obtenir un taux de succès compris entre deux limites fixées s'obtient en approximant la somme des termes du développement en série du binôme $(a + b)^n$.

6.3. Rappel concernant la loi des grands nombres

Toutes les épreuves répétées d'événements peuvent être assimilées à des tirages d'objets de différentes couleurs contenus dans une urne, l'objet tiré étant réintroduit dans l'urne afin de ne pas modifier les probabilités.

Le principe des probabilités composées, à savoir la probabilité d'un événement composé, résultant du concours de plusieurs événements indépendants les uns des autres, est le produit des probabilités simples de ces événements.

En supposant donc que des événements $A', A'', \dots$ soient tous des répétitions de l'événement A , de même que $B', B'', \dots$ des répétitions de l'événement B , on a, si p est la probabilité de A , et q celle de B :

$$p = p' = p'' = \text{etc.} \quad q = q' = q'' = \text{etc.}$$

Soit n le nombre d'épreuves, le produit $(p + q)(p' + q')(p'' + q'') \dots = (p + q)^n$, qui peut être écrit, par la formule du binôme de Newton,

$$(p + q)^n = \sum_{k=0}^n C_n^k p^k q^{n-k},$$

$$\text{avec } C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} \quad (\text{coefficients binomiaux})$$

Le terme général $C_n^k p^k q^{n-k}$ exprime la probabilité que, dans n épreuves, les événements A et B arriveront respectivement k et $n - k$ fois.

La somme des termes du développement en série de $(p + q)^n$, depuis le premier jusqu'au terme $C_n^k p^k q^{n-k}$ inclusivement, exprime la probabilité que, dans m épreuves, l'événement A n'arrivera pas moins de k fois ou que l'événement complémentaire B n'arrivera pas plus de $n - k$ fois.

Il est à noter que Karl Pearson commente¹ la note de 1733 de De Moivre en lui attribuant la priorité sur trois points :

1. Biometrika, 16, p. 402-404, 1924 et 17, p. 201-210, 1925.

- les calculs relatifs à l'intégrale de probabilité et essentiellement à la loi normale,
- la formulation et l'usage de ce qui est appelé improprement le théorème de James Stirling,
- l'énoncé du théorème établissant que la mesure de la précision de l'évaluation d'une probabilité dépend de l'inverse de la racine carrée de la taille de l'échantillon testé.

En ce qui concerne les factorielles, Pearson écrit :

« De Moivre trouve le rapport du terme maximum d'un binôme au terme distant de x de ce maximum ; Il suppose l'exposant du binôme si élevé que l'on peut pratiquement écrire

$$m! = \text{Constante} \cdot \sqrt{m} \cdot e^{-m} \cdot m^m.$$

Il détermine la constante, qu'il appelle B , par le théorème qui donne le logarithme hyperbolique de B ;

$$\log_e B = 1 - \frac{1}{12} + \frac{1}{360} - \frac{1}{1260} + \frac{1}{1680}, \text{etc.}$$

ce qui donne $\log_e B = 0,3992235$ ou $B = 2,5074$.

D'où

$$m! = 2,5074 \sqrt{m} \cdot e^{-m} \cdot m^m. »$$

6.4. Le commentaire de Laplace

Laplace, dans sa *Notice historique sur les Probabilités*¹, commente ainsi les travaux de De Moivre :

« Moivre a repris dans son Ouvrage le théorème de Jacques Bernoulli sur la probabilité des résultats déterminés par un grand nombre d'observations. Il ne se contente pas de faire voir, comme Bernoulli, que le rapport des événements qui doivent arriver approche sans cesse de celui de leurs possibilités respectives ; il donne de plus une expression élégante et simple de la probabilité que la différence de ces deux rapports est contenue dans des limites données. Pour cela, il détermine le rapport du plus grand terme du développement d'une puissance très élevée du binôme à la somme de tous ses termes, et le logarithme hyperbolique de l'excès de ce terme sur les termes qui en sont très voisins. Le plus grand terme étant alors le produit d'un nombre considérable de facteurs, son calcul numérique devient impraticable. Pour l'obtenir par une approximation convergente, Moivre fait usage d'un théorème de Stirling sur le terme moyen du binôme élevé à une haute puissance, théorème remarquable surtout en ce qu'il introduit la racine carrée du rapport, de la circonférence au rayon, dans une expression qui semble devoir être étrangère à cette transcendante. Aussi Moivre fut-il extrêmement frappé de ce résultat, que Stirling avait déduit de l'expression de la circonférence en produits infinis, expression à laquelle Wallis était

1. Cf., par exemple, *Essai philosophique sur les Probabilités*, 3^e éd., Paris, Courcier, 1816, page 212.

parvenu par une singulière analyse qui contient le germe de la théorie si curieuse et si utile des intégrales définies. »

6.5. Extrait du mémoire de 1733

Approximatio ad summam terminorum binomii $(a + b)^n$

Nous traduisons ci dessous un extrait du mémoire de 1733 de De Moivre à partir de sa propre traduction anglaise ultérieure intitulée :

« *A method of approximating the terms of the binomial $(a + b)^n$ expanded to a series from whence are deduced some practical rules to estimate the degree of assent which is to be given to experiments.* »

[Méthode pour approximer la Somme des termes du Binôme $(a + b)^n$ développé en Série, d'où sont déduites des règles pratiques pour estimer le degré d'assentiment à attribuer aux expériences¹.]

« [...] I. Il y a maintenant une douzaine d'années ou plus que j'ai trouvé ce qui suit : Si le binôme $1 + 1$ est élevé à une très grande puissance n , le rapport du terme central à la somme de tous les termes, c'est-à-dire, à 2^n , peut être exprimé par la

fraction $\frac{2A.(n-1)^n}{n^n \cdot \sqrt{n-1}}$ où A est le nombre dont le logarithme hyperbolique est

$\frac{1}{12} - \frac{1}{360} + \frac{1}{1260} - \frac{1}{1680}$ etc. Mais puisque la quantité $\frac{(n-1)^n}{n^n}$ ou $(1 - \frac{1}{n})^n$ est très près d'être donnée quand n est une grande puissance, ce qui n'est pas difficile à prouver, il suit que, pour une puissance infinie, cette quantité sera absolument donnée, et représente le nombre dont le logarithme hyperbolique est -1 ; d'où il suit, que si B désigne le nombre dont le logarithme hyperbolique est

$$-1 + \frac{1}{12} - \frac{1}{360} + \frac{1}{1260} - \frac{1}{1680} \text{ etc.}$$

l'expression écrite ci-dessus deviendra $\frac{2B}{\sqrt{n-1}}$ ou approximativement $\frac{2B}{\sqrt{n}}$. Si

nous changeons les signes de cette série, et supposons maintenant que B désigne le nombre dont le logarithme hyperbolique est $1 - \frac{1}{12} + \frac{1}{360} - \frac{1}{1260} + \frac{1}{1680}$ etc. cette expression sera changée en $\frac{2}{B\sqrt{n}}$.

Quand j'ai pour la première fois commencé cette recherche, je me suis contenté de déterminer en gros la valeur de B, en additionnant quelques termes de la série ci-dessus ; mais comme j'ai pris conscience à la fois qu'elle convergerait très lentement, et que j'avais répondu d'une façon acceptable à ma question j'ai renoncé à aller plus loin jusqu'à ce que mon savant Ami M. *James Stirling*, qui s'était appliqué après moi dans cette recherche, ait découvert que la quantité B représentait en fait la racine carrée de la circonférence d'un cercle de rayon unité, si bien que si cette

1. *The Doctrine of Chances, the second edition, by A. de Moivre*, London H. Woodfall, 1738, p. 235-243.

APPROXIMATIO AD SUMMAM TERMINORUM BINOMII

$\overline{a+b^n}$ in Seriem expansi,

Autore A. D. M. R. S. S.

Quanquam solutio Problematum ad sortem spectantium non raro exigit ut plures Termini Binomii $\overline{a+b^n}$ in summam colligantur; attamen in potestatibus excelsis res adeo laboriosa videtur, ut pauci hoc opus aggredi curaverint; *Jacobus & Nicolaus Bernoulli* viri Doctissimi primi quod sciam tentarunt quid sua industria in hoc genere prestare posset, in quo etiamsi uterque propositum summa cum laude sit affecutus, aliquid tamen ultra potest requiri, hoc est approximatio ad summam; non enim tam de approximatione videntur fuisse solliciti quam de assignandis certis limitibus quos Summa Terminorum necessario transcenderet. Quam vero viam illi tenuerint, breviter in Miscellaneis meis exposui * quæ consulat Lector si vacat, quod ipsi tamen scripserint melius erit fortasse consulere: Ego quoque in hanc disquisitionem incubui; quod autem eo me primum impulit non profectum fuit ab opinione me cæteros anteiturum, sed ab obsequio in Dignissimum virum qui mihi autor fuerat ut hæc susceperem; Quicquid est, novas cogitationes prioribus subnecto, sed eò ut connexio postremorum cum primis melius appareat, mihi necesse est ut pauca jampridem a me tradita denuo proferam.

I. Duodecim jam sunt anni & amplius cum illud inveneram; si Binomium $1+1$ ad potestatem n permagnam attollatur, ratio quam Terminus Medius habet ad summam Terminorum omnium, hoc est ad 2^n , ad hunc modum poterit exprimi $\frac{2 A \times n - 1^n}{n^n \sqrt{n-1}^n}$, ubi A eum numerum exponit cujus Logarithmus

A

rithmus

* Vide Miscellanea Analytica pag. 96. 97. 98. 99.

circonférence est appelée c , le rapport du terme central à la somme de tous les termes sera exprimé par $\frac{2}{\sqrt{nc}}$.

Mais bien qu'il ne soit pas nécessaire de connaître quelle relation peut exister entre le nombre B et la circonférence du cercle, pourvu que sa valeur soit atteinte, soit en poursuivant la série logarithmique mentionnée plus haut, soit par un autre moyen, je dois dire, avec plaisir, que cette découverte, outre qu'elle a évité le doute, a apporté une élégance singulière à la solution.

II. J'ai aussi découvert que le logarithme du rapport que le terme central, pour une grande puissance, a à un terme quelconque qui lui est distant d'un intervalle noté l , pouvait être représenté avec une bonne approximation, (en posant $m = 1/2 n$) par la quantité

$$(m + l - \frac{1}{2})\text{Log}(m + l - 1) + (m - l + \frac{1}{2})\text{Log}(m - l + 1) - 2m\text{Log}m + \text{Log} \frac{m+1}{m}.$$

COROLLAIRE 1.

Ceci étant admis, je conclus, que si m ou $\frac{1}{2}n$ est une quantité infiniment grande, alors le logarithme du rapport qu'a un terme distant du terme central d'un intervalle l à ce terme central est $\frac{-2ll}{n}$. »

Notons qu'avec nos notations modernes le terme nc de De Moivre est équivalent à $8\sigma^2\pi$, où σ est l'écart type de la courbe dont l'ordonnée du maximum est $\frac{1}{\sigma\sqrt{2\pi}}$. On a également $n = 4\sigma^2$ et l'équation de la courbe

$$\text{est } y = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}}.$$

Le corollaire 1 de De Moivre est donc le suivant : Si n est grand, on peut écrire, en notations modernes :

$$\ln \frac{P(x = \frac{n}{2} + l)}{P(x = \frac{n}{2})} \cong -\frac{2l^2}{n}, \text{ ou } P(x = \frac{n}{2} + l) \cong \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-l^2}{2\sigma^2}}.$$

Dans les corollaires (pages 238-239 de la seconde édition de 1738), de Moivre montre que la différence probable entre la fréquence observée des événements par rapport au vrai rapport des probabilités décroît comme l'inverse du carré du nombre d'essais. Il approxime également la probabilité que la fréquence observée soit comprise entre deux limites spécifiées.

De Moivre donne également dans le même mémoire la valeur de l'intégrale

$$\int_{-a}^{+a} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx, \text{ calculée comme approximation de la probabilité binomiale}$$

$$P\left(\frac{n}{2} - \frac{a\sqrt{n}}{2}\right) \leq x \leq \left(\frac{n}{2} + \frac{a\sqrt{n}}{2}\right), \text{ et trouve :}$$

a	valeur trouvée par de Moivre	valeur exacte
1	0,682688	0,682689
2	0,95428	0,95450
3	0,99874	0,99730

En partant du fait que la somme de n variables aléatoires de *Bernoulli*, valant 1 avec la probabilité p et sinon 0, suit une loi binomiale de paramètres (n, p) et en utilisant la formule de *Stirling* pour approcher la factorielle, de Moivre a montré que la distribution binomiale, pour n grand, et $p = 1/2$, tend vers la loi normale.

Laplace généralisera en 1812 ce résultat au cas d'une probabilité p quelconque.

6.6. Les causes premières du hasard selon de Moivre

De Moivre, comme la plupart des savants de son époque croyait en un dessein divin.

Matthieu Maty, dans le *Journal britannique pour les mois de septembre et d'octobre 1755*¹, que nous avons déjà cité, au sujet de la controverse entre Montmort et de Moivre rapporte quelques phrases de De Moivre à ce sujet :

« On peut considérer, disait de Moivre, le hasard et le dessein comme deux compétiteurs, qui se disputent certains événements, et dont les droits respectifs sont soumis au calcul des probabilités. Supposons que les cartes de deux jeux de piquet fussent également disposées dans l'un et dans l'autre, il y a plus de 2 631 308 fois un million multiplié trois fois par lui-même à parier contre un, que la volonté a eu part à cette disposition. Ainsi la doctrine, qui trouve le hazard où il est, s'élève par une suite de probabilités à la démonstration qu'où s'observent l'ordre, la constance, et l'uniformité, là aussi résident le dessein et le choix... Je m'estimerais heureux, disait-il aussi dans son *Épître dédicatoire à Newton* mise au-devant de la seconde édition de son livre, si ayant donné à mes lecteurs une méthode pour calculer les effets du sort dans les jeux, et proposé des règles pour estimer jusqu'où certains événements tirent leur origine de la volonté plutôt que du hasard, je pouvais en exciter d'autres à pousser plus loin ces recherches, et à apprendre de votre Philosophie à recueillir par un calcul exact les indices certains de la sagesse et de l'intelligence parfaites, que l'Univers présente dans les divers phénomènes de la nature ».

1. *Mémoire sur la Vie et sur les Écrits de Mr. De Moivre.*

*A Method of approximating the Sum of the Terms
of the Binomial $a + b^n$ expanded into a Series,
from whence are deduced some practical Rules
to estimate the Degree of Assent which is to be
given to Experiments.*

ALTHO' the Solution of Problems of Chance often require that several Terms of the Binomial $a + b^n$ be added together, nevertheless in very high Powers the thing appears so laborious, and of so great a difficulty, that few people have undertaken that Task; for besides *James* and *Nicolas Bernoulli*, two great Mathematicians, I know of no body that has attempted it; in which, tho' they have shewn very great skill, and have the praise which is due to their Industry, yet some things were farther required; for what they have done is not so much an Approximation as the determining very wide limits, within which they demonstrated that the Sum of the Terms was contained. Now the Method which they have followed has been briefly described in my *Miscellanea Analytica*, which the Reader may consult if he pleases, unless they rather chuse, which perhaps would be the best, to consult what they themselves have writ upon that Subject: for my part, what made me apply myself to that Inquiry was not out of opinion that I should excel others, in which however I might have been forgiven; but what I did was in compliance to the desire of a very worthy Gentleman, and good Mathematician, who encouraged me to it: I now add some new thoughts to the former; but in order to make their connexion the clearer, it is necessary for me to resume some few things that have been delivered by me a pretty while ago.

I. It is now a dozen years or more since I had found what follows; If the Binomial $1 + 1$ be raised to a very high Power de-

H h 2

noted

ANNUITIES
UPON
LIVES:
OR,
The VALUATION of
ANNUITIES upon any
Number of LIVES; as
also, of REVERSIONS.
To which is added,
An APPENDIX concerning the
EXPECTATIONS of LIFE, and
Probabilities of SURVIVORSHIP.

By A. DE MOIVRE. F. R. S.

L O N D O N,
Printed by WILLIAM PEARSON,
over-against Wright's-Coffee-House in
Aldersgate-Street. MDCCXXV.

6.7. Les *Annuities upon lives*

Si *The Doctrine of chances* ne comportaient pas de statistiques, de Moivre s'est, cependant, intéressé au sujet et a publié en 1725 un livre intitulé *Annuities upon lives ; or the valuation of annuities upon any number of lives ; or also, of reversions. To which is added an appendix concerning the expectations of life, and probabilities of survivorship*¹.

La question de savoir qui inventa les statistiques est complexe. L'économiste allemand Gottfried Achenwall (1719-1772) est considéré, par certains, comme l'inventeur de la statistique car ayant publié en 1748 la *Constitution des royaumes et États de l'Europe* et créé le mot *statistique*, dans le sens *études constitutionnelles*. Mais si on considère, comme le fait Karl Pearson, que les statistiques sont essentiellement l'établissement de données numériques relatives à l'État, les fondateurs de cette discipline sont les Anglais John Graunt (1620-1674) et Sir William Petty (1623-1687). Cependant, l'ouvrage de John Graunt, *Natural and Political Observations on the Bills of Mortality of London*, publié en 1662 et l'analyse de Sir William Petty dans son traité intitulé *Political Arithmetic* (1682), n'offraient pas le degré de précision souhaité par de Moivre qui se servit des tables de mortalité obtenues par Edmund Halley². Dans son mémoire Halley indique que les travaux de Graunt et Petty souffraient du fait que les populations des villes étudiées n'étaient pas stables et que des tables plus précises relatives à la ville allemande de Breslau, qui avaient été communiquées à la *Royal Society* par le professeur allemand Caspar Neumann (1648-1715), lui avaient été transmises par M. Justell. Ces tables mentionnaient à la fois l'âge et le sexe de toutes les personnes décédées durant les années 1687 à 1691, et les comparaient mensuellement aux naissances survenues dans le même temps.³

Dans son ouvrage, de Moivre postule un taux d'intérêt monétaire constant et un taux de mortalité donné par la table de E. Halley, dont il publie d'ailleurs les données, et pour établir sa théorie, il traite mathématiquement les annuités relatives aux assurances sur des vies multiples, le problème des héritages, celui des tontines et plus généralement ceux des contrats où l'âge des personnes et le niveau des taux d'intérêt sont les paramètres essentiels.

1. London, W.P. 1725.

2. *An estimate of the degrees of mortality of mankind ; drawn from curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the prices of annuities upon lives*, Philosophical Transactions, 17, n° 196, p. 596-610, 654-656, 1693.

3. Cf. notre chapitre 18.

Dr. HALLEY'S TABLE of Observations, exhibiting
the Probabilities of LIFE.

Age.	Perfons.	Age.	Perfso.	Age.	Perfso.	Age.	Perfso.	Age.	Perfso.	Age.	Perfso.
1	1000	8	680	15	628	22	586	29	539	36	481
2	855	9	670	16	622	23	579	30	531	37	472
3	798	10	661	17	616	24	573	31	523	38	463
4	760	11	653	18	610	25	567	32	515	39	454
5	732	12	646	19	604	26	560	33	507	40	445
6	710	13	640	20	598	27	553	34	499	41	436
7	692	14	634	21	592	28	546	35	490	42	427
Age.	Perfons.	Age.	Perfso.	Age.	Perfso.	Age.	Perfso.	Age.	Perfso.	Age.	Perfso.
43	417	50	346	57	272	64	202	71	131	78	58
44	407	51	335	58	262	65	192	72	120	79	49
45	397	52	324	59	252	66	182	73	109	80	41
46	387	53	313	60	242	67	172	74	98	81	34
47	377	54	302	61	232	68	162	75	88	82	28
48	367	55	292	62	222	69	152	76	78	83	23
49	357	56	282	63	212	70	142	77	68	84	20

Un excellent résumé de l'ouvrage *Annuities upon lives* a également été donné par Matthieu Maty, dans le même *Journal britannique pour les mois de septembre et d'octobre 1755*¹.

Voici les phrases de Maty :

« L'éditeur du livre de Mr. Jacques Bernoulli avoit invité Mr. de Moivre à suivre les vues du célèbre Auteur, en appliquant la science des probabilités aux usages de la vie. Notre Savant s'était excusé d'entreprendre cette nouvelle tâche. Il semble cependant avoir été déterminé par cette invitation au travail, qu'il entreprit en 1721, pour déterminer les problèmes fondés sur les probabilités de la vie humaine. L'Angleterre est peut-être le pays, où les questions relatives à la valeur des rentes viagères, des contrats de substitution, des achats d'expectatives etc., reviennent le plus souvent. On n'y suivait, avant Mr. de Moivre, qu'une routine aveugle et toujours erronée. Aussi nos insulaires reçurent-ils avec avidité les règles simples générales et précises, que Mr. de Moivre donna dans son *Traité sur les rentes à vie* publié pour la première fois en 1724 et pour la seconde en 1743. Comme la théorie sur laquelle il les fonde lui appartient en propre, je ne puis encore éviter quelques détails, que je crains de mal rendre en voulant trop les abrégier. Dès l'année 1692, le Dr. Halley avoit composé une table de mortalité sur les registres de Breslaw. Il avoit même indiqué quelques moyens de déduire de cette table la valeur des annuités sur une ou plusieurs vies. Mais ces moyens exigeaient pour chacune des vies autant d'opérations arithmétiques qu'il y avoit d'années entre l'âge actuel et celui de cent ans. S'il s'agissait de trouver les sommes et les différences de plusieurs vies, les combinaisons se multipliaient à un point prodigieux, et l'inventeur convenait lui-même que malgré les facilités que procuraient les logarithmes, il étoit à souhaiter qu'on trouvât quelque méthode plus abrégée que les siennes. Ce que Halley avoit vainement cherché ne devoit

1. *Mémoire sur la Vie et sur les Écrits de Mr. De Moivre.*

pas être facile. Mr. de Moivre s'y appliqua cependant, et les succès surpassèrent ses espérances. Il vit d'abord dans la table de Halley ce qu'il est surprenant que Halley n'y eut point vu, des intervalles de plusieurs années, pendant lesquels la vie humaine décroît uniformément. De 646 adultes de 12 ans, restes de mille qui ont eu les accidents de l'enfance à soutenir, il en meurt six dans un an, douze dans deux et ainsi de suite jusqu'à 25 ans. Chacune des quatre années suivantes en emporte sept. De 29 à 34 ans la proportion annuelle est de huit, elle est de neuf jusqu'à 42 ans, de dix jusqu'à 49, et de onze jusqu'à 54. Elle se réduit à dix jusqu'à l'âge de 70 ans, remonte à onze jusqu'à 74, et revient à dix jusqu'à 78. Les progrès de la mortalité suivent ensuite une proportion arithmétique de neuf, de huit, de sept, de six pendant les quatre années suivantes, et des vingt personnes qui restent en vie à 86 ans, une tout au plus parvient au siècle. Mr. de Moivre, non content d'avoir trouvé ces intervalles, qui seuls abrègent considérablement les calculs, observe que leurs inégalités se balancent, et qu'ils peuvent être regardés comme les parties d'une progression arithmétique, que fourniraient probablement des observations plus exactes et plus multipliées. Le premier terme de cette progression peut être fixé à 12 ans et le dernier à 86. De 74 adolescents au premier de ces deux âges, il en doit mourir un par an, et la distance de l'âge de qui que ce soit au terme le plus reculé est le complément de sa vie. À chaque âge répond une suite, qui exprime la probabilité de sa durée, et qui multipliée par celle d'une rente annuelle pour ce nombre d'années, représente sa valeur. Mr. de Moivre la somme sans peine, et trouve ainsi une formule très simple, qui s'applique à quelque âge que ce soit ; elle n'exige que quatre opérations aisées, et par le moyen des tables le moindre arithméticien peut la calculer. Les mêmes principes ont lieu dans les cas de vies conjointes, les survivances, les hypothèques, etc. Les règles de notre Mathématicien sont si faciles, qu'en les suivant, *on avance plus en un quart d'heure qu'on ne pouvait le faire par les méthodes précédentes en un quart d'an*. Les rentes évaluées de cette manière sont cependant assujetties à cette condition, que le paiement s'en fait chaque année, et que celle où meurt le possesseur est perdue pour ses héritiers. Que si l'on veut changer cette condition, et ne faire finir les paiements qu'à l'instant même de la mort, notre Mathématicien satisfait à ce nouveau problème dans un Mémoire présenté à la Société Royale en 1744, où il montre en même temps le moyen de lier entr'eux divers intervalles de la vie, et d'en calculer les probabilités, en s'en tenant aux observations. »

6.8. Biographie de De Moivre (1667-1754)

Abraham de Moivre est né à Vitry-le-François, en Champagne, le 26 mai 1667. Son père, chirurgien, vivait modestement mais a toujours considéré comme prioritaire l'éducation de ses enfants. Le jeune Abraham a d'abord été éduqué à Vitry sous la conduite d'un précepteur. Puis, bien que la famille de Moivre soit protestante, son père l'inscrit dans une école catholique tolérante de Vitry où il restera jusqu'à l'âge de onze ans. Il poursuit alors ses études dans une autre école, protestante, cette fois, à Sedan. Il est en pension chez son professeur de Grec et se montre assidu et travailleur. C'est à ce moment qu'il découvre les mathématiques par l'intermédiaire d'un traité d'arithmétique. Cette distraction est tolérée par son professeur qui regrette

tout de même que ce temps consacré aux mathématiques ne le soit pas à l'étude du Grec. En 1682, alors que le jeune Abraham a déjà passé quatre ans dans cette école protestante de Sedan, cette dernière est supprimée. Ses parents l'envoient alors à Saumur, pour y achever ses études. Son professeur de philosophie est un homme entêté, très attaché aux anciennes doctrines, méprisant ouvertement la méthode de Descartes. C'est de nouveau par lui-même que de Moivre va poursuivre sa découverte des mathématiques, cette fois en lisant le traité de Huygens sur les jeux de hasard *De ratiociniis in ludo aleae*, qui avait été publié en 1657. Entre-temps, ses parents se sont installés à Paris et leur fils poursuivra ses études au Collège d'Harcourt¹. Il y suivra des cours de physique et de mathématiques et bénéficiera, de plus, de cours particuliers. Une anecdote montre bien la passion de De Moivre pour les mathématiques². Un jour, il va avec l'un de ses parents dans une maison en Bourgogne et découvre Les Éléments d'Euclide qu'il se met à dévorer. Il est arrêté par la cinquième proposition et se met à pleurer. Ses parents ne peuvent le consoler qu'en lui expliquant cet axiome, après quoi il poursuit la lecture des six livres sans difficulté.

À la révocation de l'Édit de Nantes par Louis XIV en 1685, la vie de De Moivre va basculer. Il est d'abord emprisonné, pour ses croyances religieuses, dans le prieuré de Saint Martin. Nous ne connaissons pas exactement la durée de cet emprisonnement, ce qui est certain est, qu'une fois libéré, il émigra en Angleterre. Il a vingt et un ans (1688) lorsqu'il arrive dans ce pays et pense, à tort, avoir atteint le sommet des mathématiques. Le hasard fait qu'il est reçu chez Lord Devonshire qui vient de recevoir un exemplaire des *Principia Mathematica* que son auteur, Isaac Newton, lui a personnellement adressé. De Moivre ouvre l'ouvrage pensant le lire sans difficulté et découvre qu'il est hors de portée de ses connaissances. Nullement découragé, il s'en procure un exemplaire, en déchire, dit-on, les feuillets pour les mettre dans sa poche de façon à les étudier à chaque moment libre.

La situation de De Moivre n'est guère brillante, il souhaiterait obtenir un poste de professeur, mais étant étranger il n'a que très peu de chances d'y parvenir ; il se contente donc de quelques cours particuliers et de quelques leçons données dans des salons de thé londoniens. Sa réputation grandit néanmoins et, en 1692, il rencontre Halley³ avec qui il se lie d'une étroite amitié. Probablement grâce à son intermédiaire, il rencontre également Newton avec qui il sera également proche. Le premier article de De Moivre portera d'ailleurs sur une extension de la méthode des fluxions de Newton. Halley présente cet article à la *Royal Society* en 1695 et il est publié la même année dans *les Transactions of the Royal Society*. Deux ans plus tard il communique

1. Maintenant Lycée Saint-Louis.

2. Anecdote tirée d'*Éloges de M. Moivre*.

3. Halley est alors assistant du secrétaire de la Royal Society.

un article donnant une méthode générale permettant d'élever ou d'abaisser tout multinôme $ax + bx^2 + cx^3 + dx^4 + \dots$ à quelque puissance que ce soit. Cet article, très apprécié, lui ouvrira les portes de la *Royal Society* où il est élu membre en 1697. Sous l'influence de Halley, de Moivre s'intéressa un temps à l'astronomie. Toutefois son intérêt pour les mathématiques pures domine. En 1707, il donne différentes formules pour résoudre des équations n'ayant que des puissances impaires de l'inconnue. Ce travail lui vaut les éloges des plus grands mathématiciens, en particulier de Jacques Bernoulli et de Leibniz. Ce dernier, conscient de la valeur de De Moivre, cherche à lui obtenir une chaire de mathématiques, mais en vain.

Alors que la dispute relative à la priorité de la découverte du calcul infinitésimal entre Newton et Leibniz fait rage¹, de Moivre est engagé (le 17 avril 1712) dans le comité mis en place par la *Royal Society* pour étudier la question². La *Royal Society* est alors présidée par Newton lui-même et l'amitié existant entre Newton et de Moivre n'est certainement pas pour rien dans le choix de ce dernier.

De Moivre présente en 1711 à la *Royal Society* son étude *De Mensura sortis* sur la doctrine du hasard qui est publiée la même année dans les *Transactions*. Ce travail sera ensuite amplifié en 1718 sous le titre : *The doctrine of chances*.

De Moivre s'intéresse également, sur les conseils de Halley, aux applications pratiques des probabilités et des statistiques. Il élabore des tables de mortalité précises (*Annuities on lives*, 1725) et donne des formules permettant de calculer équitablement le montant des rentes viagères. Il est également célèbre pour avoir introduit la formule qui porte son nom $[(\cos x + i \sin x)^n = \cos nx + i \sin nx]$, formule publiée en 1722, mais dont une forme proche avait déjà été perçue par de Moivre en 1707. En 1730 il fait paraître *Miscellanea Analytica*, ouvrage dans lequel apparaît pour la première fois la formule appelée à tort « de Stirling ». On doit également à de Moivre la révision de la traduction latine de l'*Optique* de Newton. Ce dernier et de Moivre se rencontraient alors tous les soirs pour discuter de cette traduction et d'autres aspects de la physique. De Moivre avait dédié son livre *Doctrine of Chances* à Newton, et ce dernier lui retourna le compliment en répondant à ceux qui le questionnaient sur les *Principia* : « Allez voir de Moivre, il connaît ces choses beaucoup mieux que moi », ce qui n'est pas un mince compliment lorsque l'on connaît la personnalité de Newton.

L'un de ses biographes, Clerke, fait le portrait suivant de la fin de vie de De Moivre : « De Moivre est resté célibataire toute sa vie et a passé ses dernières

1. Voir par exemple : *Trente livres de mathématiques qui ont changé le monde*, Jean-Jacques Samuelli, Jean-Claude Boudenot, Éd. Ellipses, 2006 (en particulier, chapitre 12 : *Isaac Newton mathématicien et physicien*).
2. Les autres commissaires étaient MM. Arbuthnot, Hill, Halley, Jones, Machin, Burnet, Robarts, Bonet, Aston et Taylor.

années à faire des études paisibles. La littérature, antique et moderne, lui a fourni sa détente ; il a dit une fois qu'il aurait plutôt préféré être Molière que Newton ; et connaissait son œuvre ainsi que celle de Rabelais presque par cœur. Il a été toute sa vie un homme dévoué et charitable. Après qu'il ait perdu successivement la vue et l'audition, il a encore connu le plaisir d'être élu, en tant qu'associé étranger, à l'Académie des sciences de Paris le 27 juin 1754. »

Dans ses dernières années il avait considérablement besoin de sommeil : il dormait régulièrement vingt heures par jour. Lors des quatre heures d'éveil son esprit était intact il parlait de tout, dictait des réponses précises aux lettres qu'on lui adressait, intervenait sur des questions d'algèbre, etc. On raconte, qu'à partir d'un certain âge, de Moivre s'est rendu compte qu'il dormait quelques minutes de plus chaque nuit. Il a ainsi calculé le jour de sa mort, qui devait coïncider avec un sommeil de 24 heures La dernière prédiction de De Moivre s'est, dit-on, révélée juste, il est mort le 27 novembre 1754.

6.9. Principales publications de De Moivre

- *De mensura sortis, seu de probabilitate eventuum in ludis à casu fortuito pendentibus*, Philos. Trans., 1711, n° 329.
- *The Doctrine of Chances : or, a method of calculating the Probabilities of Events in Play*, London, 1718 (une seconde édition paraît en 1738 et une troisième, posthume, en 1756).
- *Miscellanea analytica de seriebus et quadraturis*, Londini, Tonson and Watts, 1730.
- *Approximatio ad summam terminorum binomii $(a + b)^n$ in seriem expansi*. Ce très important mémoire a été imprimé à quelques exemplaires pour les amis de De Moivre.
- *Annuities upon lives; or the valuation of annuities upon any number of lives; or also, of reversions. To which is added an appendix concerning the expectations of life, and probabilities of survivorship*, London, W.P., 1725.
- Diverses publications ont paru dans les Philos. Trans. de Londres entre 1695 et 1744 (n° 216, 230, 240, 265, 278, 309, 329, 341, 345, 352, 360, 373, 374, 451, 473).

6.10. Références

- « Éloge de M. de Moivre », in *Histoire et Mémoires de l'Académie royale des sciences*, année 1754, [partie Histoire] p. 175-184 (publiée en 1769).
- Matthieu Maty, « Mémoire sur la Vie et sur les Écrits de Mr. De Moivre », in *Journal britannique pour les mois de septembre et d'octobre 1755*.
- I. Schneider, *Der Mathematiker Abraham de Moivre (1667-1754)*, Arch. Hist. Ex. Sci. 5, 177-317, 1968.
- Ian Hacking, *De Moivre*, Dict. of Sc. Biog., New York, Scribner, 1970, tome 9, p. 452-454.
- R. H. Daw and E. S. Pearson, *Studies in the history of probability and statistics. XXX: Abraham de Moivre's 1733 derivation of the normal curve : a bibliographical note*, Biometrika 59, 677-680, 1972.

Le principe de probabilité inverse de Bayes

7.1. La publication du mémoire de Bayes

Le révérend Thomas Bayes est l'auteur célèbre d'un unique mémoire, publié à titre posthume, en 1763, par son ami Richard Price et intitulé *An Essay towards solving a Problem in the Doctrine of Chances*¹ (*Un Essai pour résoudre un problème dans la théorie des probabilités*). Ce travail est relatif à l'estimation des causes par lesquelles un événement peut avoir été produit, c'est-à-dire à ce que l'on appellerait aujourd'hui la *distribution a posteriori* de la probabilité p d'une loi binomiale. Price indique dans sa lettre d'introduction que Bayes avait tenté de trouver une méthode par laquelle on pourrait évaluer la probabilité qu'un événement de se produire dans certaines circonstances, en supposant que l'on ignore tout au sujet de cet événement mais que, dans des circonstances identiques l'événement se soit produit un nombre de fois connu et qu'il ne se soit pas produit un autre nombre de fois également connu.

Bayes tentait donc d'élaborer un raisonnement analogique ou inductif afin de mesurer la probabilité avec laquelle les résultats d'une expérience passée peut être appliquée au futur. On nomme aujourd'hui *induction* ou *inférence bayésienne* la démarche logique permettant de calculer ou de réviser la probabilité d'une hypothèse.

L'*Essai* de Bayes avait été précédé par la publication d'une note de deux pages *A letter from the late Reverend Mr. Thomas Bayes, F. R. S. to John Canton, M. A. and F. R. S.*². Cette note est relative à divers problèmes de séries, et en particulier au théorème de De Moivre-Stirling. On y trouve la formule que Bayes utilisera dans son *Essai* :

1. La publication comporte, en fait deux mémoires successifs. Le premier occupe les pages 370-478 du volume LIII des Philosophical Transactions pour 1763, publié en 1764. Il a été lu à la R.S. le 23 décembre 1763. Le second occupe les pages 296-325 du volume LIV de la même revue, année 1764 publiée en 1765.
2. Le papier a été lu le 23 novembre 1763 à la R. S. et publié en 1764 : Philos. Trans. R. Soc. London 53, 269-271.

$$\log x! = \log \sqrt{2\pi} + \left(x + \frac{1}{2}\right) \log x - S,$$

$$\text{avec } S = \left[\frac{1}{12x} + \frac{1}{360x^3} - \frac{1}{1260x^5} + \frac{1}{1680x^7} - \frac{1}{1188x^9} + \dots \right]$$

L'Essai proprement dit débute par l'exposé du problème abordé. Bayes écrit :

« Étant donné le nombre de fois qu'un événement inconnu s'est réalisé ou a échoué, trouver la chance que la probabilité de survenue de l'événement dans une seule épreuve soit comprise quelque part entre deux degrés de probabilité que l'on peut choisir. »

Avec des notations modernes, on écrirait l'énoncé du problème, comme suit :

Si p est le nombre de fois où l'événement s'est produit en $n = (p + q)$ épreuves et x la probabilité d'occurrence en un essai unique, on cherche la probabilité $P(a < x < b/p)$.

Dans son mémoire, Bayes établit dans ses propositions 3 et 5 la formule des probabilités conditionnelles qui s'écrit, en notations modernes,

Probabilité de A sachant B : $P(A/B) = P(A \cap B) / P(B)$.

On appelle $P(A)$ probabilité *a priori* de A ; $P(A/B)$ probabilité *a posteriori* de A sachant B ; et $P(B)$ probabilité marginale ou *a priori* de B .

Le problème de la loi binomiale (considération d'une distribution uniforme du paramètre binomial x) est alors abordé par Bayes à l'aide d'une figure géométrique à l'aide de laquelle il décrit des quantités aléatoires continues comme des surfaces. Il obtient la probabilité qu'a une épreuve unique de donner une réalisation comprise entre a et b , sachant que l'événement s'est produit auparavant p fois en $n = (p + q)$ épreuves :

$$P(a < x < b/p) = \frac{\int_a^b x^p (1-x)^q dx}{\int_0^1 x^p (1-x)^q dx} \quad (\text{C'est la proposition 10 de Bayes dans son Essai.})$$

Théorème des Probabilités Composées Probabilité de A union B (A ou B) Probabilité de A Inter B (A et B à la fois)	$P[A \cup B] = P[A] + P[B] - P[A \cap B]$
Événement non A Probabilité de A barre	$P[A \cup \bar{A}] = P[A] + P[\bar{A}] = 1$ d'où $P[\bar{A}] = 1 - P[A]$
Probabilité Conditionnelles Probabilité de A sachant B (= Probabilité d'un événement dans une « Espace des épreuves » restreint)	$P[A/B] = P[A \cap B] / P[B]$ $P[A \cap B] = P[A/B] \times P[B]$
Indépendance (A est indépendant de B si B n'intervient pas dans la probabilité de A)	$P[A/B] = P[A]$ $P[B \cap A] = P[A] \times P[B]$
Théorème de Bayes (exprime la probabilité de [B/A] en fonction de [A/B])	$P[B/A] = P[A/B] \times P[B] / P[A]$ ou $P[B/A] = P[A/B] \times P[B] / (P[A/B] \times P[B] + P[A/B \text{ barre}] \times (1 - P[B]))$

[370]

LII. *An Essay towards solving a Problem in the Doctrine of Chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S.*

Dear Sir,

Read Dec. 23, 1763. **I** Now send you an essay which I have found among the papers of our deceased friend Mr. Bayes, and which, in my opinion, has great merit, and well deserves to be preserved. Experimental philosophy, you will find, is nearly interested in the subject of it; and on this account there seems to be particular reason for thinking that a communication of it to the Royal Society cannot be improper.

He had, you know, the honour of being a member of that illustrious Society, and was much esteemed by many in it as a very able mathematician. In an introduction which he has writ to this Essay, he says, that his design at first in thinking on the subject of it was, to find out a method by which we might judge concerning the probability that an event has to happen, in given circumstances, upon supposition that we know nothing concerning it but that, under the same circum-

7.2. Extrait du mémoire de 1763

Nous traduisons ci-dessous un extrait de l'édition originale du mémoire de Bayes¹ :

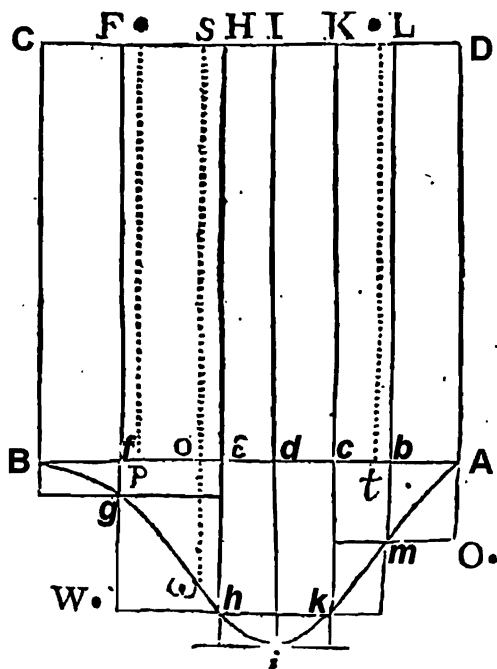
Dans sa Proposition 7, Bayes rappelle :

« Si la probabilité d'un événement est a , et la probabilité de son échec b lors d'un essai unique, la probabilité de sa réalisation p fois en $p + q$ essais est $E a^p b^q$, E étant le coefficient de $a^p b^q$ dans le développement du binôme $(a + b)^{p+q}$. »

« SECTION II (page 385)

POSTULAT 1. Je suppose qu'une table carrée ABCD est réalisée et nivelée de telle sorte que si l'on jette sur sa surface plane l'une ou l'autre des balles O ou W, il y a la même probabilité qu'elle s'immobilise en n'importe quelle partie égale de la surface et qu'elle doit s'immobiliser nécessairement quelque part sur cette surface.

POSTULAT 2. Je suppose que la balle W est jetée la première et que l'on trace, par le point où elle s'immobilise, la ligne os , parallèle à AD et rencontrant CD et AB respectivement en s et en o ; qu'on lance ensuite la balle O ($p + q$), soit n fois; et qu'on désigne son immobilisation entre AD et os après un premier jet par le nom "réalisation de l'événement M à la première épreuve". Ces choses étant supposées, on peut énoncer les lemmes suivants :



1. Il existe par ailleurs une traduction en français du mémoire complet de Bayes, précédée d'une préface de B. Bru et suivie d'une postface de J.-P. Cléro, qui est également le traducteur du texte de Bayes, publiée en 1988 (n° 18 du Cahier d'histoire et de philosophie des sciences, Paris).

LEMME 1. La probabilité pour que le point o tombe entre deux points quelconques de la ligne AB est le rapport de la distance entre les deux points et la ligne AB tout entière.

Soient deux points quelconques de la ligne AB , désignés par f et par b , par lesquels on fait passer fF et bL , parallèles à AD , et coupant CD respectivement en F et en L . Alors, si les rectangles Cf , Fb , LA sont commensurables entre eux, il est possible de diviser chacun d'eux en parties égales ; cette division étant effectuée et la balle W ayant été lancée, la probabilité pour qu'elle s'immobilise quelque part sur l'une quelconque de ces parties égales sera la somme des probabilités qu'elle a de s'immobiliser sur l'une d'elles, parce que toutes les immobilisations sur une partie quelconque du plan AC sont autant d'événements équiprobables ; et cette somme, parce que la probabilité que la balle a de s'immobiliser sur l'une quelconque des parties égales est la même, est la probabilité pour qu'elle s'immobilise sur n'importe quelle partie égale multipliée par le nombre de parties. Par conséquent, la probabilité pour que la balle W s'immobilise quelque part sur Fb est la probabilité qu'elle a de s'immobiliser sur une des parties égales multipliée par le nombre de parties égales contenues en Fb ; et la probabilité qu'elle s'immobilise quelque part sur Cf ou sur LA , autrement dit, qu'elle ne s'immobilise pas sur Fb (car elle doit s'immobiliser quelque part sur AC) est la probabilité qu'elle s'immobilise sur une des parties égales multipliée par le nombre des parties égales dans Cf et dans LA pris ensemble. Il s'ensuit que la probabilité pour qu'elle s'immobilise sur Fb est à la probabilité pour qu'elle ne s'y immobilise pas comme le nombre de parties égales sur Fb au nombre de parties égales sur Cf et LA pris ensemble, ou comme Fb à Cf et LA pris ensemble, ou comme fb à Bf et Ab pris ensemble. Par conséquent, la probabilité qu'elle s'immobilise sur fb est à la probabilité qu'elle ne le fasse pas comme fb est à Bf et Ab pris ensemble. Et la probabilité qu'elle s'immobilise sur Fb est à la probabilité qu'elle s'immobilise sur Fb ajoutée à celle qu'elle ne le fasse pas comme le rapport de fb à AB , ou comme le rapport fb à AB au rapport de AB à AB . Mais la probabilité de tout événement ajoutée à la probabilité de sa non-réalisation est le rapport d'égalité ; par conséquent, la probabilité pour qu'elle s'immobilise sur Fb est au rapport d'égalité comme le rapport de fb à AB au rapport de AB à AB , soit le rapport d'égalité ; et par conséquent, la probabilité pour qu'elle s'immobilise sur Fb est le rapport de fb à AB . Mais, par hypothèse, selon que la balle W tombe sur Fb ou non, le point o se trouvera entre f et b ou pas, et par conséquent la probabilité pour que le point o se trouve entre f et b est le rapport de fb à AB .

De même, si les rectangles Cf , Fb , LA ne sont pas commensurables, la probabilité qui vient d'être mentionnée ne peut être ni plus grande ni plus petite que le rapport de fb à AB ; en effet, s'il est moindre, laissons-le être le rapport de fc à AB , et prenons sur la ligne fb les points p et t de telle sorte que pt soit plus grand que fc , et que les trois lignes Bp , pt , tA soient commensurables (ce qu'on peut évidemment toujours effectuer en divisant AB en parties égales moindres que la moitié de cb et en prenant p et t comme points de division les plus proches de f et de c appartenant à fb). Alors, parce que Bp , pt , tA sont commensurables, les rectangles Cp , Dt et celui qui repose sur la base pt complétant le rectangle verticalement le sont aussi. Par conséquent, d'après ce qui a été dit, la probabilité pour que le point o soit compris entre p et t est le rapport de pt à AB . Mais s'il est situé entre p et t , il doit être situé entre f et b . Par conséquent, la probabilité pour qu'elle soit située entre f et b ne

peut pas être plus petite que le rapport de pt à AB , et donc doit être plus grande que le rapport de fc à AB (puisque pt est plus grand que fc). Et on peut prouver de la même façon que ladite probabilité ne peut pas être plus grande que le rapport de fb à AB ; il doit donc être le même.

LEMME 2. La balle W ayant été lancée, et la ligne os tracée, la probabilité de l'événement M en une seule épreuve est le rapport de Ao à AB .

Car, de la même façon que dans le lemme précédent, la probabilité pour que la balle O , étant lancée, s'immobilise quelque part sur Do ou entre AD et so est le rapport de Ao à Ab . Mais l'immobilisation de la balle O entre AD et so après un seul lancer est la production de l'événement M en une seule épreuve. Par conséquent le lemme est évident.

PROPOSITION 8

Supposons que sur BA l'on construise la figure $BghikmA$ comme suit : la base BA est divisée en deux parties quelconques, telles que Ab et Bb ; une perpendiculaire est tracée au point de division b sur BA et coupe la figure en m ; y , x , r représentent respectivement les rapports de bm , Ab , et Bb à AB et E est le coefficient du terme en $a^p b^q$ quand le binôme $(a + b)^{p+q}$ est développé. La courbe représente $y = E \cdot x^p r^q$. Je dis que, avant que la balle M soit lancée, la probabilité pour que le point o soit situé entre deux points quelconques, f et b , de la ligne AB , ainsi que celle pour que l'événement M se produise p fois et ne se produise pas q fois en $p + q$ essais, est le rapport de $fghikmb$, partie de la figure $BghikmA$ comprise entre les perpendiculaires fg et bm sur la ligne AB , au carré CA de côté AB .

Bayes donne ensuite une démonstration de sa proposition 8 puis le corollaire suivant et sa proposition 9 :

COROLLAIRE : Avant que la balle W soit lancée, la probabilité pour que le point o soit situé quelque part entre A et B , ou quelque part sur la ligne AB , et en outre pour que l'événement se produise p fois et ne se produise pas q fois en $p + q$ épreuves est le rapport de toute la figure AiB à CA . Mais il est certain que le point o sera quelque part sur AB . Par conséquent, avant que la balle W soit lancée, la probabilité que l'événement M se produise p fois et ne se produise pas q fois en $p + q$ épreuves est le rapport de AiB à CA .

PROPOSITION 9

Si, avant de découvrir quoi que ce soit au sujet de l'emplacement du point o , il apparaissait que l'événement M s'est produit p fois et ne s'est pas produit q fois en $p + q$ essais, et que, dans ce cas, je conjecture que le point o se trouve entre deux points quelconques de la ligne AB , comme f et b , et que, par conséquent, la probabilité de l'événement M en un seul essai se trouvait quelque part entre le rapport de Ab à AB et celui de Af à AB : la probabilité que j'aie raison serait le rapport de cette partie de la figure AiB décrite précédemment, comprise entre les perpendiculaires à AB passant par les points f et b , à toute la figure AiB .

Car étant donné deux événements subséquents, le premier étant que le point o se trouve entre f et b ; le second que l'événement M se produise p fois et ne se produise pas q fois en $p + q$ essais ; étant donné que (par le corollaire de la proposition 8) la probabilité originale du second est le rapport de AiB à CA ; il s'ensuit (par la Propo-

sition 5) que si l'on découvre en premier lieu que le second s'est produit et par là que je conjecture que le premier s'est produit aussi, la probabilité que j'aie raison est le rapport de $fghikmb$ à AiB , ce qui était à démontrer.

COROLLAIRE : Les mêmes choses étant supposées, si je conjecture que la probabilité de l'événement M se trouve quelque part entre o et le rapport de Ab à AB , ma chance d'avoir raison est le rapport de Abm à AiB . »

Bayes conclut sa proposition 9, page 392 de l'édition originale du mémoire, par un *scholium* qui est la base de toute l'inférence statistique :

« SCHOLIUM : Par la proposition précédente, il est évident que, dans le cas d'un événement tel que celui que j'appelle ici M , à partir du nombre de fois qu'il se produit et qu'il ne se produit pas, sans en savoir davantage sur son compte, on peut faire une conjecture à l'endroit de sa probabilité et, par les méthodes ordinaires d'évaluation des grandeurs des aires mentionnées ici, voir la chance que cette conjecture soit bonne. Et que la même règle soit adéquate dans le cas d'un événement dont nous ignorons totalement la probabilité avant d'avoir expérimenté à son sujet, cela me semble manifeste [...]

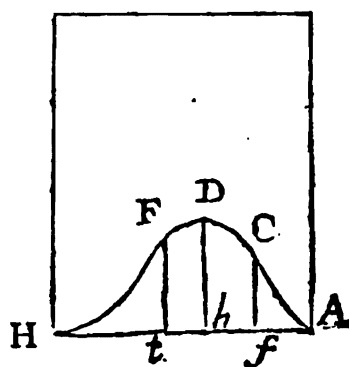
Il s'ensuit par conséquent que je tiendrai pour exact que la règle concernant l'événement M dans la proposition 9 est également la règle que l'on doit utiliser pour tout événement dont nous ignorons la probabilité totalement avant d'avoir effectué des essais et cela semble être la conséquence de la considération suivante : l'on n'a aucune raison de penser, qu'au cours d'un certain nombre d'épreuves, il se produirait un nombre quelconque de fois plutôt qu'un autre. Car, dans ces conditions, on peut raisonner à son sujet exactement comme si sa probabilité avait été préalablement non fixée puis avait été déterminée de telle façon qu'elle ne donne aucune raison de penser qu'au cours d'un certain nombre d'épreuves, l'événement se produirait un nombre quelconque de fois plutôt qu'un autre. Mais c'est exactement ce qui se passe pour l'événement M . Car avant que la balle W soit lancée, ce qui détermine sa probabilité en un seul essai (par le corollaire de la proposition 8), la probabilité que l'événement a de se produire p fois et de ne pas se produire q fois en $p + q = n$ essais, est le rapport de AiB à CA ; ce rapport est le même quand $p + q$, soit n , est donné, quel que soit le nombre p , comme il apparaîtra en calculant la grandeur de AiB par la méthode des fluxions. Par conséquent, avant de découvrir l'emplacement du point o ou le nombre de fois que l'événement M s'est produit en n épreuves, on ne peut avoir de raison de penser qu'il puisse se produire un nombre possible de fois plutôt qu'un autre.

Il s'ensuit par conséquent que je tiendrai pour exact que la règle concernant l'événement M dans la proposition 9 est aussi la règle qu'on doit utiliser pour tout événement dont nous ignorons la probabilité avant d'avoir effectué et observé une épreuve quelconque le concernant. J'appellerai un tel événement : un événement inconnu.

COROLLAIRE : De là, si l'on contracte les ordonnées de la figure AiB dans le rapport de E à 1, ce qui ne crée aucun changement dans la proportion des parties de la figure délimitées par elles, et en appliquant ce qui a été dit de l'événement M à un événement inconnu, nous avons la proposition suivante qui donne les règles pour trouver la probabilité d'un événement à partir du nombre de fois qu'il se produit ou ne se produit pas.

PROPOSITION 10

Soit une figure décrite à partir d'une base quelconque AH et ayant pour équation $y = x^p r^q$; où y, x, r , sont respectivement les rapports (la base étant prise comme dénominateur commun) d'une ordonnée de la figure prise à angle droit sur la base, du segment de base intercepté entre l'ordonnée et l'origine A de la base, et de l'autre segment de la base compris entre l'ordonnée et le point H. Je dis alors que si un événement inconnu est arrivé p fois, et a échoué q fois en $p + q$ essais, et que si vous prenez deux points quelconques f et t sur la base et érigez les ordonnées fC et tF à angle droit, la chance que la probabilité de l'événement soit comprise entre les rapports Af/AH et At/AH , est le rapport $tFCf$, la partie de la figure comprise entre les deux ordonnées, à $ACFH$, la figure entière reposant sur la base AH. Cela est évident par la proposition 9 et le scholium et le corollaire qui suivent cette proposition [...]. »



Les résultats finaux calculés par Bayes (pages 399 et 400) sont les suivants :

« RÈGLE 1

Si rien n'est connu d'un événement, sinon qu'il s'est produit p fois et ne s'est pas produit q fois en $p + q = n$ essais, et si, à partir de là, je conjecture que la probabilité de sa production en un seul essai est comprise entre deux degrés de probabilité X et x , la chance que j'ai d'avoir raison dans ma conjecture est $(n + 1)E$ multiplié par la différence entre la série :

$$\frac{X^{p+1}}{p+1} - \frac{qX^{p+2}}{p+2} + \frac{q(q-1)X^{p+3}}{2(p+3)} - etc$$

et la série :

$$\frac{x^{p+1}}{p+1} - \frac{qx^{p+2}}{p+2} + \frac{q(q-1)x^{p+3}}{2(p+3)} - etc$$

E étant le coefficient de $a^p b^q$ lorsque $(a + b)^n$ est développé.

Voilà la règle que l'on peut utiliser quand q est un petit nombre ; mais si q est grand et p petit, changez partout dans les séries proposées ci-dessus p en q , q en p , x en r ou en $(1 - x)$, et X en $R = (1 - X)$; ce qui ne changera pas la différence entre les deux séries. »

Le texte écrit par Bayes s'achève ici.

Price ajoute qu'une seconde règle est nécessaire lorsque p et q sont de grands nombres, et il écrit :

« RÈGLE 2

Si on ne connaît rien d'un événement sinon qu'il s'est produit p fois et qu'il ne s'est pas produit q fois en $p + q = n$ essais et que de là je conjecture que la probabilité de sa production en un seul essai est comprise entre $(p/n) + z$ et $(p/n) - z$; si $m^2 = n^3 / pq$, $a = p/n$, $b = q/n$, E étant le coefficient de $a^p b^q$ lorsque $(a + b)^n$ est développé et si

$$\Sigma = \frac{(n+1)}{n} \cdot \frac{\sqrt{2pq}}{\sqrt{n}} \cdot E \cdot a^p b^q \text{ multiplié par la série}$$

$$mz - \frac{m^3 z^3}{3} + \frac{n-2}{2n} \cdot \frac{m^5 z^5}{5} - \frac{(n-2)}{2n} \cdot \frac{(n-4)}{3n} \cdot \frac{m^7 z^7}{7} + etc.$$

La probabilité que j'ai d'être dans le vrai est plus grande que :

$$\frac{2\Sigma}{1 + 2Ea^p b^q + \frac{2a^p b^q}{n}}$$

et moindre que

$$\frac{2\Sigma}{1 - 2Ea^p b^q - \frac{2a^p b^q}{n}}$$

et si $p = q$, ma chance est exactement de $2\Sigma \cdot \frac{n}{...}$

Pour que cette règle soit utilisable dans tous les cas, il est uniquement nécessaire de connaître comment trouver dans des limites adéquates la valeur de $Ea^p b^q$ et celle de la série $mz - \frac{m^3 z^3}{3} + etc...$

En rapport avec la règle I, M. Bayes a prouvé que, en supposant que K signifie le rapport du quart d'une circonférence à son rayon, $Ea^p b^q$ sera égal à $\frac{\sqrt{n}}{2\sqrt{Kpq}}$

fois le rapport dont le logarithme hyperbolique est

$$\frac{1}{12} \left[\frac{1}{n} - \frac{1}{p} - \frac{1}{q} \right] - \frac{1}{360} \left[\frac{1}{n^3} - \frac{1}{p^3} - \frac{1}{q^3} \right] + \frac{1}{1260} \left[\frac{1}{n^5} - \frac{1}{p^5} - \frac{1}{q^5} \right] - ...etc...$$

où les coefficients numériques peuvent être trouvés comme suit. Appelons les A, B, C, D, E, etc.

$$\text{Alors } A = \frac{1}{2.2.3} = \frac{1}{4.3}, \quad B = \frac{1}{2.4.5} - \frac{A}{3}, \quad C = \frac{1}{2.6.7} - \frac{10B-A}{5},$$

$$D = \frac{1}{2.8.9} - \frac{35C+21B+A}{7}, \quad E = \frac{1}{2.10.11} - \frac{126C+84D+36B+A}{9},$$

$$F = \frac{1}{2.12.13} - \frac{462D+330C+165E+55B+A}{11}, \text{ etc. où les coefficients de B, C,}$$

D, E, F, etc. dans les valeurs de D, E, F, etc. sont les deuxième, troisième, quatrième, etc. plus hauts coefficients dans le développement de $(a+b)^7$, $(a+b)^9$, $(a+b)^{11}$, etc., en attribuant pour chaque valeur particulière le plus petit de ces coefficients à B, le suivant en grandeur à la lettre la plus éloignée de B, le suivant à C, le suivant à la lettre la plus éloignée moins une, le suivant à D, le suivant à la lettre la plus éloignée moins deux etc.

En ce qui concerne la valeur de la série

$$mz - \frac{m^3 z^3}{3} + \frac{n-2}{2n} \cdot \frac{m^5 z^5}{5} - \frac{(n-2)}{2n} \cdot \frac{(n-4)}{3n} \cdot \frac{m^7 z^7}{7} + etc.$$

il a observé qu'elle pouvait être calculée directement quand mz est plus petit que 1, ou même pas plus grand que $\sqrt{3}$, mais quand mz est beaucoup plus grand, le calcul devient impraticable, dans lequel cas il montre une voie pour trouver facilement deux valeurs presque égales entre lesquelles se trouve la vraie valeur de la série.

Le théorème qu'il donne dans ce but est le suivant :

Soit K, comme précédemment, le rapport du quart de la circonférence d'un cercle à son rayon, et H le rapport dont le logarithme hyperbolique est

$$\frac{2^2-1}{2n} - \frac{2^4-1}{360n^3} + \frac{2^6-1}{1260n^5} - \frac{2^8-1}{1680n^7} + etc.$$

Alors la série $mz - \frac{m^3 z^3}{3} + etc...$ sera plus grande ou plus petite que la série

$$\frac{Hn}{n+1} \cdot \frac{\sqrt{K}}{\sqrt{2}} - \frac{n}{n+2} \cdot \frac{\left(\frac{2m^2 z^2}{n}\right)^{\frac{n}{2}+1}}{2mz} + \frac{n^2}{n+2} \cdot \frac{\left(\frac{2m^2 z^2}{n}\right)^{\frac{n}{2}+2}}{(n+4) \cdot (4m^3 z^3)} - \frac{3n^3}{n+2} \cdot \frac{\left(\frac{2m^2 z^2}{n}\right)^{\frac{n}{2}+3}}{(n+4) \cdot (n+6) \cdot (8m^5 z^5)} +$$

continué jusqu'à un nombre quelconque de termes, selon que le dernier terme est précédé d'un signe positif ou négatif.

En substituant ces valeurs de $Ea^p b^q$ et de

$$mz - \frac{m^3 z^3}{3} + \frac{n-2}{2n} \cdot \frac{m^5 z^5}{5} - \frac{(n-2)}{2n} \cdot \frac{(n-4)}{3n} \cdot \frac{m^7 z^7}{7} + etc.$$

dans la deuxième règle on obtient une troisième règle qui doit être utilisée lorsque mz est d'une quelconque grandeur considérable.

RÈGLE 3

Si on ne connaît rien d'un événement sinon qu'il s'est produit p fois et qu'il ne s'est pas produit q fois en $p+q=n$ essais et que de là je conjecture que la probabilité de sa production en un seul essai est comprise entre $(p/n) + z$ et $(p/n) - z$, ma chance d'avoir raison est plus grande que

$$\frac{b\sqrt{Kpq}}{2\sqrt{Kpq} + bn^{1/2} + bn^{-1/2}} \cdot \left[2H - \frac{\sqrt{2}}{\sqrt{K}} \cdot \frac{n+1}{n+2} \cdot \frac{1}{mz} \cdot \left(1 - \frac{2m^2 z^2}{n} \right)^{\frac{n}{2}+1} \right]$$

et plus petite que

$$\frac{b\sqrt{Kpq}}{2\sqrt{Kpq} - bn^{1/2} - bn^{-1/2}} \text{ multiplié par}$$

$$2H - \frac{\sqrt{2}}{\sqrt{K}} \cdot \frac{n+1}{n+2} \cdot \frac{1}{mz} \cdot \left(1 - \frac{2m^2 z^2}{n}\right)^{\frac{n}{2}+1} + \frac{\sqrt{2}}{\sqrt{K}} \cdot \frac{n}{n+2} \cdot \frac{n+1}{n+4} \cdot \frac{1}{2m^3 z^3} \cdot \left(1 - \frac{2m^2 z^2}{n}\right)^{\frac{n}{2}+2}$$

où m , K , h , et H sont des quantités déjà définies. »

« Un appendice contenant une application des règles précédentes à quelques cas particuliers » est ajouté par Price à fin du mémoire (pages 404 à 418).

Il est à noter que Price a publié¹ un supplément intitulé *A Demonstration of the Second Rule in the Essay towards the Solution of a Problem in the Doctrine of Chances, published in the Philosophical Transactions*, vol. LIII. *Communicated by the Rev. Mr. Richard Price, in a Letter to Mr. John Canton, M. A. F. R. S.*, où il développe une démonstration, donnée par Bayes, de la seconde règle de l'Essai.

7.3. Le commentaire de Laplace

Laplace, dans son *Mémoire sur la probabilité des causes par les événements*², écrit en 1774 :

« L'incertitude des connaissances humaines porte sur les événements ou sur les causes des événements ; si l'on est assuré, par exemple, qu'une urne ne renferme que des billets blancs et noirs dans un rapport donné, et que l'on demande la probabilité qu'en prenant au hasard un de ces billets il sera blanc, l'événement alors est incertain, mais la cause dont dépend la probabilité de son existence, c'est-à-dire le rapport des billets blancs aux noirs, est connue.

Dans le problème suivant : *Une urne étant supposée renfermer un nombre donné de billets blancs et noirs dans un rapport inconnu, si l'on tire un billet et qu'il soit blanc, déterminer la probabilité que le rapport des billets blancs aux noirs est celui de p à q* ; l'événement est connu et la cause inconnue.

On peut ramener à ces deux classes de problèmes tous ceux qui dépendent de la théorie des hasards ; nous ne discuterons ici que ceux de la seconde classe, et pour cela nous établirons le principe suivant :

PRINCIPE – Si un événement peut être produit par un nombre n de causes différentes, les probabilités de l'existence de ces causes prises de l'événement sont entre elles comme les probabilités de l'événement prises de ces causes, et la probabilité de l'existence de chacune d'elles est égale à la probabilité de l'événement prise de cette cause divisée par la somme de toutes les probabilités de l'événement prises de chacune de ces causes. »

Avec des notations modernes, on écrit que si un événement E peut être le résultat d'une cause A ou d'une cause B , le principe de Laplace s'écrit :

$$\frac{P(A/E)}{P(B/E)} = \frac{P(E/A)}{P(E/B)}$$

1. Phil. Trans., vol. LIV, p. 296-325, 1765.

2. Mémoires de l'Académie des sciences (savants étrangers), 6, 1774, 99, p. 621-656.

[296]

LII. A Demonstration of the Second Rule in the Essay towards the Solution of a Problem in the Doctrine of Chances, published in the Philosophical Transactions, Vol. LIII. Communicated by the Rev. Mr. Richard Price, in a Letter to Mr. John Canton, M. A. F. R. S.

Dear Sir,

Nov. 26, 1764.

Read Dec. 6,
1764.

I Send you the following *Supplement to the Essay on a Problem in the Doctrine of Chances*, hoping that you may not think it improper to be communicated to the Royal Society. I should not have troubled you again in this way had I not found that some additions to my former papers were necessary in order to explain some passages in them, and particularly what is hinted in the note at the end of the Appendix. "I have first given the deduction of Mr. Bayes's second rule chiefly in his own words; and then added, as briefly as possible, the demonstrations of several propositions, which seem to improve considerably the solution of the problem, and to throw light on the nature of the curve by the quadrature of which this solution is obtained." Perhaps,

et la probabilité de l'une des causes :

$$P(A / E) = \frac{P(E / A)}{P(E / A) + P(E / B)} .$$

Un excellent commentaire du très important *Essai* de Bayes a été fait par Dale¹ :

« Qu'est-ce qu'un lecteur peut espérer trouver dans cet *Essai*. En ce qui concerne les probabilités il va rechercher une forme ou une autre de ce qui est devenu connu sous le nom de théorème de Bayes, et cette recherche sera satisfaite. En plus, il trouvera une discussion claire de la distribution binomiale et s'il cherche plus profondément il trouvera [...] la première énonciation d'un résultat de logique probabiliste impliquant une probabilité conditionnelle. L'*Essai* intéressera les mathématiciens pour l'évaluation de la fonction Bêta incomplète. Notons également l'utilisation d'approximations faites ici et dans le Supplément dû à Bayes et à Price, et l'attention portée aux erreurs comportées par ces approximations. L'*Essai*, donc, principalement et sans doute justement apprécié pour la solution du problème posé par Bayes, peut aussi l'être pour sa contribution aux mathématiques pures. »

7.4. Biographie de Bayes (1701-1761)

Le père de Thomas Bayes, Joshua Bayes (1671-1746), a été l'un des six premiers pasteurs non-conformistes² à être ordonné (en 1694) en Angleterre. Il se marie le 23 octobre 1700 avec Anne Carpenter (~ 1675-1733). Le couple s'installe à Londres et c'est là que leur premier enfant, Thomas, notre futur mathématicien naît. Sa date de naissance exacte n'est pas connue, mais une étude très détaillée a été faite par Stephen Stigler dont la conclusion est que la date la plus probable est le 10 septembre 1701. D'autres enfants viendront et Thomas aura trois frères et trois sœurs.

Étant le fils d'un Pasteur non-conformiste l'accès à l'école lui était très difficile et c'est chez lui qu'il a reçu sa première éducation. Certains de ses biographes indiquent qu'il aurait eu de Moivre comme précepteur ; cela n'est pas impossible car ce dernier donnait alors des cours à Londres, mais c'est toutefois assez improbable. Il aurait ensuite suivi les cours de la *Fund Academy* à Tenter Alley qui était le seul endroit de la région à ouvrir ses portes à des élèves non anglicans. Ce qui est sûr, c'est que Thomas s'est vite révélé excellent en littérature grecque et latine.

En 1719, Thomas Bayes s'inscrit à l'université d'Édimbourg où il étudie la logique et la théologie. Le choix de cette université écossaise était presque impératif dans la mesure où les universités d'Oxford ou de Cambridge n'autorisaient pas l'inscription d'étudiants non-conformistes. Il est probable

1. A. I. Dale, *Most honourable remembrance: The life and work of Thomas Bayes* (New York-Berlin-Heidelberg, 2003).

2. Un pasteur non-conformiste est un pasteur protestant qui ne suit pas la religion anglicane.

que c'est à Édimbourg que Bayes a commencé à étudier les mathématiques, sans doute avec James Gregory¹, qui avait succédé à son frère David comme professeur de mathématiques à l'université d'Édimbourg. Les études de Thomas se poursuivent jusqu'en 1722. Il quitte l'université avec un diplôme de Master, mais, au plan religieux il n'est pas encore ordonné pasteur et reste novice. Il n'existe pratiquement aucune information sur la vie de Bayes entre 1722 et 1728. On le trouve sur une liste de pasteurs presbytériens à Londres en 1728. Il assiste alors son père, à l'église presbytérienne Leather Lane de Londres, jusque dans les années 1733-1734. Sa première publication théologique, *Divine benevolence*, date de cette époque (1731). Il devient alors pasteur dans la chapelle presbytérienne Mount Sion de Tunbridge Wells, petite ville touristique et thermale du Kent très appréciée à l'époque, située à une cinquantaine de kilomètres au sud-est de Londres. Il restera dans cette ville jusqu'à la fin de sa vie. En 1746, William Whiston, le successeur de Newton à la chaire lucasienne de Cambridge, passe par Tunbridge Wells et rend visite à Thomas Bayes.

En 1742, Thomas Bayes est élu membre de la *Royal Society*. Conformément à la procédure de l'époque, il devait assister, avant l'élection, à une réunion en présence de son « parrain », en l'occurrence John Belchier, chirurgien éminent à l'hôpital de Londres. Cette réunion s'est tenue le 25 mars 1742 et deux semaines plus tard, le 8 avril, Bayes devint effectivement membre de la *Royal Society*. On peut lire sur son « certificat » d'admission : « Le révérend Thomas Bayes, de Tunbridge Wells, désirant avoir l'honneur d'être élu dans notre Société, nous le proposons et le recommandons en tant que Gentleman dont le mérite est connu, habile en géométrie et dans toutes les parties des mathématiques et des sciences physiques, tout à fait qualifié pour être un membre digne de notre Société. »

Quels étaient les travaux de Bayes qui lui avaient valu une telle réputation ? Il est évident qu'il s'agit de l'*Introduction to the doctrines of fluxions* publiée en 1736 qui répondait au célèbre ouvrage du métaphysicien Berkeley *The Analyst, or a discourse addressed to an Infidel Mathematician* dans lequel Berkeley jugeait les fondements logiques du calcul des fluxions de Newton fort douteux. De fait, le mémoire de Bayes reste l'une des meilleures réponses formulées aux critiques de Berkeley.

1. Ce James Gregory est le neveu du célèbre James Gregory (1638-1675). Ce dernier, astronome et mathématicien, a été le premier à réaliser un télescope à miroir utilisable (1663) et le premier professeur de mathématiques à l'université d'Édimbourg. C'est David Gregory, autre neveu du grand James, et également célèbre, qui lui a succédé à ce poste. Mathématicien et astronome comme son oncle, David Gregory contribuera en effet à répandre en Angleterre les travaux de Newton. James (II) Gregory, qui enseigna à Bayes, était le successeur de David.

À la mort de son père, en 1746, Bayes hérite d'une fortune appréciable ainsi que d'une riche bibliothèque. C'est peut-être l'une des raisons pour lesquelles, dès 1749, il a désiré cesser ses fonctions de pasteur. Sans doute faute de successeur, il n'aura gain de cause qu'en 1752. Il continue toutefois à vivre dans cette ville de Tunbridge Wells qu'il aime tant. Durant ces années, à partir de 1747, Bayes se penche également sur la façon d'obtenir une démonstration correcte de la formule de Stirling. Il n'a toutefois jamais publié son résultat.

Il est probable que l'intérêt de Thomas Bayes pour les probabilités est apparu aux alentours des années 1755 après qu'il ait eu à analyser (à « reviewer ») un article de Thomas Simpson¹. Ce dernier avait examiné un cas spécial de la loi de grands nombres : la moyenne d'un ensemble d'observations est une meilleure estimation d'un paramètre que celle donnée par une seule observation. Dans une lettre à John Canton, Bayes montre que ceci ne peut pas être vrai en présence de biais de mesure.

En dehors des mathématiques (séries infinies, probabilités, mais aussi trigonométrie, géométrie, équations différentielles) Bayes a eu plusieurs autres domaines d'intérêt scientifiques, en particulier les sciences physiques (électricité, optique, acoustique) et la mécanique céleste.

La renommée de Bayes reste évidemment liée à son célèbre *Essay towards solving a problem in the doctrine of chances* qui est la première tentative pour établir les fondements de l'inférence statistique, qui a été publié, à titre posthume en 1763 par son ami Richard Price.

À partir de 1755, la santé de Bayes se dégrade. Il souffre de rhumatisme et son état physique passe par des hauts et des bas. Sentant sa fin proche il fait part de ses dernières volontés le 12 décembre 1760 et s'éteint quatre mois plus tard, le 7 avril 1761 à Tunbridge Wells, à l'âge de cinquante-neuf ans. Il a été enterré dans le caveau familial de Bunhill Fields, près de Moorgate. Le caveau a été entretenu et l'on peut y lire maintenant : « En reconnaissance des travaux importants de Thomas Bayes en probabilité, ce caveau a été restauré en 1969 grâce aux contributions reçues des statisticiens du monde entier ».

Le cas de Thomas Bayes est assez unique dans l'histoire des sciences. À son époque on a fait très peu de cas de sa vie et de son œuvre. Ce n'est qu'au XX^e siècle que ses travaux ont été réellement mis en valeur. La première biographie sur Thomas Bayes n'a été rédigée qu'en 1958, par G.A. Barnard (1915-2002). Quant à l'adjectif *bayésien*, il n'a été introduit qu'en 1950 par Fisher.

1. Cf. notre chapitre 10, § 10.2.

7.5. Ouvrages de Thomas Bayes

- *Divine Benevolence: Or, An Attempt to prove that the Principal End of the Divine Providence and Government is the Happiness of his Creatures*, John Noon, London, 1731. (publication anonyme).
- *An Introduction to the Doctrine of Fluxions, and a Defence of the Mathematicians against the objections of the Author of The Analyst*, John Noon, London, 1736 (publication anonyme).
- *A letter from the late Reverend Mr. Thomas Bayes, F. R. S. to John Canton, M. A. and F. R. S.*, Philos. Trans. R. Soc., London, 53, 269-271 (1763).
- *An essay towards solving a problem in the doctrine of chances*, Philos. Trans. R. Soc., London, 53, 370-418 (1763).
- *A demonstration of the second rule in the Essay towards the Solution of a Problem in the Doctrine of Chances, published in the Philosophical Transactions, vol. 53. Communicated by the Rev. Mr. Richard Price, in a letter to Mr. John Canton, M. A., F. R. S.*, Philos. Trans. R. Soc., London, 54, 296-325 (1764).

7.6. Références

- G. A. Barnard, *Thomas Bayes, a biographical note*, Biometrika, 45, 293-295, 1958.
- I. Hacking, *Thomas Bayes*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 1, p. 531-532. 1970.
- P. Dupont, *Un joyau dans l'histoire des sciences : le mémoire de Thomas Bayes de 1763*, Rend. Sem. Mat. Univ. Politec., Torino, 37 (3), 1979, p. 105-138.
- S. M. Stigler, *The history of statistics*, Belknap, Cambridge, 1986.
- A. Hald, *Evaluations of the Beta probability integral by Bayes and Price*, AHES, 41, 139-156, 1990.
- A. I. Dale, *A history of inverse probability*, Springer-Verlag, 1991.
- A. Hald, *A History of mathematical statistics from 1750 to 1930*, chapitre 8, Wiley, 1998.
- A. I. Dale, *Most honourable remembrance: The life and work of Thomas Bayes*, New York-Berlin-Heidelberg, 2003.
- O. B. Sheynin, *On the history of Bayes's theorem*, Math. Scientist, vol. 28, p. 37-42, 2003.

Buffon

Buffon est entré à l'Académie des sciences à l'âge de vingt-sept ans, grâce à ses travaux sur les probabilités. Après des études de mathématiques, de médecine et de botanique, Buffon présenta en 1733 deux mémoires, l'un sur le calcul des probabilités *Solutions de problèmes sur le jeu du Franc-Carreau* et l'autre de *Fine mécanique* qui retinrent l'attention des commissaires de l'Académie des sciences.

8.1 Le jeu du franc-carreau

Les *Solutions de problèmes sur le jeu du Franc-Carreau* ont ensuite été publiées¹ en 1777 dans l'*Essai d'arithmétique morale*. On y trouve le problème bien connu de l'aiguille.

Voici un extrait du texte de Buffon :

« L'Analyse est le seul instrument dont on se soit servi jusqu'à ce jour dans la science des probabilités, pour déterminer et fixer les rapports du hasard ; la Géométrie paraissait peu propre à un ouvrage aussi délié ; cependant si l'on y regarde de près, il sera facile de reconnaître que cet avantage de l'Analyse sur la Géométrie est tout à fait accidentel, et que le hasard selon qu'il est modifié et conditionné, se trouve du ressort de la géométrie aussi bien que celui de l'analyse ; pour s'en assurer, il suffira de faire attention que les jeux et les questions de conjecture ne roulent ordinairement que sur les rapports de quantités discrètes ; l'esprit humain plus familier avec les nombres qu'avec les mesures de l'étendue les a toujours préférés ; les jeux en sont une preuve, car leurs lois sont une arithmétique continue ; pour mettre donc la Géométrie en possession de ses droits sur la science du hasard, il ne s'agit que d'inventer des jeux qui roulent sur l'étendue et sur ses rapports, ou calculer le petit nombre de ceux de cette nature qui sont déjà trouvés ; le jeu du franc-carreau peut nous servir d'exemple : voici ses conditions qui sont fort simples.

Dans une chambre parquetée ou pavée de carreaux égaux, d'une figure quelconque, on jette en l'air un écu ; l'un des joueurs parie que cet écu après sa chute se trouvera à franc-carreau, c'est-à-dire sur un seul carreau ; le second parie que cet écu se trouvera sur deux carreaux, c'est-à-dire qu'il couvrira un des joints qui les séparent ; un troisième joueur parie que l'écu se trouvera sur deux joints, un quatrième parie

1. Buffon G., « Essai d'arithmétique morale », in *Histoire naturelle, générale et particulière, servant de suite à l'Histoire naturelle de l'Homme*, Supplément, tome quatrième, Paris, Imprimerie Royale, 1777, p. 46-148.

que l'écu se trouvera sur trois, quatre, ou six joints : on demande le sort de chacun de ces joueurs.

Je cherche d'abord le sort du premier joueur et du second : pour le trouver, j'inscris dans l'un des carreaux une figure semblable, éloignée des côtés du carreau, de la longueur du demi-diamètre de l'écu ; le sort du premier joueur sera à celui du second comme la superficie de la couronne circonscrite est à la superficie de la figure inscrite : cela peut se démontrer aisément, car tant que le centre de l'écu est dans la figure inscrite, cet écu ne peut être que sur un seul carreau, puisque par construction cette figure inscrite est partout éloignée du contour du carreau, d'une distance égale au rayon de l'écu ; et, au contraire, dès que le centre de l'écu tombe au dehors de la figure inscrite, l'écu est nécessairement sur deux ou plusieurs carreaux, puisque alors son rayon est plus grand que la distance du contour de cette figure inscrite ou contour du carreau ; or, tous les points où peut tomber ce centre de l'écu sont représentés dans le premier cas par la superficie de la couronne qui fait le reste du carreau ; donc le sort du premier joueur est au sort du second, comme cette première superficie est à la seconde : ainsi pour rendre égal le sort de ces deux joueurs, il faut que la superficie de la figure inscrite soit égale à celle de la couronne, ou, ce qui est la même chose, qu'elle soit la moitié de la surface totale du carreau.

Je me suis amusé à en faire le calcul, et j'ai trouvé que pour jouer à jeu égal sur des carreaux carrés, le côté du carreau devait être au diamètre de l'écu, comme

$$\frac{1}{1 - \sqrt{\frac{1}{2}}} ; \text{ c'est-à-dire à peu près trois et demi fois plus grand que le diamètre de}$$

la pièce avec laquelle on joue.

Pour jouer sur des carreaux triangulaires équilatéraux, le côté du carreau doit être

$$\text{au diamètre de la pièce, comme } 1 / \left(\frac{\frac{1}{2}\sqrt{3}}{3 + 3\sqrt{\frac{1}{2}}} \right), \text{ c'est-à-dire presque six fois plus grand que le diamètre de la pièce.}$$

Sur des carreaux en losange, le côté du carreau doit être au diamètre de la pièce,

$$\text{comme } 1 / \left(\frac{\frac{1}{2}\sqrt{3}}{2 + \sqrt{2}} \right), \text{ c'est-à-dire presque quatre fois plus grand.}$$

Enfin, sur des carreaux hexagones, le côté du carreau doit être au diamètre de la

$$\text{pièce, comme } 1 / \left(\frac{\frac{1}{2}\sqrt{3}}{1 + \sqrt{\frac{1}{2}}} \right), \text{ c'est-à-dire presque le double.}$$

Je n'ai pas fait le calcul pour d'autres figures, parce que celles-ci sont les seules dont on puisse remplir un espace sans y laisser des intervalles d'autres figures ; et je n'ai pas cru qu'il fût nécessaire d'avertir que les joints des carreaux ayant quelque largeur, ils donnent de l'avantage au joueur qui parie pour le joint, et que par conséquent l'on fera bien, pour rendre le jeu encore plus égal, de donner aux carreaux carrés un peu plus de trois et demie fois, aux triangulaires six fois, aux

losanges quatre fois, et aux hexagones deux fois la longueur du diamètre de la pièce avec laquelle on joue.

Je cherche maintenant le sort du troisième joueur qui parie que l'écu se trouvera sur deux joints ; et, pour le trouver, j'inscris dans l'un des carreaux une figure semblable, comme je l'ai déjà fait ; ensuite je prolonge les côtés de cette figure inscrite jusqu'à ce qu'ils rencontrent ceux du carreau, le sort du troisième joueur sera à celui de son adversaire, comme la somme des espaces compris entre le prolongement de ces lignes et les côtés du carreau est au reste de la surface du carreau. Ceci n'a besoin, pour être pleinement démontré, que d'être bien entendu.

J'ai fait aussi le calcul de ce cas, et j'ai trouvé que, pour jouer à jeu égal sur des carreaux carrés, le côté du carreau doit être au diamètre de la pièce, comme $1 / \left(\frac{1}{\sqrt{2}} \right)$ c'est-à-dire plus grand d'un peu moins d'un tiers.

Sur des carreaux triangulaires équilatéraux, le côté du carreau doit être au diamètre de la pièce, comme $1/(1/2)$, c'est-à-dire le double.[...].

J'omets ici la solution de plusieurs autres cas, comme lorsque l'un des joueurs parie que l'écu ne tombera que sur un joint ou sur deux, sur trois, etc. Ils n'ont rien de plus difficile que les précédents ; et d'ailleurs on joue rarement ce jeu avec d'autres conditions que celles dont nous avons fait mention.

Mais si au lieu de jeter en l'air une pièce ronde, comme un écu, on jetait une pièce d'une autre figure comme une pistole d'Espagne carrée, ou une aiguille, une baguette, etc., le problème demanderait un peu plus de géométrie, quoiqu'en général il fût toujours possible d'en donner la solution par des comparaisons d'espaces, comme nous allons le démontrer.

Je suppose que dans une chambre, dont le parquet est simplement divisé par des joints parallèles, on jette en l'air une baguette, et que l'un des joueurs parie que la baguette ne croisera aucune des parallèles du parquet, et que l'autre au contraire parie que la baguette croisera quelques-unes de ces parallèles ; on demande le sort de ces deux joueurs. On peut jouer ce jeu sur un damier avec une aiguille à coudre ou une épingle sans tête. »

Aujourd'hui on énonce la solution comme suit :

Soit c la largeur des planches, et d la longueur de l'aiguille. On suppose $d \leq c$ que. On note r la distance aléatoire entre le centre de l'aiguille et la jointure de deux planches la plus proche, et θ l'angle aléatoire de l'aiguille avec la jointure. Il est évident que l'aiguille intersecte une jointure si $\frac{d}{2} \sin \theta > r$.

Si on suppose, que les jets d'aiguilles sont faits de telle sorte que les lois de r et θ soient uniformes sur respectivement $[0, c/2]$ et $[0, \pi/2]$, la probabilité

de tomber sur une rainure du parquet est donc de
$$\frac{\int_0^{\pi/2} \frac{d}{2} \sin \theta d\theta}{\frac{\pi}{2} \cdot \frac{c}{2}} = \frac{2d}{\pi c}$$

En estimant cette probabilité sur un grand nombre d'expériences, on peut déduire une estimation du nombre π .

8.2. L'Histoire Naturelle

Dans son ouvrage célèbre intitulé *Histoire Naturelle*, Buffon se sert, également, de notions élémentaires de probabilité pour étayer ses théories concernant la formation des planètes du système solaire. Selon son hypothèse les six planètes alors connues, Mercure, Vénus, Terre, Mars, Jupiter, Saturne sont le produit d'une collision entre le Soleil et une comète. Buffon constate que les planètes tournent dans le même sens autour du Soleil et presque dans le même plan. Cela suppose nécessairement, selon lui quelque chose de commun dans leur mouvement d'impulsion, et fait soupçonner qu'il leur a été communiqué par une seule et même cause. Buffon écrit¹ :

« Cette idée sur la cause du mouvement d'impulsion des planètes paraîtra moins hasardée lorsqu'on rassemblera toutes les analogies qui y ont rapport, et qu'on voudra se donner la peine d'estimer les probabilités. La première est cette direction commune de leur mouvement d'impulsion, qui fait que les six planètes vont toutes d'occident en orient ; et il y a déjà 4 à parier contre un qu'elles n'auraient pas eu ce mouvement dans le même sens, si la même cause ne l'avait pas produit, ce qu'il est aisé de prouver par la doctrine des hasards.

Cette probabilité augmentera prodigieusement par la seconde analogie, qui est que l'inclinaison des orbites n'excède pas 7 degrés et demi ; car en comparant les espaces on trouve qu'il y a 24 contre un pour que deux planètes se trouvent dans des plans plus éloignés et par conséquent 24^5 ou 7 962 624 à parier contre un, que ce n'est pas par hasard qu'elles se trouvent toutes six ainsi placées et renfermées dans l'espace de 7 degrés et demi, ou, ce qui revient au même, il y a cette probabilité qu'elles ont quelque chose de commun dans le mouvement qui leur a donné cette position. Mais que peut-il y avoir de commun dans l'impression d'un mouvement d'impulsion, si ce n'est la force et la direction des corps qui le communiquent ? On peut donc conclure avec une très grande vraisemblance que les planètes ont reçu leur mouvement d'impulsion par un seul coup. »

Le raisonnement de Buffon est évident : La probabilité que les six planètes aient un mouvement dans la même direction étant faible, on peut supposer que ce fait n'est pas dû au hasard mais résulte d'une cause à déterminer. On notera que Daniel Bernoulli avait déjà abordé ce même problème en 1735 dans un mémoire portant le titre *Recherches physiques et astronomiques sur le problème proposé pour la seconde fois par l'Académie Royale des sciences de Paris : Quelle est la cause physique de l'inclinaison des plans des orbites des planètes par rapport au plan de l'équateur de la révolution du soleil autour de son axe ; Et d'où vient que les inclinaisons de ces orbites sont différentes entre elles.*²

1. *Histoire Naturelle*, tome 1, Imprimerie Royale, 1749, p. 131.

2. *Recueil des pièces qui ont remporté le Prix de l'Académie Royale des sciences*, 3, p. 93-122, 1735. Le même mémoire rédigé en latin occupe les pages 125-144.

8.3. Des probabilités de la durée de la vie

Buffon avait rédigé en 1745-1746 un rapport académique sur le travail de Deparcieux, *Sur les probabilités de la vie*. Il inséra en 1749, dans le tome II de l'Histoire naturelle générale et particulière un chapitre d'une quinzaine de pages intitulé *De la vieillesse et de la mort* dans lequel il analysait une table de mortalité, élaborée par Dupré de Saint-Maur (1695-1774) et résultant de la compilation des registres de trois paroisses parisiennes et de douze paroisses rurales, portant sur environ 24 000 décès. À partir de ces données, Buffon calculait le temps restant probablement à vivre aux différents âges de la vie, mais on ignore sa technique d'évaluation. Le choix de cette table, qui, par ailleurs ne distingue pas les décès selon le sexe, est justifié par Buffon en arguant que les autres tables disponibles sont spécifiques à certaines catégories de personnes, les citadins pour les tables de Graunt, Halley, Kerseboom et Simpson, ou les ecclésiastiques et les rentiers pour Deparcieux.

En 1777, Condorcet publie un mémoire intitulé *Des probabilités de la durée de la vie*¹, beaucoup plus développe, où il donne, pour chacun des âges compris entre la naissance et 100 ans, la probabilité de vivre encore n années, n étant tabulé tous les cinq ans à partir de l'âge de deux ans. Les calculs utilisent, pour l'âge x , le rapport du nombre de survivants à l'âge $x + n$ par rapport au nombre de survivants à l'âge x . L'ensemble de ces résultats occupe 120 pages. Condorcet en déduit, par exemple que les deux tiers des humains périssent, selon ses calculs, avant l'âge de 39 ans. D'autres tables et observations sur le nombre de naissances, mariages et décès relatifs à la ville de Paris, pour la période 1709-1766, suivent.

8.4. L'Essai d'arithmétique morale

Buffon définit le but de son ouvrage², paru en 1777 mais composé selon Gouraud³, vers 1760, comme suit :

« La mesure des choses incertaines fait ici mon objet, je vais tâcher de donner quelques règles pour estimer les rapports de vraisemblance, les degrés de probabilité, le poids des témoignages, l'influence des hasards, l'inconvénient des risques, et juger en même temps de la valeur réelle de nos craintes et de nos espérances. »

-
1. *Histoire naturelle, générale et particulière, servant de suite à l'Histoire naturelle de l'Homme*, Supplément, tome quatrième, Paris, Imprimerie Royale, 1777, p. 149-264.
 2. « Essai d'arithmétique morale », in *Histoire naturelle, générale et particulière, servant de suite à l'Histoire naturelle de l'Homme*, Supplément, tome quatrième, Paris, Imprimerie Royale, 1777, p. 46-148.
 3. Gouraud, C., *Histoire de calcul des probabilités*, Paris, Durand, 1848, p. 54.

Ernest Coumet¹ estime que l'ensemble de l'ouvrage constitue un essai sur les mesures, les sections I à XXII étudiant la mesure de l'incertitude, les sections XXIII à XXXIII les mesures en mathématiques et les deux dernières sections, celles des abstractions.

Buffon examine et commente donc, dans ce mémoire, la relation qui existe entre probabilité, évidence et certitude et cherche à mesurer numériquement la certitude physique et la certitude morale.

« Dans les sciences physiques » énonce-t-il, « l'évidence est remplacée par la certitude ; l'évidence n'est pas susceptible de mesure, parce qu'elle n'a qu'une seule propriété absolue, qui est la négation nette ou l'affirmation de la chose qu'elle démontre ; mais la certitude n'étant jamais d'un positif absolu, a des rapports que l'on doit comparer et dont on peut estimer la mesure. La certitude physique, c'est-à-dire la certitude de toutes la plus certaine, n'est néanmoins que la probabilité, presque infinie qu'un effet, un événement qui n'a jamais manqué d'arriver, arrivera encore une fois ; par exemple, puisque le soleil s'est toujours levé, il est dès lors physiquement certain qu'il se lèvera demain ; une raison pour être, c'est d'avoir été, mais une raison pour cesser d'être, c'est d'avoir commencé d'être ; et par conséquent l'on ne peut pas dire qu'il soit également certain que le soleil se lèvera toujours, à moins de lui supposer une éternité antécédente, égale à la perpétuité subséquente, autrement il finira puisqu'il a commencé. Car nous ne devons juger de l'avenir que par la vue du passé ; dès qu'une chose a toujours été, ou s'est toujours faite de la même façon, nous devons être assurés qu'elle sera ou se fera toujours, de cette même façon : par *toujours*, j'entends un très long temps, et non pas une éternité absolue. »

Pour Buffon, la connaissance s'acquiert par l'observation ou par l'analogie. L'expérience résulte de l'itération d'une même observation, et l'analogie peut être définie comme la somme des rapports avec les connaissances déjà acquises. Il écrit :

« Mais si l'expérience est la base de nos connaissances physiques et morales, l'analogie en est le premier instrument ; lorsque nous voyons qu'une chose arrive constamment d'une certaine façon, nous sommes assurés par notre expérience qu'elle arrivera encore de la même façon ; et lorsque l'on nous rapporte qu'une chose est arrivée de telle ou telle manière, si ces faits ont de l'analogie avec les autres faits que nous connaissons par nous-mêmes, dès lors nous les croyons ; au contraire, si le fait n'a aucune analogie avec les effets ordinaires, c'est-à-dire, avec les choses qui nous sont connues, nous devons en douter ; et s'il est directement opposé à ce que nous connaissons, nous n'hésitons pas à le nier. »

La vérité physique, pour Buffon, « n'est qu'une probabilité, mais une probabilité si grande, qu'elle équivaut à une certitude. » La certitude morale, qui est une

1. Séminaire d'*Histoire du calcul des probabilités et de la statistique* de l'École des hautes études en sciences sociales, mars 1989.

certitude par analogie, diffère de la certitude physique et de l'évidence mathématique et elle tient « le milieu entre le doute et la certitude physique. »

Passant au choix d'une unité de mesure pour estimer les certitudes physique et morale par le calcul des probabilités, Buffon, en ce qui concerne la dernière, affirme simplement qu'une probabilité de $1/10\ 000$ est assimilable à une probabilité nulle. Voici son raisonnement :

« En consultant les Tables de mortalité, je vois qu'on en peut déduire, qu'il n'y a que dix mille cent quatre-vingt-neuf à parier contre un, qu'un homme de cinquante-six ans vivra plus d'un jour. Or comme tout homme de cet âge, où la raison a acquis toute sa maturité et l'expérience toute sa force, n'a néanmoins nulle crainte de la mort dans les vingt-quatre heures, quoiqu'il n'y ait que dix mille cent quatre-vingt-neuf à parier contre un, qu'il ne mourra pas dans ce court intervalle de temps ; j'en conclus, que toute probabilité égale ou plus petite, doit être regardée comme nulle, et que toute crainte ou toute espérance qui se trouve au-dessous de dix mille, ne doit ni nous affecter, ni même nous occuper un seul instant le cœur ou la tête. »

Ce principe des probabilités négligeables ne conduit pas, cependant, Buffon à une assimilation du probable au certain, et il affirme que toute connaissance du réel ne peut être que probable.

Ayant construit la valeur de la certitude morale sur la crainte de la mort, Buffon procède ensuite à la comparaison de la crainte et de l'espérance et déclare que la perte affecte l'homme toujours plus que le gain. Il adopte, de fait, le concept d'utilité marginale de Daniel Bernoulli. Nous avons, dans le chapitre consacré à ce dernier, cité le texte de Buffon relatif au problème de Saint-Petersbourg. Ce paradoxe fut étudié non seulement par Daniel Bernoulli, mais aussi par Cramer, Buffon, d'Alembert et beaucoup d'autres, car il soulevait un problème d'espérance mathématique face, en quelque sorte, à une espérance morale.

La notion d'utilité marginale a conduit Buffon à condamner tous les jeux de hasard. Voici son argumentation :

« C'est au jeu en général, au jeu le plus égal, et par conséquent le plus honnête que je trouve une essence vicieuse, je comprends même sous le nom de jeu, toutes les conventions, tous les paris où l'on met au hasard une partie de son bien pour obtenir une pareille partie du bien d'autrui ; et je dis qu'en général le jeu est un pacte mal entendu, un contrat désavantageux aux deux parties, dont l'effet est de rendre la perte toujours plus grande que le gain ; et d'ôter au bien pour ajouter au mal. La démonstration en est aussi aisée qu'évidente. Prenons deux hommes de fortune égale, qui, par exemple aient chacun cent mille livres de bien, et supposons que ces deux hommes jouent en un ou plusieurs coups de dés cinquante mille livres, c'est-à-dire la moitié de leur bien ; il est certain que celui qui gagne, n'augmente son bien que d'un tiers, et que celui qui perd, diminue le sien de moitié, car chacun d'eux avait cent mille livres avant le jeu ; mais après l'événement du jeu, l'un aura cent cinquante mille livres, c'est-à-dire un tiers de plus qu'il n'avait, et l'autre

n'a plus que cinquante mille livres, c'est-à-dire moitié moins qu'il n'avait ; donc la perte est d'une sixième partie plus grande que le gain ; car il y a cette différence entre le tiers et la moitié ; donc la convention est nuisible à tous deux, et par conséquent essentiellement vicieuse.

Ce raisonnement n'est point captieux, il est vrai et exact, car quoique l'un des joueurs n'ait perdu précisément que ce que l'autre a gagné, cette égalité numérique de la somme, n'empêche pas l'inégalité vraie de la perte et du gain n'est qu'apparente et l'inégalité réelle [...] »

Philosophe, naturaliste, mathématicien, Buffon a abordé dans ses ouvrages non seulement le calcul des probabilités, mais également l'astronomie, la physique, la physiologie, et l'histoire naturelle. Ses œuvres complètes ont été éditées à Paris en 1884-1885, en 14 volumes, par J.-L. Lanessan.

8.5. Biographie Georges-Louis Le Clerc, Comte de Buffon (1707-1788)

Georges-Louis Le Clerc, futur comte de Buffon¹, est né à Montbard en 1707, dans une riche famille de la noblesse de robe. Son père, Benjamin Le Clerc est président du grenier à sel de Montbard, et sa mère Anne-Christine Marlin élèvera les cinq enfants du couple (Georges-Louis est l'aîné). Son père hérite de la fortune d'un parent en 1717 et achète les propriétés de la seigneurie de Buffon, située à six kilomètres de Montbard.

Georges-Louis fait des études secondaires dans un collège de jésuites à Dijon, étudie le droit à l'université de cette même ville et obtient sa licence en 1726. Préférant toutefois les sciences, il part étudier les mathématiques et la botanique à Angers en 1728. Là, il se plonge dans les mathématiques, lit Newton et suit des cours de médecine. En 1732, il s'installe à Paris et commence à signer *Buffon*. Ses premiers travaux portent sur les mathématiques, son domaine de prédilection, et il présente en 1733 un mémoire à l'Académie des sciences, dont Maupertuis et Clairaut font un compte rendu élogieux. Ce mémoire *Sur le jeu du franc-carreau* introduit pour la première fois le calcul différentiel et le calcul intégral en probabilité. Par ailleurs il traduit de l'anglais plusieurs ouvrages de mathématiques de Newton et de botanique de Hales. Il fait la connaissance de Voltaire et d'autres intellectuels, et entre à l'Académie des sciences en 1734, à l'âge de 26 ans au poste d'adjoint dans la section mécanique. Il s'intéresse alors à tous les domaines de la nature : botanique, biologie, chimie, géologie, utilise les mathématiques comme un outil, et effectue des recherches sur les organes reproducteurs des animaux. Anglophile, il correspond abondamment avec plusieurs savants, séjourne à Londres en 1738 et se fait élire à la *Royal Society* en 1739. Nommé Intendant du Jardin du roi² en 1739, il va l'agrandir et en enrichir les collections, tout en

1. Il devient comte de Buffon en 1773.

2. Aujourd'hui et depuis la révolution Jardin des Plantes.

travaillant à son *Histoire naturelle* œuvre à laquelle il va désormais consacrer une grande partie de son temps. Il connaît la gloire dès la parution des trois premiers volumes en 1749. L'Académie française lui ouvre ses portes en 1753 et il sera également membre de toutes les grandes académies européennes.

Buffon se marie en 1752, à l'âge de quarante-cinq ans, à Marie-Françoise de Saint-Belin Malain, jeune femme de noblesse ruinée. Celle-ci voue une grande affection à son mari qui l'a arrachée au couvent, même s'il n'est pas d'une extrême fidélité, mais elle meurt en 1769.

En cette même année 1752, Buffon vérifie les hypothèses de Benjamin Franklin sur la foudre et l'électricité en installant un paratonnerre sur sa demeure.

Buffon fait œuvre de vulgarisation et, suivant ses vœux, est lu par le plus grand nombre. Son style soigné, souvent empreint de lyrisme, fait qu'il est considéré par ses contemporains comme un grand écrivain et même un grand poète. Son *Histoire naturelle*¹ connaît un succès qui rivalise avec celui de l'*Encyclopédie* de Diderot et d'Alembert. Buffon pressent la théorie de l'évolution, sans cependant adhérer à l'hypothèse du transformisme présentée par Maupertuis en 1751. Il attribue les variations observées à un rassemblement spontané de « molécules organiques ». Notons qu'en mathématiques, son résultat le plus connu est le problème de l'aiguille qui permet de déterminer expérimentalement la valeur de π .

Buffon meurt en 1788, d'une crise de gravelle, quelques mois avant la Révolution française.

8.6. Principaux ouvrages de Buffon (1707-1788)

- *La statique des végétaux, et l'analyse de l'air, expériences nouvelles lues à la Société royale de Londres, par M. Hales... ouvrage traduit de l'anglais, par M. de Buffon*, Paris, Debure l'aîné, 1735.
- *La méthode des Fluxions...* (de Newton), Paris, Debure l'aîné, 1740.
- *Histoire naturelle...*, Déterville, 1798...
- *Histoire naturelle, générale et particulière*, Imprimerie royale, 1749-1804.
- *Histoire naturelle des oiseaux*, Imprimerie royale, 1770-83.
- *Histoire naturelle des minéraux*, Imprimerie royale, 1783-1788.

1. Trente-six volumes seront publiés de son vivant entre 1749 et 1785 et huit autres après sa mort.

8.7. Références

- J. Roger, *Buffon*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 2, p. 576-582, 1970.
- J. Roger, *Buffon : un philosophe au Jardin du Roi*, Fayard, Paris, 1989.
- Pierre Gascar, *Buffon*, Gallimard, Paris, 1983.
- Th. Martin, *Buffon et l'arithmétique politique*, Math. Inf. Sci. hum., 37^e année, n° 148, p. 5-30, 1999.
- Yves Lassius, *Buffon, la nature en majesté*, Découvertes Gallimard, 2007.

D'Alembert

« Je suis très porté à croire que la matière dont il s'agit [la théorie des probabilités], ne peut être soumise, au moins à plusieurs égards, à un calcul exact et précis, également net dans ses principes et dans ses résultats¹. »

Quelles sont les contributions de d'Alembert relatives aux probabilités et aux statistiques ?

Karl Pearson² juge que ces contributions sont nulles, mais intéressantes à étudier, car, si le mathématicien génial du *Traité de dynamique* de 1743 et des *Recherches sur la précession des équinoxes et sur la nutation de l'axe de la terre* de 1749 a pu ne pas comprendre certains aspects de la théorie des probabilités cela doit inciter tous les scientifiques à la prudence quant à la formulation des principes logiques de ce sujet, et à la réalisation d'expérimentations lorsqu'elles sont possibles.

Les publications de d'Alembert, dans le domaine qui nous occupe, ont été relatives à trois domaines : les jeux de hasard, les avantages et les désavantages de l'inoculation, et les problèmes d'annuités et de rentes.

9.1. Doutes et questions sur le calcul des probabilités

D'Alembert a exposé ses vues sur le sujet dans un mémoire intitulé *Doutes et questions sur le calcul des probabilités*³ et publié en 1773. On peut y lire :

« On se plaint assez communément que les formules des mathématiciens, appliquées aux objets de la nature, ne se trouvent que trop en défaut. Personne néanmoins n'avait encore aperçu ou cru apercevoir cet inconvénient dans le *calcul des probabilités*. J'ai osé le premier proposer des doutes sur quelques principes qui servent de base à ce calcul. De grands géomètres ont jugé ces doutes *dignes d'attention*, d'autres grands géomètres les ont trouvés *absurdes* ; car pourquoi adoucerais-je les termes dont ils se sont servis ? La question est de savoir s'ils ont eu tort de les employer, et

1. D'Alembert, « Réflexions philosophiques et mathématiques sur l'application du calcul des probabilités à l'inoculation de la petite vérole », in *Mélanges de littérature, d'histoire, et de philosophie*, tome V, 1767.
2. *The history of statistics in the 17th and 18th centuries*. Giffin and Co. 1978.
3. *Œuvres philosophiques, historiques et littéraires de D'Alembert*, vol. IV, Paris, Bastien, 1805, p. 289-315.

en ce cas ils auraient doublement tort. Leur décision, qu'ils n'ont pas jugé à propos de motiver, a encouragé des mathématiciens médiocres, qui se sont hâtés d'écrire sur ce sujet, et de m'attaquer sans m'entendre. Je vais tâcher de m'expliquer si clairement, que presque tous mes lecteurs seront à portée de me juger.

Je remarquerai d'abord qu'il ne serait pas étonnant que des formules où on se propose de calculer l'*incertitude* même, pussent, à certains égards au moins, participer à cette incertitude, et laisser dans l'esprit quelques nuages sur la vérité rigoureuse du résultat qu'elles fournissent. Mais je n'insisterai point sur cette réflexion, trop vague pour qu'on puisse en rien conclure. Je ne m'arrêterai point non plus à faire voir que *la théorie des probabilités*, telle qu'elle est présentée dans les livres qui en traitent, n'est sur bien des points ni aussi lumineuse, ni aussi complète qu'on pourrait le croire ; ce détail ne pourrait être entendu que des mathématiciens ; et encore une fois je veux tâcher ici d'être entendu de tout le monde. J'adopte donc, ou plutôt j'admets pour bonne dans la rigueur mathématique, *la théorie ordinaire des probabilités*, et je vais seulement examiner si les résultats de cette théorie, quand ils seraient hors d'atteinte dans l'abstraction géométrique, ne sont pas susceptibles de restriction, lorsqu'on applique ces résultats à la nature.

Pour m'expliquer de la manière la plus précise, voici le point de la difficulté que je propose.

Le calcul des probabilités est appuyé sur cette supposition, que toutes les combinaisons différentes d'un même effet sont également possibles. Par exemple, si on jette une pièce en l'air cent fois de suite, on suppose qu'il est également possible que *pile* arrive cent fois de suite, ou que *pile* et *croix* soient *mêlés*, en suivant d'ailleurs entre eux telle succession particulière qu'on voudra ; par exemple, *pile* au premier coup, *croix* aux deux coups suivants, *pile* au quatrième, *croix* au cinquième, *pile* au sixième, au septième, etc.

Ces deux cas sont sans doute également possibles, mathématiquement parlant ; ce n'est pas là le point de la difficulté, et les mathématiciens médiocres dont je parlais tout à l'heure ont pris la peine fort inutile d'écrire de longues dissertations pour prouver cette égale possibilité. Mais il s'agit de savoir si ces deux cas, également possibles *mathématiquement*, le sont aussi *physiquement* et dans l'ordre des choses ; s'il *est physiquement* aussi possible que le même effet arrive cent fois de suite, qu'il l'est que ce même effet soit mêlé avec d'autres suivant cette loi qu'on voudra marquer. »

D'Alembert prend alors l'exemple du paradoxe de Saint-Pétersbourg, que nous avons exposé dans notre chapitre relatif à Daniel Bernoulli. Après avoir exposé le principe du jeu et le paradoxe de l'espérance mathématique infinie qui lui est associé, il écrit :

« Je demande s'il faut aller chercher bien loin la raison de ce paradoxe, et s'il ne saute pas aux yeux que cette prétendue *somme infinie* due par Paul au commencement du jeu, n'est infinie, en apparence, que parce qu'elle est appuyée sur une supposition fausse, savoir sur la supposition *que pile* peut n'arriver jamais, et que le jeu peut durer éternellement ?

Il est pourtant vrai, et même évident, que cette supposition est possible dans la rigueur mathématique. Ce n'est donc que *physiquement parlant* qu'elle est fausse.

Il est donc faux, *physiquement parlant*, que *pile* puisse n'arriver jamais.

Il est donc impossible, *physiquement parlant*, que *croix* arrive une infinité de fois de suite.

Donc, *physiquement parlant*, *croix* ne peut arriver de suite qu'un nombre fini de fois.

Quel est ce nombre ? c'est ce que je n'entreprends point de déterminer. Mais je vais plus loin, et je demande par quelle raison *croix* ne saurait arriver une infinité de fois de suite, *physiquement parlant* ? On ne peut en donner que la raison suivante : *c'est qu'il n'est pas dans la nature qu'un, effet soit toujours et constamment le même, comme il n'est pas dans la nature que tous les hommes et tous les arbres se ressemblent.*

Je demande ensuite s'il est plus possible, *physiquement parlant*, que le même effet arrive un très grand nombre de fois de suite, dix mille fois, par exemple, qu'il ne l'est que cet effet arrive une infinité de fois de suite ? Par exemple, est-il possible, *physiquement parlant*, que si on jette une pièce en l'air dix mille fois de suite, il vienne de suite dix mille fois *croix* ou *pile* ? Sur cela j'en appelle à tous les joueurs. »

De plus, d'Alembert ne croit pas à l'indépendance dans une répétition d'événements comme le lancer de pièces. La sortie de plusieurs *pile* consécutifs rendrait, selon lui, plus probable la sortie de *croix* au coup suivant.

Sa croyance principale est un principe métaphysique selon lequel « *il n'est pas dans la nature qu'un, effet soit toujours et constamment le même.* »

Voici ce qu'écrivait d'Alembert à la fin de son article :

« En finissant cet écrit, je tombe par hasard sur l'article *Fatalité* du dictionnaire Encyclopédique, article qu'on reconnaîtra aisément pour l'ouvrage d'un homme d'esprit et d'un philosophe ; et voici ce que j'y trouve, à propos du prétendu. *bonheur* ou *malheur* dans le jeu. "Ou il faut avoir égard aux coups passés pour estimer le coup prochain, ou il faut considérer le coup prochain, indépendamment des coups déjà joués ; *ces deux opinions ont leurs partisans.* Dans le premier cas, l'analyse des hasards me conduit à penser que si les coups précédents m'ont été favorables, le coup prochain me sera contraire ; que si j'ai gagné tant de coups, il y a tant à parier que je perdrai celui que je vais jouer, *et vice-versa.* Je ne pourrai donc jamais dire ; je suis en malheur, et je ne risquerai pas ce coup-là ; car je ne pourrais le dire que d'après les coups passés qui m'ont été contraires ; mais ces coups passés doivent plutôt me faire espérer que le coup suivant me sera favorable. Dans le second cas, c'est-à-dire, *si on regarde le coup prochain comme tout à fait isolé des coups précédents,* on n'a point de raison d'estimer que le coup prochain sera favorable plutôt que contraire, ou contraire plutôt que favorable ; ainsi on ne peut pas régler sa conduite au jeu, d'après l'opinion du destin, du bonheur, ou du malheur."

De ce passage je tire deux conséquences. La première, que, suivant l'auteur de cet excellent article, on peut se partager sur la question, *s'il est également probable qu'un effet arrive ou n'arrive pas, lorsqu'il est déjà arrivé plusieurs fois de suite.* Or, il me suffit que cela soit regardé comme douteux, pour m'autoriser à croire que l'objet de l'écrit précédent n'est pas aussi étrange que d'habiles mathématiciens l'ont imaginé. La seconde conséquence, c'est que l'analyse des hasards, telle que la conçoit l'auteur de l'article, *donne moins de probabilité aux combinaisons qui*

renferment la répétition successive du même effet, qu'aux combinaisons ou cet effet est mêlé avec d'autres. Or, cela ne se peut dire que de l'analyse des hasards considérée physiquement ; car à l'envisager du seul côté mathématique, toutes les combinaisons, comme nous l'avons dit, sont également possibles. Je crois donc pouvoir regarder l'auteur de l'article *Fatalité* comme partisan de l'opinion que j'ai tâché d'établir ; et un partisan de ce mérite me persuade de nouveau que cette opinion n'est pas une absurdité. »

9.2. L'article « Croix-Pile » de l'Encyclopédie

Par ailleurs, d'Alembert a écrit, vers 1755, le célèbre article « Croix ou pile » de l'Encyclopédie qui est reproduit ci-dessous :

« CROIX OU PILE, analyse des hasards.

Ce jeu, qui est très connu, et qui n'a pas besoin de définition, nous fournira les réflexions suivantes. On demande combien il y a à parier qu'on amènera *croix* en jouant deux coups consécutifs. La réponse qu'on trouvera dans tous les auteurs et suivant les principes ordinaires, est celle-ci. Il y a quatre combinaisons.

Premier coup.	Deuxième coup.
<i>Croix</i>	<i>Croix</i>
<i>Pile</i>	<i>Croix</i>
<i>Croix</i>	<i>Pile</i>
<i>Pile</i>	<i>Pile</i>

De ces quatre combinaisons, une seule fait perdre et trois font gagner, il y a donc 3 contre 1 à parier en faveur du joueur qui jette la pièce. S'il pariait en trois coups, on trouveroit huit combinaisons, dont une seule fait perdre, et sept font gagner ; ainsi, il y aurait 7 contre 1 à parier. Voyez COMBINAISON ET AVANTAGE. Cependant cela est-il bien exact ? Car, pour ne prendre ici que le cas de deux coups, ne faut-il pas réduire à une les deux combinaisons qui donnent *croix* au premier coup ? Car, dès qu'une fois *croix* est venu, le jeu est fini, et le second coup est compté pour rien. Ainsi, il n'y a proprement que trois combinaisons de possibles :

Croix, premier coup

Pile, *Croix*, premier et second coup.

Pile, *pile*, premier et second coup.

Donc il n'y a que 2 contre 1 à parier. De même, dans le cas de trois coups, on trouvera :

Croix

Pile, *croix*.

pile, *pile*, *croix*.

Pile, *pile*, *pile*.

Donc il n'y a que 3 contre 1 à parier. Ceci est digne, ce me semble, de l'attention des calculateurs, et irait à réformer bien des règles unanimement reçues sur les jeux de hasard. »

D'Alembert pose la question de savoir si le principe de calcul adopté par tous est réellement applicable. La seule réponse qu'il est aujourd'hui possible de

ENCYCLOPÉDIE *MÉTHODIQUE.*

MATHÉMATIQUES,

*Par MM. D'ALEMBERT, l'Abbé BOSSUT, DE LA LANDE,
le Marquis de CONDORCET, CHARLES, &c.*

TOME SECOND.



A PARIS,

Chez PANCKOUCKE, Libraire, Hôtel de Thou, rue des Poitevins;

A LIÈGE,

Chez PLONTEUX, Imprimeur des Etats.

M. DCC. LXXXV.

AVEC APPROBATION, ET PRIVILÈGE DU ROI.

faire est que l'expérience montre que, lorsque le nombre d'essais est grand, la fréquence d'amener croix en deux coups tend vers $3/4$ et que le principe de calcul est donc correct.

9.3. L'inoculation¹

Ce mémoire est essentiellement un commentaire et une critique du papier de Daniel Bernoulli de 1760 sur le même sujet. D'Alembert a lu son propre papier à l'Académie des sciences le 12 novembre 1760. Nous avons indiqué dans notre chapitre consacré à Daniel Bernoulli que ce dernier avait tenté de démontrer que, malgré les risques, la généralisation de la pratique dite de l'inoculation par la substance contenue dans les vésicules d'un malade de la variole permettrait de gagner deux à trois ans d'espérance de vie à la naissance.

Cette technique suscita cependant l'hostilité de nombreux médecins.

Voici les phrases de d'Alembert précisant le but de son mémoire :

« Je me propose ici trois objets :

- 1). J'examinerai successivement les différentes manières dont on a calculé jusqu'ici les avantages de *l'inoculation*, et j'essaierai de prouver que dans ces divers calculs, on n'a point, ce me semble, envisagé la question sous son véritable point de vue.
- 2). Je montrerai même que les avantages de cette opération, sous quelque aspect qu'on veuille les présenter, sont très difficiles à apprécier d'une manière satisfaisante, *si l'on convient que cette opération peut causer la mort*.
- 3). Je tâcherai de faire voir ensuite que *l'inoculation* peut être soutenue par d'autres raisons, qui non seulement doivent empêcher de la proscrire, mais qui paraissent même propres à l'autoriser. »

D'Alembert discute tout d'abord les hypothèses de mortalité adoptées par Bernoulli. Il insiste, aussi sur la difficulté, voire l'impossibilité de calculer les avantages de l'inoculation pour un individu. Il ajoute que même si le risque est le même à tout âge, il n'est pas perçu de la même manière à tout âge et par tout individu.

Il analyse les arguments favorables à l'inoculation, comme suit :

« Enfin, quand même le risque de mourir de l'inoculation, sagement administrée, serait plus grand que celui de mourir de la petite vérole naturelle dans le courant du même mois, ce risque, s'il n'était en effet que de 1 sur 1200, serait encore plus petit que celui de mourir de la petite vérole naturelle dans l'espace de trois mois. Car le nombre de ceux qui meurent à Paris de la petite vérole, année commune, est tout au moins de 1 sur 1 000 en trois mois ; donc le risque de mourir de la petite vérole naturelle en trois mois, serait au moins égal, et vraisemblablement supérieur à celui

1. *Réflexions sur l'inoculation*, Œuvres complètes, tome 1, Paris, Belin et Bossange, 1821, p. 465.

OPUSCULES MATHÉMATIQUES.

MÉMOIRES sur différens sujets de GÉOMÉTRIE;
de MÉCANIQUE, D'OPTIQUE, D'ASTRONOMIE &c.

Par M. D'ALEMBERT, de l'Académie Française, des
Académies Royales des Sciences de France, de Prusse &
d'Angleterre, de l'Académie Royale des Belles - Lettres de
Suède, & de l'Institut de Bologne.

TOME PREMIER.



A PARIS,

Chez DAVID, rue & vis-à-vis la grille des Mathurins.

M. DCC. LXI.

AVEC APPROBATION ET PRIVILÈGE DU ROI.

de mourir en un mois de l'inoculation. Or risquer de mourir au bout d'un mois, ou dans l'espace de trois, est à peu près la même chose pour le commun des hommes. On ne devrait donc pas balancer à préférer celui de ces deux risques, qui délivre pour toujours de la crainte de la petite vérole. Par là on aurait l'avantage de s'assurer à la fois une vie plus longue et une plus grande tranquillité [...] Lorsqu'il est question d'un avantage, même éloigné, il y a une infinité de cas, surtout dans le cours de la vie, où une probabilité très petite de danger, qui balance cet avantage, doit être traitée comme si elle était absolument nulle. Ce principe, pour le dire en passant, est très important dans la théorie des jeux de hasard, et peut servir à résoudre des questions épineuses et délicates, qui n'ont point été résolues jusqu'ici, ou qui l'ont été mal, mais qui ne sont pas quant à présent de notre objet.

Voilà, ce me semble, ce qu'on peut dire de plus fort en faveur de l'inoculation ; cette manière d'en calculer l'avantage, quoiqu'elle ait échappé à ses plus zélés partisans, est, si je ne me trompe, la moins sujette aux objections qu'il est possible. Il est vrai qu'elle ne donne pas et ne saurait donner la valeur précise, mathématique et rigoureuse, de l'avantage qu'il y a à se faire inoculer ; mais elle montre, et cela suffit, que l'avantage est très considérable ; je ne suis donc pas surpris que cet avantage détermine un grand nombre de citoyens à subir l'inoculation, ou à la faire subir aux personnes qui les intéressent. »

Pour tenter de rester impartial, d'Alembert souligne que des doutes subsistent sur les risques liés à l'inoculation, et qu'il faut faire la différence entre les avantages pour les individus et ceux pour la population totale. Ainsi, si l'inoculation enlevait, au hasard, un cinquième d'une population et permettait aux autres de vivre jusqu'à cent ans, elle serait avantageuse pour la population mais risquée individuellement.

Au total, voici les conclusions de d'Alembert :

- « 1). Il y a lieu de croire qu'on ne meurt jamais de l'inoculation, quand elle est sagement administrée, et après un examen convenable.
- 2). Il est extrêmement rare, pour n'en pas dire davantage, qu'un inoculé sur qui l'opération a réussi, ait repris la petite vérole.
- 3). S'il n'est pas démontré en rigueur que l'inoculation augmente la vie moyenne des hommes, il est encore moins prouvé qu'elle la diminue ; il est même vraisemblable qu'elle doit l'augmenter, puisqu'elle délivre, ou absolument, ou presque absolument, d'une cause de mort, sans qu'il soit prouvé qu'elle en substitue d'autres à la place.

Il faut donc bien se garder, ce me semble, d'arrêter ou de retarder les progrès de cette opération. [...]

Je suis donc bien éloigné de dissuader mes concitoyens d'une pratique dont l'utilité paraît, au moins jusqu'ici, beaucoup mieux constatée que ses inconvénients. Les objections proposées dans les deux premières parties de cet écrit, n'attaquent que les mathématiciens qui pourraient trop se presser de réduire cette matière en équations et formules, mais je crois d'ailleurs en avoir dit assez pour faire voir, *que si les avantages de l'inoculation ne sont pas de nature à être appréciés mathématiquement, ils n'en paraissent pas moins réels.* »

9.4. Le mémoire sur la durée de la vie

Les *Opuscles mathématiques* de d'Alembert ont été écrits sur une période de vingt-cinq ans entre 1758 et 1783. Huit volumes ont paru¹ totalisant près de 4 000 pages.

Le mémoire (en fait un extrait de lettre écrite en 1767) *Sur la durée de la vie* occupe les pages 92 à 98 du volume IV, publié en 1768.

D'Alembert y analyse la probabilité de la vie et insiste sur la différence entre les notions de moyenne et de médiane, utilisées dit-il par Daniel Bernoulli (moyenne) et Buffon (médiane) dans leurs tables de mortalité.

Une analyse de ce papier a été faite par K. Pearson².

P. Crépel³ a, par ailleurs indiqué que le tome IX des *Opuscles* de d'Alembert, rédigé entre 1781 et sa mort en 1783 et inédit comporte un texte intitulé *Sur les annuités* comportant 91 articles. La structure du papier a été reconstituée par P. Crépel. Il comporte « une comparaison entre trois méthodes de calcul des rentes viagères : 1) par le calcul de la vie moyenne par Deparcieux, 2) par (ce que P. Crépel appelle) la vie médiane, 3) par la vie moyenne de toutes les personnes vivant à une époque. »

9.5. Biographie de d'Alembert (1717-1783)

Jean-Le-Rond d'Alembert naît le 16 novembre 1717 d'un amour illégitime entre la marquise de Tencin et le chevalier Destouches, alors officier d'artillerie qui deviendra général en 1720. Il est abandonné par sa mère à sa naissance sur les marches de l'église Saint Jean-Le-Rond à Paris⁴. Un commissaire du quartier le découvre et décide, compte tenu de son état (il est « incroyablement maigre et chétif ») de lui trouver une nourrice. Il la trouve en la personne d'une vitrière, Madame Rousseau, femme d'un artisan, que d'Alembert considérera toute sa vie comme sa mère. Il reste d'ailleurs jusqu'à l'âge de 48 ans dans cet humble, mais accueillant, foyer. Son père, le Chevalier Destouches, retrouve la trace du nouveau-né, lui laisse une pension de 1 200 livres (qui par testament se prolonge après sa mort qui survient quand d'Alembert n'a que dix ans), et s'assure qu'il bénéficiera d'une bonne éducation. D'Alembert va d'abord en pension et s'y montre un brillant élève. Il entre en 1730, à douze ans, au Collège des Quatre-Nations et suit l'enseignement des jansénistes. L'entrée dans ce Collège, réservé à ceux qui

1. *Opuscles mathématiques*, Paris, David Briasson, Jombert, 1761-1780.

2. Pearson K., *The history of statistics*, 1978, p. 562-564.

3. Crépel P., « Les dernières perfidies de d'Alembert », *Math. and Social Sciences*, 44^e année, n° 176, p. 61-87.

4. C'est à cause de ce lieu que, pendant un peu plus de vingt ans, il est appelé Jean-Baptiste-le-Rond. À la fin de ses études il se fera appeler d'Alembert.

possédaient des lettres de noblesse, semble bien montrer que les origines réelles de d'Alembert étaient connues. Ses dispositions intellectuelles sont vite reconnues, en particulier son commentaire sur l'*Épître de Saint Paul aux romains* fait voir en lui un nouveau Pascal. « Pour rendre la ressemblance plus parfaite, dit Condorcet¹, on lui fit suivre des leçons de mathématiques ». Le Collège des Quatre-Nations est le premier à posséder une chaire de mathématiques², et d'Alembert se passionne pour ce domaine que lui fait découvrir son excellent professeur, Monsieur Caron. D'Alembert devient, en 1735, « Bachelier ès Arts », mais ses connaissances acquises de façon autodidacte vont bien au-delà de ce diplôme. D'Alembert passerait volontiers ses journées à faire des mathématiques, mais cela ne mène à aucune situation et la pension laissée par son père se révélerait vite insuffisante. Il se décide donc à faire des études de droit, et il obtient un titre d'avocat en 1738. Il poursuit ensuite ses études, « fait médecine » et essaie d'oublier les mathématiques. Mais sa passion est la plus forte, il abandonne la médecine et décide, quel qu'en soit le prix à payer, de se consacrer aux mathématiques.

D'Alembert transmet alors plusieurs mémoires à l'Académie des sciences (le premier date de 1739), ce qui lui vaut d'être remarqué par Clairaut. Grâce à ce dernier, notre mathématicien entre en 1741, à l'âge de 24 ans, dans cette prestigieuse Académie comme « associé astronome adjoint ».

D'Alembert présente en 1742 à l'Académie son célèbre *Traité de dynamique*, qui sera publié l'année suivante. Ce traité marque le fondement de la mécanique rationnelle. Dès cette époque d'Alembert se montre profondément newtonien. Comme le dit J.-P. Luminet : « Pour d'Alembert, et jusque dans ses vieux jours, la science ne commençait pas avec Newton, elle s'arrêtait après lui ». Dans le *Traité de dynamique*, d'Alembert expose son *Principe*. En rapportant la dynamique tout entière aux seules trois lois du mouvement de Newton, le principe de d'Alembert permet de trouver le mouvement de systèmes soumis à des forces de liaison. On trouve également là les prémices des travaux de Lagrange. D'Alembert met en œuvre cette dynamique dans l'étude de l'oscillation des corps pesants et de la rotation des corps autour de leur centre de gravité. Il applique également, dans son *Traité de l'équilibre et du mouvement des fluides* (1744), les principes de la dynamique à la mécanique des fluides.

La contribution de d'Alembert en mécanique des fluides est à la fois vaste et importante. En 1746 il présente un ouvrage intitulé *Réflexions sur la cause*

-
1. Condorcet, « Éloge de M. d'Alembert », *Histoire de l'Académie royale des sciences avec les mémoires de mathématique et de physique tirés des registres de cette Académie, année 1783*, publiée en 1786, p. 76.
 2. On compte parmi les premiers titulaires de cette chaire, Pierre Varignon à qui l'on doit la diffusion en France, sous une forme beaucoup plus accessible que l'original, des *Principia* de Newton.

générale des vents, qui lui vaut d'être élu à l'Académie de Berlin. En 1752 il poursuit en ce sens avec l'étude de la résistance des fluides. Malheureusement il se trouve en compétition avec Euler et les deux savants deviennent vite rivaux. Cette double rivalité, avec Clairaut à l'Académie de Paris et avec Euler à l'Académie de Berlin, laissera des traces. D'Alembert, poussé à publier rapidement, est moins clair que ses compétiteurs à qui l'histoire attribuera quelquefois la paternité de résultats que d'Alembert avait présentés le premier.

Une autre contribution majeure de d'Alembert est son étude des cordes vibrantes. Il fournit dans ce travail, qui date de 1747, non seulement la première équation d'onde ainsi que la solution générale associée, mais surtout, serions-nous tentés de dire, il introduit pour la première fois les équations aux dérivées partielles en physique. Cet apport sera immortalisé par l'équation qui porte son nom.

En dynamique, d'Alembert contribue largement à l'étude des problèmes de la « mécanique céleste », en particulier du problème des trois corps. Il aborde ce thème dès 1745, et en donne une solution approchée à la même époque, et indépendamment de ses « concurrents » Clairaut et Euler. Cela lui permet d'étudier la précession des équinoxes et la rotation de l'axe de la Terre¹, il frôle la notion de nutation, qui ne sera finalement découverte que par Bradley. L'essentiel de l'œuvre en physique et mathématique de d'Alembert date de la période allant de 1742 à 1754. Mais là ne s'arrête pas son œuvre. L'année 1754 marque une transition dans la vie de d'Alembert. D'une part, il est élu à l'Académie française, d'autre part il fait la connaissance dans un salon parisien d'une jeune fille de 22 ans, Julie de Lespinasse, qui comme lui est le fruit d'un amour illégitime.

Physicien et mathématicien, d'Alembert est également philosophe. C'est en 1745 qu'il entre en relation avec Denis Diderot. Ils signent ensemble un contrat pour traduire la *Cyclopaedia* d'Ephraim Chambers. Mais le projet évolue, de là va naître l'*Encyclopédie*. L'ouvrage envisagé est monumental, il est conçu comme un *Dictionnaire raisonné des arts, des sciences et des métiers*. Comme l'indique Condorcet, d'Alembert et Diderot se proposent « de réunir tout ce qui a été découvert dans les sciences, ce que l'on a pu connaître des productions du globe, les détails des arts que les hommes ont inventés, les principes de la morale, ceux de la politique et de la législation, les lois qui gouvernent les sociétés, la métaphysique des langues et les règles de la grammaire, l'analyse de nos facultés et jusqu'à l'histoire de nos opinions ». La contribution de d'Alembert porte naturellement sur les sciences (mathématiques, physique, astronomie, etc.) mais également sur des questions plus générales des connaissances. Au total ce n'est pas moins de six cents articles

1. C'est en 1749 qu'il publie ses *Recherches sur la précession des équinoxes*.

qu'il écrira, sans compter son célèbre *Discours préliminaire de l'Encyclopédie* (1751). La contribution de Diderot est encore plus colossale, il rédige plusieurs milliers d'articles ! La coutume voulant que quelques « grands » académiciens des sciences entrent à l'Académie française¹ on n'est pas surpris de voir d'Alembert élu dans cette institution en 1754². En 1759, d'Alembert se brouille avec Diderot, démissionne, mais continue la rédaction des articles scientifiques de l'*Encyclopédie*.

Nous avons signalé la rencontre, en 1754 de d'Alembert avec Julie de Lespinasse. Peu de temps après, la jeune fille tombe malade et c'est notre mathématicien qui la soigne avec dévouement. L'amitié se transforme en liaison, probablement aux alentours de 1760. Quelques années plus tard (1763), d'Alembert séjourne trois mois à la cour de Frédéric II. Le Roi de Prusse lui offre de rester et de présider l'Académie de Berlin mais d'Alembert refuse. Tout comme il avait refusé, peu avant, l'offre de l'Impératrice de Russie de devenir le précepteur de son fils. Ses refus sont probablement liés au fait qu'il ne voulait pas s'éloigner de Julie. Lorsque d'Alembert tombe gravement malade, en 1765, Julie prend soin de lui, tandis que son médecin lui conseille de quitter le logis de sa famille adoptive, les Rousseau, car les conditions d'hygiène devenaient incompatibles avec son état de santé. C'est probablement autour de cette époque que d'Alembert se met en ménage avec Julie. C'est également à ce moment que notre mathématicien renoue avec Diderot. L'année suivante les derniers volumes de l'*Encyclopédie* sont achevés. D'Alembert a alors une vie mondaine et participe au salon que Julie a ouvert en 1764. Cela n'empêche pas d'Alembert, qui a vécu une grande partie de sa vie dans deux mondes très différents³, de rester très modeste. « M. d'Alembert, dit Condorcet, établit pour principe de morale l'obligation de ne pas regarder comme légitime l'usage de son superflu lorsque d'autres hommes sont privés du nécessaire ; et de ne disposer soi-même que de la portion de la fortune qui est formée non au dépens du nécessaire des autres, mais par la réunion d'une partie de leur superflu ». D'Alembert est élu en 1772 secrétaire perpétuel de l'Académie française. Il se fait un devoir de continuer la rédaction de l'histoire de cette Compagnie, que ses prédécesseurs avaient négligée. Il s'engage à écrire la vie de tous les Académiciens morts entre 1700 et 1772. Ces *Éloges historiques* vont occuper une grande partie de son temps pendant les trois années suivantes. Au total c'est près de soixante-dix éloges que d'Alembert rédige. La lecture de ses *Éloges* aux séances publiques de l'Académie française lui vaut d'ailleurs un grand succès

1. Citons : Fontenelle, Maupertuis, Bougainville, Buffon, Condorcet, etc.

2. L'Académie donne en 1756 à d'Alembert le titre de pensionnaire surnuméraire (avec l'attribution d'une pension de 2 400 livres), mais ce n'est qu'en 1765 qu'il deviendra pensionnaire de droit.

3. D'Alembert a, jusqu'à 48 ans, pris l'habitude de travailler le matin dans le logis de ses parents adoptifs et de se rendre l'après midi dans les salons ou à l'Académie.

dans le public mondain. À la fin de sa vie, d'Alembert est membre de presque toutes les Académies d'Europe.

D'Alembert meurt le 29 octobre 1783, à l'âge de soixante-six ans, suite à des calculs rénaux.

9.6. Principaux ouvrages scientifiques de d'Alembert

- *Traité de dynamique*, Paris, David, 1743.
- *Réflexions sur la cause générale des vents...*, Paris, David l'aîné, 1747.
- *Recherches sur la précession des équinoxes et sur la nutation de l'axe de la terre dans le système newtonien...*, Paris, David, 1749.
- *Essai d'une nouvelle théorie de la résistance des fluides...*, Paris, David l'aîné, 1752.
- *Opuscules mathématiques ou Mémoires sur différens sujets de géométrie, de mécanique, d'optique, d'astronomie*, Paris, David, [puis] Briasson [puis] C.-A. Jombert, 1761-1780.

9.7. Références

- J. Morton Briggs, *Alembert*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 1, p. 110-117, 1970.
- J. Bertrand, *D'Alembert*, 1889.
- R. Grimsley, *Jean d'Alembert, 1717-1783*, Oxford, 1963.
- V. Le Ru, *D'Alembert philosophe*, Librairie Philosophique J. Vrin, coll. « Mathesis », Paris, 1994.

Les contributions de Cotes et Simpson à la théorie des erreurs

10.1. Cotes

Roger Cotes est considéré à juste titre comme le premier mathématicien ayant étudié la combinaison des observations et la minimisation des erreurs. Son mémoire intitulé *Aestimatio errorum in mixta mathesi* est inclus dans la publication posthume de ses travaux, faite en 1722, par Robert Smith, à Cambridge, sous le titre *Harmonia mensuratorum*.

Avant Cotes, il était d'usage, en astronomie par exemple, de prendre la moyenne arithmétique d'une série de mesures.

Cotes proposa une valeur moyenne pondérée, la valeur adoptée étant le barycentre des différentes valeurs affectées chacune d'un poids spécifique. À la page 22 de son mémoire, on peut lire le texte qui est reproduit ci-dessous :

22

ÆSTIMATIO ERRORUM

Ad eundem fere modum in aliis casibus Limites inveniuntur Errorum qui ex minus accuratis observationibus ortum ducunt, quin & Positiones ad Observandum commodissimæ deprehenduntur: ut mihi vix quidquam ulterius desiderari videatur postquam ostensum fuerit quæ ratione Probabilitas maxima in his rebus haberi possit, ubi diversæ Observationes, in eundem finem institutæ, paullulum diversas ab invicem conclusiones exhibent. Id autem fiet ad modum sequentis Exempli. Sit p locus Objecti alicujus ex Observatione prima definitus, q, r, s ejusdem Objecti loca ex Observationibus subsequenter; sint insuper P, Q, R, S pondera reciproce proportionalia spatiis Evagationum, per quæ se diffundere possint Errores ex Observationibus singulis prodeuntes, quæque dantur ex datis Errorum Limitibus; & ad puncta p, q, r, s posita intelligantur pondera P, Q, R, S , & inveniatur eorum gravitatis centrum Z : dico punctum Z fore Locum Objecti maxime probabilem, qui pro vero ejus loco tutissime haberi potest.

$p \cdot$ q
 $\cdot Z$
 $r \cdot$ s

HARMONIA MENSURARUM,

SIVE

ANALYSIS & SYNTHESIS

PER RATIONUM & ANGULORUM MENSURAS

P R O M O T Æ :

ACCEDUNT ALIA

OPUSCULA MATHEMATICA:

P E R

ROGERUM COTESIUM.

————— *Si propius ftes,*
Te capient magis. —————

EDIDIT ET AUXIT

ROBERTUS SMITH

Collegii S. Trinitatis apud Cantabrigienses Socius;

Astronomiæ & Experimentalis Philosophiæ

Post COTESIUM Professor.

CANTABRIGIÆ, MDCCXXII.

Ce texte peut se traduire comme suit (à partir de : sit p locus...) :

« Soit p l'emplacement d'un objet quelconque défini par une première observation et q, r, s , les emplacements du même objet obtenues par des observations subséquentes. Soient aussi P, Q, R, S , les poids réciproquement proportionnels aux déplacements qui peuvent résulter des erreurs d'observation individuelles, et qui sont donnés par les limites données des erreurs. Les poids P, Q, R, S , sont imaginés placés en p, q, r, s , et leur centre de gravité Z est trouvé : Je dis que le point Z est l'emplacement le plus probable de l'objet, et peut être pris en toute sécurité pour son véritable emplacement. »

Laplace, en 1812, dans sa *Théorie Analytique des probabilités*¹ confirme la priorité de la méthode de Cotes qui est l'ancêtre de la méthode des moindres carrés :

« Cotes est, si je ne me trompe, le premier qui ait donné une règle générale pour faire concourir à la détermination d'un élément, plusieurs observations, proportionnellement à leur influence. En considérant chaque observation comme une fonction de l'élément, et regardant l'erreur de l'observation comme une différentielle infiniment petite ; elle sera égale à la différentielle de la fonction, prise par rapport à cet élément. Plus le coefficient de la différentielle de l'élément sera considérable, moins il faudra faire varier l'élément, pour que le produit de sa variation, par ce coefficient, soit égal à l'erreur de l'observation ; ce coefficient exprimera donc l'influence de l'observation sur la valeur de l'élément. Cela posé, Cotes représente toutes les valeurs de l'élément données par chaque observation par les parties d'une droite indéfinie, toutes ces parties ayant une commune origine. Il conçoit ensuite, à leurs autres extrémités, des poids proportionnels aux influences respectives des observations. La distance de l'origine commune des parties, au centre commun de gravité de tous ces poids, est la valeur qu'il choisit pour l'élément. »

Ajoutons que le papier de Cotes permet d'entrevoir le lien entre les paramètres *variance* et *écart-type* d'une série et la physique : l'espérance mathématique $E(X)$ correspond au *centre de gravité* de points pondérés. $E(X^2)$ n'est autre que le *moment d'inertie* par rapport à l'origine. La variance $V(X)$ apparaît en tant que le moment d'inertie des masses par rapport à leur centre de gravité.

Pearson, en 1893, a introduit le terme de *standard deviation* (déviations standard) pour signifier ce que l'on appelle aujourd'hui l'*écart-type*, racine carrée de la *variance*.

10.2. Simpson

En 1740, Simpson publia un court traité intitulé *The nature and laws of chance*. Deux ans plus tôt avait paru la deuxième édition, très sensiblement augmentée, du livre de De Moivre *Doctrine of Chances*. Dans son ouvrage, qui était vendu, soulignait-il, à un *prix modéré*, Simpson tentait de présenter

1. Page 346 de l'édition originale.

les résultats de De Moivre sous une forme simplifiée, mais il revendiquait cependant la solution de deux problèmes pour lesquels De Moivre n'avait donné que les résultats. Simpson s'étant à nouveau largement inspiré d'un ouvrage de De Moivre *Annuities on Lives*, en publiant son *Annuities and Reversions* en 1742, une dispute prit naissance entre les deux mathématiciens. Cependant, les *Annuities and Reversions* utilisaient des tables récentes basées sur dix années des *Bills of mortality of London*, alors que l'ouvrage de De Moivre de 1725 était basé sur les tables de la ville de Breslau données par Halley. De Moivre répondit en sortant sa deuxième édition des *Annuities on Lives* en 1743 avec une préface offensive vis-à-vis de Simpson. Ce dernier crut devoir répondre rapidement par un pamphlet de 16 pages dont le titre indique clairement le contenu : *Appendix containing some remarks on a late book on the same subject, with answers to some personal and malignant misrepresentations, in the preface thereof*.

Le mémoire de Simpson dont nous allons parler maintenant a paru dans les *Philosophical Transactions*¹, en 1755 et a pour titre *A Letter to the Right Honourable George, Earl of Macclesfield, President of the Royal Society, as to the Advantage of taking the Mean of a Number of Observations in practical Astronomy*.

XIX. *A Letter to the Right Honourable George Earl of Macclesfield, President of the Royal Society, on the Advantage of taking the Mean of a Number of Observations, in practical Astronomy: By T. Simpson, F. R. S.*

My Lord,

Read April 10, 1755. **I**T is well known to your Lordship, that the method practised by astronomers, in order to diminish the errors arising from the imperfections of instruments, and of the organs of sense, by taking the Mean of several observations, has not been so generally received, but that some persons, of considerable note, have been of opinion, and even publicly maintained, that one single observation, taken

1. Volume 49, p. 82-93, 1755.

Simpson aborde le problème des observations astronomiques en se focalisant sur les erreurs de positionnement du corps observé.

Il appelle x l'erreur, qu'il considère comme une variable aléatoire distribuée uniformément sur les entiers entre $-a$ et $+a$, dont la fonction génératrice est :

$$g(t) = \frac{t^{-a} + t^{-a+1} + \dots + t^a}{2a+1} = \frac{1 - t^{2a+1}}{(2a+1)t^a(1-t)}.$$

Soit

$s_n = x_1 + \dots + x_n$, la somme des erreurs,

la probabilité $P(s_n = s) = p_n(s)$ est trouvée, à partir de la fonction génératrice, comme étant

$$p_n(s) = (2a+1)^{-n} \sum_i (-1)^i \binom{n}{i} \binom{n+s+na-1-(2a+1)i}{n-1}$$

avec $0 \leq i \leq \left\lfloor \frac{s+na}{2a+1} \right\rfloor$.

Cette probabilité peut être interprétée comme celle d'obtenir $s + n(a+1)$ points en n tirages avec un dé à $(2a+1)$ faces.

Or, parmi les différents problèmes liés aux jeux de dés, de Moivre avait étudié la distribution de la somme des faces de plusieurs dés¹. Le résultat indique que cette distribution de la somme des faces de n dés ayant chacun f faces numérotées 1 à f est donnée par les coefficients du développement de

$$(1 + r + r^2 + r^3 + \dots + r^{f-1})^n = \frac{(1 - r^f)^n}{(1 - r)^n}.$$

À partir du résultat de De Moivre, Simpson écrit la fonction génératrice pour une distribution des erreurs triangulaire et symétrique et compare les probabilités d'obtenir une erreur bornée entre deux limites et une erreur moyenne bornée entre ces mêmes limites. Il conclut :

« Prendre la moyenne d'un certain nombre d'observations diminue grandement les chances pour les petites erreurs et élimine pratiquement toute possibilité de grandes erreurs. »

Il donne un exemple numérique sur 6 observations.

Comme l'indique Stigler², Thomas Bayes signala, dans une lettre à John Canton de la *Royal Society*, que Simpson n'avait pas tenu compte du fait que, dans son approche, les observations devaient être exemptes d'erreurs systématiques.

1. *Doctrine of Chances*, 1738, p. 35-37.

2. Stigler S. M., *The history of statistics*, Belknap, Cambridge, M.A., 1986, p. 94-95.

Simpson indiqua donc clairement cette hypothèse lorsqu'il publia une version révisée de son travail¹ en 1757. Il étendit alors son analyse à une distribution triangulaire continue au lieu d'une distribution discrète.

On trouvera des détails sur les deux papiers de Simpson de 1755 et 1757 dans diverses analyses².

10.3. Biographie de Cotes (1682-1716)

Roger Cotes est né le 10 juillet 1682 à Burbage, dans le Leicestershire. Dès l'âge de 12 ans on remarque chez lui un talent mathématique exceptionnel. En 1699 il entre au *Trinity College* de Cambridge. Il en devient Fellow en 1705 et est nommé, l'année suivante, à l'âge de 23 ans, professeur d'astronomie et de physique. Il est élu membre de la Royal Society le 30 novembre 1711. Peu de temps après il est ordonné diacre (20 mars 1713), puis prêtre (31 mai 1713). Durant la période allant de 1709 à 1713, Roger Cotes passe une grande partie de son temps à éditer la seconde édition des *Principia* de Newton. Il ne publiera qu'un seul article dans sa vie, *Logometria*. Il s'agit, selon la présentation qu'il en fait à Newton « d'une nouvelle sorte de construction en géométrie, qui me semble facile, simple et générale ».

Cotes possédait une avance substantielle dans les domaines tels que les logarithmes, le calcul intégral et les méthodes numériques de calcul (en particulier l'interpolation), ce qui a fait dire au grand Newton : « s'il avait vécu [plus longtemps] nous aurions sans doute appris des choses nouvelles ».

Roger Cotes est mort le 5 juin 1716, de violentes fièvres et a été enterré quelques jours plus tard dans la chapelle du *Trinity College*.

10.4. Ouvrages de Cotes et références

- Préface de la deuxième édition des *Principia* de Newton, Cambridge, 1713.
- *Logometria*, Philosophical Transactions of the Royal Society, 29, p. 5-47, 1714.
- *Harmonia mensurarum*, édité par R. Smith, Cambridge, 1722.
- *Hydrostatical and pneumatical lectures*, éditées par Smith, London, 1738.
- *De descensu gravium, de motu pendulorum in cycloide et de motu projectilium*. Cambridge, 1770.

1. *Miscellaneous tracts*, Nourse, London, 1757.

2. A. Hald, *A History of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 35-39.

Références

- J. M. Dubbey, *Cotes*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 3, p. 430-432, 1970.
- R. Gowing, *Roger Cotes, natural philosopher*, New York, 1983.

10.5. Biographie de Simpson (1710-1761)

Thomas Simpson est né le 20 août 1710 à Market Bosworth, dans le comté de Leicester, en Angleterre. Son père était tisserand. Après de courtes études dans cette ville Thomas devient lui-même tisserand. Une éclipse solaire qui a lieu en 1724 attise son intérêt pour l'astronomie et pour les mathématiques. Il étudie alors en autodidacte les mathématiques tant et si bien qu'il quitte son domicile pour devenir instituteur à Nuneatin. Il y rencontre une jeune fille, Swinfield, qu'il épouse en 1730. Le couple aura une fille, Elizabeth, née en 1736 et un fils, Thomas, né en 1738. En 1733, les Simpson quittent Nuneatin pour s'installer à Derby où ils ne restent pas très longtemps puisqu'en 1736 on les retrouve à Londres. Thomas Simpson reprend alors son métier de tisserand et devient membre de la *Société Mathématiques Spitalfields* qui réunit des passionnés de mathématiques de différentes professions : la moitié des membres sont tisserands, l'autre moitié est constituée de brasseurs, de boulangers, de chaudronniers ou de fabricants de briques. Simpson, comme ses autres collègues de la Société Spitalfields enseigne les mathématiques dans les cafés. Le principe est simple : les « auditeurs » payent un droit d'entrée de 1 penny et boivent un café en écoutant les leçons.

En 1743, Simpson est nommé responsable des mathématiques à la toute nouvelle Académie militaire royale, fondée deux ans plus tôt. Il s'intéresse alors aux questions d'ingénierie et en particulier aux problèmes posés par les fortifications. En 1745 il est élu membre de la *Royal Society* et sera plus tard, en 1758, admis à l'Académie des sciences de Suède. Ces honneurs récompensent des travaux de grande qualité : en 1737, Simpson a publié *Un nouveau traité sur les fluxions*, sujet qui n'est dominé à l'époque que par quelques mathématiciens en Europe. Mais le travail qui le fait vraiment connaître est relatif à l'interpolation et aux méthodes numériques d'intégration. On lui doit en particulier la méthode dite de « Newton-Raphson » pour la résolution numérique des équations $f(x) = 0$, méthode qu'il a été le premier à introduire en 1740. En cette même année, Simpson publie un traité sur les lois du hasard. Il se base sur des travaux de De Moivre ce qui entraîne un conflit de priorité sur l'introduction des probabilités dans le calcul des annuités. Simpson applique également l'approche statistique au problème de la mesure comme en témoigne son article, publié en 1757 et intitulé : *Une tentative de montrer l'avantage de prendre la moyenne d'un*

certain nombre d'observations dans l'astronomie pratique. Il aborde également des problèmes de mécanique et de mécanique céleste, tels que celui de la précession des équinoxes. En 1754 il devient éditeur du journal *Ladies Diary*, journal dans lequel il avait l'habitude de publier, depuis 1736, des réponses à des problèmes posés par ce magazine.

Thomas Simpson s'éteint dans la ville où il est né, Market Bosworth, le 14 mai 1761.

10.6. Ouvrages de Simpson et références

- *A new Treatise of Fluxions*, London, 1737.
- *The nature and laws of chance*, London, 1740.
- « Essays on Several Curious and Useful Subjects », in *Speculative and Mix'd Mathematicks. Illustrated by a Variety of Examples*, London, Woodfall, 1740.
- *The Doctrine of Annuities and Reversions*, London, Nourse, 1742
- *A Letter to the Right Honourable George, Earl of Macclesfield, President of the Royal Society, as to the Advantage of taking the Mean of a Number of Observations in practical Astronomy*, Philos. Trans., volume 49, p. 82-93, 1755.
- *Miscellaneous tracts*, Nourse, London, 1757.
- *Essays on Several Curious and Useful Subjects, In Speculative and Mix'd Mathematicks. Illustrated by a Variety of Examples*, London, Woodfall, 1740.

Références

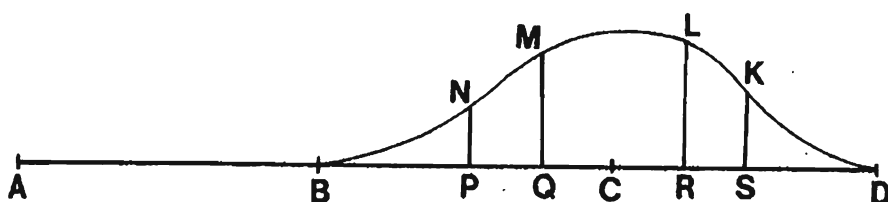
- P. J. Wallis, *Simpson Thomas*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 12, p. 443-445, 1970.
- F. M. Clarke, *Thomas Simpson and his times*, New York, 1929.

Lambert

11.1. La méthode du maximum de vraisemblance¹

En 1760, dans son ouvrage intitulé *Photometria*², Lambert a discuté des mesures d'une grandeur (dans son cas, une intensité lumineuse) et de l'estimation d'une valeur vraie de cette grandeur basée sur n observations indépendantes. En fait, il cherche la valeur moyenne des erreurs de manière qu'elle s'écarte le moins de la valeur moyenne vraie, c'est-à-dire qu'elle ait la plus grande probabilité. Il suppose que la distribution est connue, à part la valeur moyenne vraie.

Lambert trace alors une distribution empirique figurée ci dessous :



AC étant la valeur vraie à déterminer, Lambert suppose quatre mesures et appelle les ordonnées QM, PN, RL, SK, les occurrences (fréquences) des erreurs CQ, CP, CR, et CS.

Il appelle celles-ci occurrences absolues ou vraies pour les distinguer des occurrences observées. Il ajoute que celles-ci doivent coïncider avec celles-là si l'expérience est répétée à l'infini.

Il suppose également que les quantités AP, AQ, AR, AS ont été observées n , m , l , k fois et recherche tous les cas possibles pour lesquels ceci peut se produire. À partir de la théorie des combinaisons et des permutations, il déduit :

1. Ces travaux de Lambert ont été découverts par O. B. Sheynin :
 - Sheynin, O. B., *Origin of the theory of errors*, Nature, London, 211, 1966, p. 1003-4.
 - Sheynin, O. B., *J. H. Lambert's work on probability*, AHES, 7, 1971, p. 244-256.
2. Lambert J. H., *Photometria sive de Mensura et Gradibus Luminis, Colorum et Umbrae*, Detleffsen, Augsburg, 1760, paragraphes 272-305.

I. H. LAMBERT
ACADEMIAE SCIENTIARVM ELECTO-
RALIS BOICAE, ET SOCIETATIS PHYSICO-ME-
DICAE BASILIENSIS MEMBRI, REGIAE SOCIETATI
SCIENTIARVM GOETTINGENSI COMMERCIO
LITERARIO ADIVNCTI

PHOTOMETRIA

SIVE

DE

MENSURA ET GRADIBVS
LUMINIS,
COLORVM ET VMBRAE.



AUGUSTAE VINDELICORVM,
Sumptibus VIOVAE EBERHARDI KLETT
Typis CHRISTOPHORI PETRI DETLEFFSEN.
MDCCLX.

1. Si le nombre des observations faites est N et que chacune d'entre elles se produit seulement une fois, le nombre des cas possibles est le même que le nombre de permutations, c'est-à-dire : $N!$

2. Si une observation arrive p fois, le nombre des cas possible sera $\frac{N!}{p!}$.

Lambert indique alors que les deux conclusions précédentes s'obtiennent lorsque toutes les observations possibles sont également probables. Il ajoute que :

3. Si une ou plusieurs observations sont plus probables que d'autres et que cependant chacune se produit une seule fois, il faudra multiplier le nombre de cas $N!$ par les coefficients de probabilité, c'est-à-dire par les occurrences vraies.

4. Si les probabilités sont différentes et si certaines observations se produisent plusieurs fois, le nombre trouvé en 2) doit être multiplié par les occurrences vraies, élevées chacune à la puissance qui correspond à la fréquence observée.

Dans le cas des quatre observations, on a :

- le nombre des observations est $m + n + k + 1$;
- les occurrences observées sont m, n, l, k .

Donc si l'on appelle L le nombre de cas possibles :

$$L = \frac{(m+n+l+k)! \cdot PN^n \cdot QM^m \cdot RL^l \cdot SK^k}{m! \cdot n! \cdot l! \cdot k!}.$$

On voit que le nombre de cas possibles est proportionnel au produit des occurrences vraies élevées aux puissances égales aux occurrences observées.

On cherche le cas le plus probable c'est-à-dire celui dont l'occurrence est la plus grande, il faut donc que $PN^n \cdot QM^m \cdot RL^l \cdot SK^k$ soit maximal car cette quantité est proportionnelle au nombre de cas possibles et l'occurrence la plus probable sera celle pour laquelle cette quantité sera maximale. En prenant les logarithmes et en différentiant, on obtient

$$n \frac{d(PN)}{PN} + m \frac{d(QM)}{QM} + l \frac{d(RL)}{RL} + k \frac{d(SK)}{SK} = 0$$

Lambert considère alors les droites tangentes aux points N, M, L, K, et leurs sous-tangentes¹ ν , μ , λ , χ , la probabilité, c'est-à-dire le maximum de vraisemblance² sera obtenu pour :

$$\frac{n}{\nu} + \frac{m}{\mu} + \frac{1}{\lambda} + \frac{k}{\chi} = 0$$

Dans cette formule les sous-tangentes sont positives ou négatives.

En 1765, dans un ouvrage de mathématiques en trois volumes³, Lambert est revenu sur le problème de la théorie des erreurs dans un chapitre intitulé *Das Mittel zwischen den Fehlern*⁴ (la moyenne entre les erreurs). Cette fois, il formule une loi de distribution des erreurs semi-circulaire. On trouve, par exemple, une analyse des travaux Lambert sur les erreurs dans l'ouvrage de Hald⁵.

11.2. Les autres travaux de Lambert relatifs aux probabilités

Un autre mémoire relatif aux probabilités a été publié par Lambert en 1771. Il a pour titre *Examen d'une espèce de Superstition ramenée au calcul des probabilités*⁶.

Lambert y analyse le jeu de rencontre, cite Euler, retrouve les résultats de ce dernier et donne une solution générale.

Todhunter mentionne que le jeu de rencontre, ou sa variante le jeu des treize, a été discuté par Montmort, De Moivre, Euler, Lambert, Laplace et Michaelis. Voici l'énoncé du jeu par Euler⁷ :

« Le jeu de rencontre est un jeu de hasard, où deux personnes ayant chacune un jeu entier de cartes, en tirent à la fois une carte après l'autre, jusqu'à ce qu'il arrive, qu'elles rencontrent la même carte : et alors une des deux personnes gagne. Or, lorsqu'une telle rencontre n'arrive point du tout, alors c'est l'autre des deux personnes

1. La sous-tangente est la partie de l'axe d'une courbe qui est comprise entre l'ordonnée correspondante à un point de cette courbe et la tangente correspondante.
2. Si la distribution de l'erreur d'une mesure est $p(x) = f(x - \mu)$, où μ est la vraie valeur, n_i étant la fréquence de x_i et $\sum n_i = n$, aujourd'hui, l'équation du maximum de vraisemblance devant être résolue en μ , est la suivante : $\frac{d \ln L}{d \mu} = \sum \frac{n_i}{p_i} \frac{dp_i}{d \mu} = 0$, où $L = \frac{n!}{n_1! \dots n_i!} p_1^{n_1} \dots p_i^{n_i}$.

Cf. Hald A., *A history of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 79-83.

3. *Beyträge zum Gebrauche der Mathematik und deren Anwendung*, Berlin, 1765, 1770, 1772.
4. *Beyträge I*, p. 296-313.
5. Hald A., *A history of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 79-83.
6. *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles Lettres* pour 1771.
7. *Hist. de l'Acad. Berlin*, pour 1751, 1753, p. 255-270 (Eneström 201).

qui gagne ; Cela posé, on demande la probabilité, que l'une et l'autre de ces deux personnes aura de gagner. »

Le calcul est complexe. Le résultat est le suivant : Lorsque le nombre de cartes est supérieur à 12, l'espérance de l'un des joueurs est 0,6321120558 et celle de l'autre 0,367879442. Si le nombre de cartes tend vers l'infini les espérances tendent vers $1 - \frac{1}{e}$ et $\frac{1}{e}$.

Sheynin indique qu'on trouve également dans le *Neues Organon*¹ de Lambert des considérations sur l'équipossibilité qui préfigurent certains propos de Poincaré sur les conventions et sur la distinction entre des séquences d'événements gouvernés par le hasard et celles qui dérivent d'une certaine loi.

Au sujet de l'équiprobabilité, Laplace écrit dans son introduction de la deuxième édition de sa *Théorie analytique* de 1814 :

« 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 de cas favorables à l'événement dont on cherche à déterminer 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 le dénominateur est le nombre de tous les cas possibles. »

Ce principe, selon lequel des probabilités égales doivent être assignées à chacune des alternatives s'il n'y a aucune raison de préférer l'une d'elles à une autre a été appelé depuis *le principe de raison insuffisante*. Keynes² l'a appelé en 1921, *principe d'indifférence*.

11.3. L'équipossibilité selon Laplace et Poincaré

Dans son papier de 1776, *Recherches sur l'intégration des équations différentielles aux différences finies, et sur leur usage dans la théorie des hasards* Laplace passe de l'équiprobabilité à l'équipossibilité et affirme : « l'équipossibilité des événements simples est une notion fondamentale dans la théorie des probabilités. » Il indique que dans un ensemble d'événements ne pouvant pas être considérés comme également possibles, on peut chercher des sous-ensembles dans lesquels les événements le sont.

Voici, par ailleurs, les propos d'Henri Poincaré sur l'équipossibilité³ :

1. *Neues Organon*... Leipzig, 1764. On trouvera des études relatives à cet ouvrage de Lambert in : Fanfalone G., *Nouvel Organon. Phénoménologie (traduction partielle)*, Paris, Vrin, 2002 ; Martin T., « Logique du probable de Jacques Bernoulli à J. H. Lambert », *Jal élec. d'histoire des probabilités et de la statistique*, vol. 2, n°1b, 2006.
2. Keynes J. M., *A treatise of probability*, Macmillan, London, 1921, p. 41.
3. Poincaré H., *La science et l'hypothèse*, Paris, Flammarion, 1908, p. 213-216.

« 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 connaît pas ? Cependant, beaucoup de savants éminents se sont occupés de ce calcul, et l'on ne saurait nier que la science n'en ait tiré quelque profit. Comment expliquer cette apparente contradiction ?

La probabilité a-t-elle été définie ? Peut-elle même être définie ? Et, si elle ne peut l'être, comment ose-t-on en raisonner ? La définition, dira-t-on, est bien simple : la probabilité d'un événement est le rapport du nombre des cas favorables à cet événement au nombre total des cas possibles.

Un exemple simple va faire comprendre combien cette définition est incomplète. Je jette deux dés ; quelle est la probabilité pour que l'un des deux dés au moins amène un six ? Chaque dé peut amener six points différents : le nombre des cas possibles est $6 \times 6 = 36$; le nombre des cas favorables est 11. la probabilité est $11/36$.

C'est là la solution correcte. Mais ne pourrais-je pas dire tout aussi bien : Les points amenés par les deux dés peuvent former $\frac{6 \times 7}{2} = 21$ combinaisons différentes ? Parmi ces combinaisons 6 sont favorables ; la probabilité est $6/21$.

Pourquoi la première manière d'énumérer les cas possibles est-elle plus légitime que la seconde ? En tous cas, ce n'est pas notre définition qui nous l'apprend.

On est donc réduit à compléter cette définition en disant : "... au nombre total des cas possibles, pourvu que ces cas soient également probables". Nous voilà donc réduits à définir le probable par le probable.

Comment saurons-nous que deux cas possibles sont également probables ? Sera-ce par une convention ? Si nous plaçons au début de chaque problème une convention explicite, tout ira bien ; nous n'aurons plus qu'à appliquer les règles de l'arithmétique et de l'algèbre et nous irons jusqu'au bout du calcul sans que notre résultat puisse laisser place au doute ; mais dès que nous en voudrions faire la moindre application, il faudra démontrer que notre convention était légitime, et nous nous retrouverons en face de la difficulté que nous avons cru éluder.

Dira-t-on que le bon sens suffit pour nous apprendre quelle convention il faut faire ? Hélas ! M. Bertrand s'est amusé à traiter un problème simple :

"Quelle est la probabilité pour que, dans une circonférence, une corde soit plus grande que le côté du triangle équilatéral inscrit ?" L'illustre géomètre a adopté successivement deux conventions que le bon sens semblait également imposer, et il a trouvé avec l'une $1/2$ et avec l'autre $1/3$. »

Dans son ouvrage sur les probabilités Poincaré a étudié le problème des épreuves répétées¹.

Il écrit :

« Deux événements contraires A et B, ont pour probabilités respectives p et q . Ainsi une urne contient p boules blanches et q boules noires,

$$p = \frac{\mu}{\mu + \nu}, q = \frac{\nu}{\mu + \nu} ;$$

1. Poincaré H., *Calcul des probabilités*, deuxième édition, Paris, Gauthier-Villars, 1912, p. 113-117.

on en tire un très grand nombre de boules m , en remettant chaque fois la boule sortie dans l'urne. Si on a tiré α boules blanches, il y a beaucoup de chances d'après le théorème de Bernoulli que $\frac{\alpha}{mp}$ diffère peu de l'unité. »

Après calculs, Poincaré démontre le fait suivant :

Si on dispose d'observations des événements A et B consignées dans un tableau il est possible de voir si les différences observées sont dues au hasard, ou si le hasard n'intervient pas seul.

On compare, dit Poincaré, pour un certain nombre de cas, le rapport du nombre des arrivées de A à celles de B.

« Soient N et N' le nombre total des arrivées de A et de B ; on aura

$$p = \frac{N}{N+N'}, q = \frac{N'}{N+N'}.$$

Divisons maintenant le tableau en plusieurs séries ; soit, dans chacune de ces séries, m le nombre total des épreuves et

$$\alpha = mp + \lambda\sqrt{m}$$

le nombre total d'épreuves favorables à A. On calculera λ pour chaque série ; on cherchera la moyenne arithmétique de λ^2 ; si cette moyenne est égale à pq , le hasard intervient seul ; si elle est plus petite que pq , il intervient une cause indépendante du hasard. »

11.4. Biographie de Lambert (1728-1777)

La famille de Johann Heinrich Lambert est originaire de Lorraine. En 1635 cette famille s'installe à Mulhouse. Lukas Lambert, le père de notre savant était tailleur, comme l'était également son propre père. Il épouse, en 1724, Elizabeth Schmerber. Ils ont cinq fils et deux filles. Johann Heinrich Lambert naît le 26 août 1728 à Mulhouse et c'est dans cette ville qu'il ira à l'école jusqu'à l'âge de douze ans, étudiant en particulier le français et le latin. Ensuite il aide son père, mais continue d'étudier seul le soir, en particulier les sciences. Cet apprentissage autodidacte devient difficile lorsque, à l'âge de quinze ans il travaille à l'extérieur de la maison familiale pour gagner un peu d'argent. Il quitte ensuite Mulhouse pour un poste d'employé dans une ferronnerie à Seppois, au sud de Mulhouse. Ses connaissances lui permettent de devenir tuteur privé. À dix-sept ans il devient le secrétaire de Johann Rudolf Iselin, l'éditeur du journal *Basler Zeitung*. Il peut alors consacrer plus de temps à ses lectures en mathématiques, astronomie et philosophie. Lorsqu'il a vingt ans, en 1748, Lambert devient tuteur du petit-fils et du cousin (tous les deux âgés de 11 ans) du comte Peter von Salis dans la petite ville suisse de Chur. Cette position est idéale, il peut profiter de la riche bibliothèque du comte et y fabriquer ses propres instruments astronomiques. Il se fait rapidement remarquer : il est élu à la société littéraire de Chur ainsi qu'à la Société Scien-

tifique Suisse sise à Bâle. Il publie son premier article, consacré à la théorie de la chaleur, en 1755 dans la revue *Acta Helvetica*. En 1756, il emmène les deux garçons alors âgés de 19 ans, dont il a assuré l'éducation, pour un voyage en Europe. La première étape est Göttingen où il reste presque un an. Il va ensuite à Utrecht où il passe deux années au cours desquelles il visite les principales villes hollandaises. Sur place, il écrit son premier livre, publié à La Haye et consacré à l'étude de la propagation de la lumière à travers divers milieux. Le voyage se poursuit ensuite à Paris, où il rencontre d'Alembert, puis à Marseille, Nice, Turin et Milan. À son retour il décide de quitter la famille von Salis pour se consacrer pleinement à la science. Il passe quelques mois à Zürich où il fait des observations astronomiques, puis il retourne chez lui à Mulhouse. En 1759 il se rend à Augsbourg où il publie deux livres : *Photometria* et *Cosmologische Briefe*¹. C'est dans l'ouvrage *Photometria* que se trouve la fameuse loi, dite de Beer-Lambert, suivant laquelle l'intensité d'un rayon lumineux décroît exponentiellement avec la distance parcourue dans un milieu absorbant².

Preuve de la réputation grandissante de Lambert, Euler, en 1760, l'invite à être professeur d'astronomie à l'Académie des sciences de Saint-Petersbourg. Il est ensuite invité par l'Académie bavaroise des sciences, à Munich, pour réorganiser cette Académie suivant le modèle de celle de Berlin. Il rencontre beaucoup d'obstacles à ce projet et quitte Munich en 1762. Il se rend à Milan, puis à Leipzig, ville où il réussit à publier en 1764 son premier ouvrage de philosophie : *Neues Organon*. Il est alors invité par Euler à rejoindre l'Académie de Berlin où se trouve également Lagrange. Les débuts sont difficiles car Frédéric II refuse la nomination de Lambert à cause de son non-conformisme : il s'habille de façon non conventionnelle et a des comportements étranges. Toutefois, Frédéric II ne tarde pas à s'apercevoir de la brillante intelligence de notre savant et de sa perspicacité. Finalement, Lambert sera admis à l'Académie et y restera douze ans, jusqu'à sa mort. En 1766, Lambert écrit un article, *Theorie der Parallellinien*, dans lequel il étudie le postulat des parallèles. En supposant que ce postulat est faux il réussit à trouver de nombreux résultats de la géométrie non euclidienne, en particulier le fait que dans ce cas la somme des angles d'un triangle n'est pas égale à 2π . Lambert est également connu pour ses travaux, publiés en 1768, montrant que le nombre π est irrationnel. Il conjecture également que les nombres

1. Dans cet ouvrage Lambert fait œuvre de pionnier en cosmologie. Il anticipe, avec plus de cent cinquante ans d'avance, l'existence des galaxies, des super-galaxies et des systèmes plus complexes (l'équivalent de nos superamas), lesquels tournent autour de leur centre respectif qui serait constitué d'un corps opaque et très dense (l'équivalent de nos trous noirs). Notons qu'à la fin du XIX^e siècle l'univers est encore assimilé à notre galaxie et qu'il faudra attendre le premier quart du XX^e siècle pour que l'existence de galaxies externes soit démontrée.
2. Il est à noter que Bouguer avait publié trente ans plus tôt, en 1729, la loi exponentielle dans son ouvrage *Essai d'optique sur la gradation de la lumière*.

e et π sont transcendants. Il s'intéresse aux probabilités et formule, en 1772 une loi sur la mortalité. Garibaldi et Penco indiquent, à propos de la contribution de Lambert aux probabilités : « Les lignes tracées par Jacques Bernoulli dans son *Ars conjectandi* sont développées d'une manière décisive par Lambert, dont les contributions fondamentales à la théorie des erreurs ont été réévaluées ces dernières années. Les vastes activités de Lambert ont de multiples facettes, et couvrent l'optique, la cosmologie et la géodésie et l'ont mis en contact avec les principaux scientifiques et les philosophes de son temps, en particulier Kant. Dans son propre opus de philosophie fondamentale, *Neues Organon*, Lambert a développé une théorie remarquable de logique probabiliste qui, à notre connaissance, a échappé à l'attention des éminents érudits du domaine tel que Keynes. La logique probabiliste est le "troisième type général de probabilité" qui suit dans l'ordre d'exposition la "probabilité *a priori*" typique des jeux et la "probabilité *a posteriori*" que l'on trouve en statistique ».

Après une vie bien remplie, faite de voyages et de recherches profondes sur des sujets variés, Lambert s'est éteint à Berlin le 25 septembre 1777.

11.5. Principaux ouvrages de Lambert

- *Die freye Perspektive, oder Anweisung, Jeden Perspektivischen Ausrisz von freyen Stücken und ohne Grundrisz zu verfertigen*, Zürich, Heidegger und Compagnie, 1759.
- *Photometria sive de Mensura et Gradibus Luminis, Colorum et Umbrae*, Detleffsen, Augsburg, 1760.
- *Cosmologische Briefe über die Einrichtung des Weltbaues*, Augsburg, E. Kletts-Witwe, 1761.
- *Neues Organon oder Gedanken über die Erforschung und Bezeichnung des Wahren und dessen Unterscheidung vom Irrthum und Schein*, Leipzig, J. Wendler, 1764.
- *Pyrometrie*, Berlin, Haude und Spener, 1779.
- *Opera mathematica*, 2 vol., (A. Speiser éd.), Füssli, Zürich, 1946-1948.

11.6. Références

- C. J. Scriba, *J. H. Lambert*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 7, p. 595-600, 1970.
- O. B. Sheynin, *J. H. Lambert's work on probability*, AHES, 7, 1971, p. 244-256.

Laplace

La théorie des Probabilités tient à des considérations si délicates, qu'il n'est pas surprenant qu'avec les mêmes données deux personnes trouvent des résultats différents, surtout dans les questions très compliquées.

(*Essai philosophique sur les probabilités*, 1814)

Todhunter, dans son *History of the Theory of probability* (Cambridge 1865), affirme qu'au total, cette science doit plus à Laplace qu'à tout autre mathématicien. Pour sa part Poisson, en 1829, qualifie la Théorie analytique des probabilités d'*ouvrage aussi éminemment remarquable par la variété des questions qui y sont traitées, que par la généralité de méthodes que Laplace a imaginées pour les résoudre*¹.

Cet ouvrage principal de Laplace a été publié dans sa première version en 1812. Avant cette date, Laplace avait publié une douzaine de mémoires relatifs aux probabilités, et il en publia encore cinq autres après 1812².

12.1. Les premiers mémoires 1774-1776

12.1.1. *Mémoire sur les suites récurro-récurrentes et sur leurs usages dans la théorie des hasards*

Le premier papier de Laplace sur le sujet est intitulé *Mémoire sur les suites récurro-récurrentes et sur leurs usages dans la théorie des hasards*³ et fut présenté à l'Académie début février 1772.

Dans ce papier apparaît la première définition par Laplace de la notion de probabilité. Il appelle sa définition *Principe* :

1. Poisson S. D., *Annales de Chimie et de Physique*, tome XL, 1829.
2. Cf., par exemple : Todhunter I., *A History of the mathematical theory of probability*, Cambridge, 1865, chapitre XX ; Sheynin O. B., *Laplace's work on probability*, AHES, vol. 16, 1976, p. 137-187 ; Dale A. I., *A history of inverse probability*, Springer Verlag, 1991, chapitre 6.
3. *Mémoires de l'Académie Royale des Sciences de Paris (Savants étrangers)*, tome VI, p. 353-371, 1774.

« La probabilité d'un événement est égale à la somme des produits de chaque cas favorable par sa probabilité divisée par la somme des produits de chaque cas possible par sa probabilité, et si chaque cas est également probable, la probabilité de l'événement est égale au nombre de cas favorables divisé par le nombre de tous les cas possibles. »

Cette définition est identique à celle de De Moivre dans son *De mensura sortis* de 1712 et de tous les prédécesseurs de Laplace, sauf Bayes.

De nombreux problèmes de probabilités nécessitant la sommation de séries, de Moivre avait traité le problème, en 1730, dans son livre *Miscellanea Analytica*. Laplace aborde les suites récurrentes en connexion avec la solution d'une équation aux différences finies dépendant d'une seule variable. Les suites récurro-récurrentes mettent en jeu une équation aux différences finies dépendant de deux variables. Laplace applique ces outils mathématiques à trois problèmes, la durée d'un jeu, un problème de loterie et celui de savoir, si dans un tas de pièces on en prend un nombre au hasard, quelle est la probabilité que ce nombre sera pair ou impair.

12.1.2. Mémoire de la Probabilité des causes par les événements

Le second mémoire de Laplace, *Mémoire de la Probabilité des causes par les événements*¹ a également été imprimé en 1774.

Laplace énonce comme un principe (le principe de probabilité inverse²) le théorème de Bayes de la probabilité des causes prises *a priori* égales. Il écrit :

« L'incertitude des connaissances humaines porte sur les événements ou sur les causes des événements ; si l'on est assuré, par exemple, qu'une urne ne renferme que des billets blancs et noirs dans un rapport donné, et que l'on demande la probabilité qu'en prenant au hasard un de ces billets il sera blanc, l'événement alors est incertain, mais la cause dont dépend la probabilité de son existence, c'est-à-dire le rapport des billets blancs aux noirs, est connu.

Dans le problème suivant : Une urne étant supposée renfermer un nombre donné de billets blancs et noirs dans un rapport inconnu, si l'on tire un billet et qu'il soit blanc, déterminer la probabilité que le rapport des billets blancs aux noirs est celui de p à q ; l'événement est connu et la cause inconnue. [...]

Si un événement peut être produit par un nombre n de causes différentes, les probabilités de l'existence de ces causes prises de l'événement sont entre elles comme les probabilités de l'événement prises de ces causes, et la probabilité de l'existence de chacune d'elles est égale à la probabilité de l'événement prise de cette cause, divisée par la somme de toutes les probabilités de l'événement prises de chacune de ces causes.

1. *Mémoires de l'Académie des sciences de Paris (Savants étrangers)*, tome VI, p. 621-656, 1774.

2. La paternité du nom « probabilité inverse » a été attribuée par Fisher, sans doute par erreur, à De Morgan qui avait publié en 1838 *An essay on probabilities and their applications to life contingencies and insurance offices*.

La question suivante éclaircira ce principe, en même temps qu'elle en fera voir l'usage :

Je suppose que l'on me présente deux urnes A et B dont la première contienne p billets blancs et q billets noirs, et la seconde, contienne p' billets blancs et q' billets noirs ; je tire de l'une de ces urnes (j'ignore de laquelle) $f + h$ billets, dont f sont blancs et h sont noirs ; on demande, cela posé, quelle est la probabilité que l'urne dont j'ai tiré ces billets est A ou qu'elle est B.

En supposant que cette urne soit A, la probabilité d'en tirer f billets blancs et h billets noirs est :

$$\frac{(f+1)(f+2)\dots(f+h)p(p-1)\dots(p-f+1)q(q-1)\dots(q-h+1)}{1.2.3\dots h(p+q)(p+q-1)\dots(p+q-f-h+1)}$$

Soit K cette quantité ; si l'on suppose maintenant que l'urne dont j'ai tiré les billets est B, la probabilité d'en tirer f billets blancs et h billets noirs se déterminera en changeant, dans K , p et q en p' et q' ; soit K' ce que devient alors cette expression. Cela posé, les probabilités que l'urne dont j'ai tiré les billets est A ou B sont entre elles, par le principe énoncé ci dessus, comme K est à K' ; la probabilité que cette urne est A est égale à $\frac{K}{K+K'}$ et celle qu'elle est B est égale à $\frac{K'}{K+K'}$. »

Avec des notations modernes le principe exposé ci-dessus par Laplace peut s'écrire à l'aide des deux équations suivantes :

$$\frac{\Pr[A_i / E]}{\Pr[A_j / E]} = \frac{\Pr[E / A_i]}{\Pr[E / A_j]}, \quad i, j \in \{1, 2, \dots, n\}, i \neq j$$

et

$$\Pr[A_i / E] = \frac{\Pr[E / A_i]}{\sum_1^n \Pr[E / A_j]}, \quad i \in \{1, 2, \dots, n\} .$$

Laplace poursuit :

« Nous allons présentement appliquer ce principe à la résolution de quelques problèmes.

Problème I. Si une urne renferme une infinité de billets blancs et noirs dans un rapport inconnu, et que l'on en tire $p + q$ billets dont p soient blancs et q soient noirs ; on demande la probabilité qu'en tirant un nouveau billet de cette urne il sera blanc.

Solution. Le rapport du nombre de billets blancs au nombre total de billets contenus dans l'urne peut être un nombre quelconque compris depuis 0 jusqu'à 1 ; or, si l'on prend un de ces nombres x pour représenter ce nombre inconnu, la probabilité de tirer de l'urne p billets blancs et q billets noirs est, dans ce cas $x^p(1-x)^q$; partant la probabilité que x est le vrai rapport du nombre de billets blancs au nombre total des billets est par le principe de l'article précédent égale à $\frac{x^p(1-x)^q dx}{\int x^p(1-x)^q dx}$,

l'intégrale étant prise de manière qu'elle soit nulle lorsque $x = 0$, et qu'elle finisse lorsque $x = 1 \dots$ »

Laplace poursuit plus loin :

« La solution de ce problème donne une méthode directe pour déterminer la probabilité des événements futurs d'après ceux qui sont déjà arrivés [...] »

et il énonce le théorème suivant :

« On peut supposer les nombres p et q tellement grands, qu'il devienne aussi approchant que l'on voudra de la certitude que le rapport du nombre de billets blancs au nombre total de billets renfermés dans l'urne est compris entre les deux limites $\frac{p}{p+q} - \omega$ et $\frac{p}{p+q} + \omega$, ω pouvant être supposé moindre qu'aucune grandeur donnée. »

Pour démontrer ce théorème, Laplace reprend l'expression précédente donnée pour la distribution du rapport x , soit $\frac{x^p(1-x)^q dx}{\int x^p(1-x)^q dx}$ et évalue

l'intégrale de cette expression entre $\frac{p}{p+q} - \omega$ et $\frac{p}{p+q} + \omega$. Il développe en série $x^p(1-x)^q$ autour de son maximum obtenu pour $\frac{p}{p+q}$. Il note $x = \frac{p}{p+q} + z$, et obtient, avec des notations modernes :

$$P\left(\left|\frac{p}{p+q} - x\right| \leq \omega \mid p, q\right) \cong \frac{2}{\sigma\sqrt{2\pi}} \int_0^\omega e^{-\frac{z^2}{2\sigma^2}} dz.$$

Le troisième problème étudié par Laplace dans son *Mémoire de la Probabilité des causes par les événements* est le suivant : Déterminer le milieu que l'on doit prendre entre trois observations données d'un même phénomène.

La restriction à trois observations limite, à l'évidence, l'intérêt du cas étudié, mais on trouvera dans les ouvrages de Todhunter¹ (article 865) et Dale² des commentaires relatifs à cette étude.

12.1.3. Recherches sur l'intégration des équations différentielles aux différences finies, et sur leur usage dans la théorie des hasards

Ce troisième papier de Laplace³ a été publié en 1776.

Le début du mémoire reprend la théorie des séries développée dans son papier antérieur, puis on trouve une cinquantaine de pages relatives aux probabilités. On peut y lire :

1. 1865.

2. A. Dale, *A history of inverse probability?*, Springer Verlag, 1991.

3. *Mémoires de l'Académie des sciences de Paris (Savants étrangers)*, tome VII, p. 37-232, 1776.

« Il importe de fixer le sens de ces mots *hasard* et *probabilité*. Nous regardons une chose comme l'effet du hasard, lorsqu'elle n'offre à nos yeux rien de régulier, ou qui annonce un dessein, et que nous ignorons d'ailleurs les causes qui l'ont produite. Le hasard n'a donc aucune réalité en lui-même ; ce n'est qu'un terme propre à désigner notre ignorance sur la manière dont les différentes parties d'un phénomène se coordonnent entre elles et avec le reste de la Nature.

La notion de probabilité tient à cette ignorance. Si nous sommes assurés que, sur deux événements qui ne peuvent exister ensemble, l'un ou l'autre doit nécessairement arriver, et que nous ne voyons aucune raison pour laquelle l'un arriverait plutôt que l'autre, l'existence et la non-existence de chacun d'eux est également probable. Pareillement, si de trois événements qui s'excluent mutuellement, l'un doit nécessairement arriver, et que nous ne voyons aucune raison pour laquelle l'un arriverait plutôt que l'autre, leur existence est également probable, mais la non-existence de chacun d'eux est plus probable que son existence, et cela dans le rapport de 2 à 1, parce que sur trois cas également probables il y en a deux qui lui sont favorables, et un seul qui lui est contraire.

Le nombre des cas possibles restant le même, la probabilité d'un événement croît avec le nombre des cas favorables ; au contraire, le nombre des cas favorables restant le même, elle diminue à mesure que le nombre des cas possibles augmente ; en sorte qu'elle est en raison directe du nombre des cas favorables et dans l'inverse du nombre de tous les cas possibles.

La probabilité de l'existence d'un événement n'est ainsi que le rapport du nombre des cas favorables à celui de tous les cas possibles, lorsque nous ne voyons d'ailleurs aucune raison pour laquelle l'un de ces cas arriverait plutôt que l'autre. Elle peut être conséquemment représentée par une fraction dont le numérateur est le nombre des cas favorables, et le dénominateur celui de tous les cas possibles.

Semblablement, la probabilité de la non-existence d'un événement est le rapport du nombre des cas qui lui sont contraires à celui de tous les cas possibles, et doit être par conséquent exprimée par une fraction dont le numérateur est le nombre des cas contraires, et le dénominateur celui de tous les cas possibles.

Il suit de là que la probabilité de l'existence d'un événement ajoutée avec la probabilité de sa non-existence fait une somme égale à l'unité qui représente conséquemment la certitude entière, car il est visible qu'un événement doit nécessairement ou bien arriver ou manquer.

D'ailleurs, une chose arrive certainement lorsque tous les cas possibles lui sont favorables, et la fraction qui exprime sa probabilité est alors l'unité elle-même. La certitude peut donc être représentée par l'unité, et la probabilité par une fraction de la certitude ; elle peut approcher de plus en plus de l'unité, et même en différer moins que d'aucune quantité donnée ; mais elle ne peut jamais devenir plus grande. La théorie des hasards a pour objet de déterminer ces fractions, et l'on voit par là que c'est le supplément le plus heureux que l'on puisse imaginer à l'incertitude de nos connaissances.

La certitude et la probabilité, telles que nous venons de les définir, sont évidemment comparables entre elles et peuvent être soumises à un calcul rigoureux ; il n'en est pas ainsi des états différents de l'esprit humain lorsqu'il voit que tous les cas possibles favorisent un événement, ou lorsque, dans ce nombre, il en aperçoit plusieurs qui lui sont contraires. Ces deux états sont absolument incomparables, et l'on ne peut dire

du premier qu'il soit double, ou triple du second, parce que la vérité est indivisible. Il arrive ici la même chose que dans toutes les Sciences physico-mathématiques ; nous mesurons l'intensité de la lumière, les différents degrés de chaleur des corps, leurs forces, leurs résistances, etc. Dans toutes ces recherches, les causes physiques de nos sensations, et non les sensations elles-mêmes sont l'objet de l'Analyse.

La probabilité des événements sert à déterminer l'espérance ou la crainte des personnes intéressées à leur existence, et c'est sous ce point de vue que la science des hasards est une des plus utiles de la vie civile. Ce mot *espérance* a différentes acceptions : il exprime ordinairement l'état de l'esprit humain lorsqu'il doit lui arriver un bien quelconque dans certaines suppositions qui ne sont que vraisemblables. Dans la théorie des chances, l'espérance est le produit de la somme espérée par la probabilité de l'obtenir. Pour distinguer les deux acceptions de ce terme, j'appellerai la première *espérance morale*, et la seconde, *espérance mathématique*.

Concevons n personnes qui aient une égale probabilité d'obtenir la somme a , et que cette somme doive certainement appartenir à l'une d'entre elles ; la probabilité totale étant 1, ou égale à la certitude, il est visible que la probabilité de chacune de ces personnes est $1/n$, et conséquemment leur espérance mathématique a/n . C'est aussi la somme qui devrait leur revenir, si elles voulaient, sans courir les risques de l'événement, partager la somme entière a . »

12.2. Le mémoire *Sur les Probabilités*

Dans une lettre à Lagrange du 11 août 1780¹ Laplace indique : « On va imprimer incessamment un Mémoire de moi dont l'objet principal est la manière de remonter des événements aux causes. » Le volume pour 1778 de l'Histoire de l'Académie Royale des sciences paru en 1781 contient un résumé explicite² de l'article de Laplace qui cite Bayes et Price, alors que l'article lui-même intitulé *Sur les Probabilités* ne les mentionne pas.

Voici un extrait de ce résumé :

« Toutes les questions du Calcul des Probabilités peuvent se réduire à une seule hypothèse, à celle d'une certaine quantité de boules de différentes couleurs mêlées ensemble, dont on suppose qu'on tire au hasard différentes boules dans un certain ordre ou dans certaines proportions. Si on suppose connu le nombre de boules de chaque espèce, on a le calcul ordinaire des probabilités tel que les Géomètres du dernier siècle l'ont considéré : mais si l'on suppose le nombre de boules de chaque espèce inconnu, et que par le nombre de boules de chaque espèce qu'on a tirées, on veuille juger ou de la proportion du nombre de ces boules, ou de la probabilité de les tirer dans la suite suivant certaines lois, on a une nouvelle classe de problèmes. Ces questions dont il paraît que M^{rs} Bernoulli et Moivre avaient eu l'idée, ont été examinées depuis par Mrs Bayes et Price ; mais ils se sont bornés à exposer les principes qui peuvent servir à les résoudre. M. de la Place les a considérées avec plus d'étendue, et il y a appliqué l'analyse. On peut supposer le nombre de boules fini ou infini ; s'il est fini, les questions dépendent du calcul intégral aux différences

1. Œuvres de Lagrange, tome 14, page 95, Gauthier-Villars, 1892.

2. *Histoire de l'Académie Royale des sciences, année 1778* (parue en 1781), p. 43.

finies ; s'il est infini, on n'a besoin que du calcul intégral proprement dit ; Le cas du nombre infini est celui qui a lieu lorsqu'on applique les questions aux événements naturels : en effet, il est aisé de voir qu'alors elles embrassent l'immensité des temps, et que le nombre des combinaisons est infini.

Supposons qu'on ait tiré un certain nombre de boules noires et un certain nombre de boules blanches, on peut demander combien y a-t-il à parier que dans un nombre de coups donné, on tirera plus de boules blanches que de boules noires, ou en général quelle sera la probabilité des différents événements qu'on peut imaginer devoir arriver ? Si on applique ensuite ce que la théorie apprend sur ces questions abstraites à des événements naturels, comme à la proportion entre le nombre de naissances des garçons et des filles : on partira d'abord d'un fait ; par exemple, qu'il est prouvé, par une longue suite d'observations, qu'il naît à Paris un plus grand nombre de garçons que de filles et dans une certaine proportion. On peut demander alors quelle est la probabilité que, dans l'avenir, le nombre des garçons surpassera celui des filles, et cette probabilité est la même que celle de l'existence d'une cause déterminante à laquelle il faut attribuer ce phénomène, dont par conséquent il est raisonnable de rechercher les causes physiques. »

Dans son mémoire proprement dit, Laplace écrit¹ ;

« Je me propose de traiter dans ce mémoire deux points importants de l'analyse des hasards [...] : le premier a pour objet la manière de calculer la probabilité des événements composés d'événements simples dont on ignore les possibilités respectives ; l'objet du second est l'influence des événements passés sur la probabilité des événements futurs, et la loi suivant laquelle, ils nous font connaître les causes qui les ont produits. Ces deux objets qui ont beaucoup d'analogie entr'eux, tiennent à une métaphysique très délicate. »

Il ajoute, plus loin² :

« Dans l'analyse des hasards, on se propose de connaître les probabilités des événements composés suivant une loi quelconque, d'événements simples dont les possibilités sont données : celles-ci peuvent être déterminées de ces trois manières ; 1°. *a priori*, lorsque par la nature même des événements, on voit qu'ils sont possibles dans un rapport donné ; [...]. 2° *a posteriori*, en répétant un grand nombre de fois l'expérience qui peut amener l'événement dont il s'agit, et en examinant combien de fois il est arrivé ; 3° enfin par la considération des motifs qui peuvent nous déterminer à prononcer sur l'existence de cet événement ; si, par exemple, les adresses respectives de deux joueurs *A* et *B* sont inconnues, comme on n'a aucune raison de supposer *A* plus fort que *B*, on en conclut que la probabilité de *A* pour gagner une partie est 1/2. Le premier de ces moyens donne la possibilité absolue des événements ; le second la fait connaître à peu près [...], le troisième ne donne que leur possibilité relative à l'état de nos connaissances. »

Après avoir exploré les lois des tirages de boules dans des urnes, Laplace aborde le problème de la proportion des naissances de garçons et de filles. Il reprend les calculs de Bernoulli sur la loi des grands nombres, sans cependant

1. *Id.*, p. 227-332

2. *Id.*, p. 228-229

citer Bernoulli, afin de trouver un intervalle de confiance pour la probabilité de la naissance d'un garçon.

Laplace écrit :

« Lorsqu'on n'a aucune donnée *a priori* sur la possibilité d'un événement, il faut supposer toutes les possibilités, depuis zéro jusqu'à l'unité, également probables ; ainsi, l'observation pouvant seule nous instruire sur le rapport des naissances des garçons et des filles, on doit, à ne considérer la chose qu'en elle-même et abstraction faite des événements, supposer la loi de possibilité des naissances d'un garçon ou d'une fille constante depuis zéro jusqu'à l'unité, et partir de cette hypothèse dans les différents problèmes que l'on peut se proposer sur cet objet.

Supposons, par exemple, que l'on ait observé que, sur $p + q$ enfants, il est né p garçons et q filles, et que l'on cherche la probabilité P que, sur $m + n$ enfants qui doivent naître, il y aura m garçons et n filles ; si l'on nomme x la probabilité qu'un enfant qui doit naître sera un garçon, et $1 - x$ celle qu'il sera fille, en désignant

$$\frac{1.2.3...(m+n)}{1.2.3...m.1.2.3...n}$$

par λ , on aura $\lambda x^p (1-x)^q$ pour la probabilité que, sur $p + q$ enfants, il y aura p garçons et q filles [...] [cet événement sera appelé E].

Pareillement si l'on désigne par γ le produit

$$\frac{1.2.3...(p+q)}{1.2.3...p.1.2.3...q}$$

on aura $\gamma \lambda x^{p+m} (1-x)^{q+n}$ pour la probabilité que, sur $p + q$ enfants qui naîtront d'abord, il y aura p garçons et q filles, et que, sur $m + n$ enfants qui naîtront ensuite, il y aura m garçons et n filles, [événement appelé E + e]. Maintenant, x étant susceptible de prendre toutes les valeurs depuis $x = 0$ jusqu'à $x = 1$, et toutes ces valeurs étant, *a priori*, également probables, il faut, pour avoir la véritable probabilité de E, multiplier $\lambda x^p (1-x)^q$ par adx , a étant constant, et prendre l'intégrale $\lambda \int ax^p (1-x)^q dx$ (depuis $x = 0$ jusqu'à $x = 1$) ; la valeur de a se déterminera en observant que, a devant nécessairement tomber entre 0 et 1, on a $\int adx = 1$, l'intégrale étant prise depuis $x = 0$ jusqu'à $x = 1$, ce qui donne $a = 1$.

On aura, semblablement $\gamma \lambda \int x^{p+m} (1-x)^{q+n} dx$ pour la probabilité de l'événement E + e, donc la probabilité cherchée P, que sur $m + n$ enfants qui doivent naître, il y aura m garçons et n filles sera [...]

$$\frac{\gamma \int x^{p+m} (1-x)^{q+n} dx}{\int x^p (1-x)^q dx}$$

les intégrales du numérateur et du dénominateur étant prises depuis $x = 0$ jusqu'à $x = 1$. »

Après calculs, Laplace montre que si m et n sont très petits par rapport à p et q on a

$$P = \gamma \frac{p^m q^n}{(p+q)^{m+n}}.$$

Cherchant l'intervalle de confiance, Laplace appelle θ une quantité très petite et indique que la probabilité que la possibilité de la naissance d'un garçon est



M É M O I R E S
D E
MATHÉMATIQUE
E T
D E P H Y S I Q U E ,
T I R É S D E S R E G I S T R E S
de l'Académie Royale des Sciences.
Année M. DCCLXXXII.

M É M O I R E

*Sur les approximations des Formules qui sont
fonctions de très-grands nombres.*

Par M. DE LA PLACE.

ON est souvent conduit dans l'Analyse, & principalement dans celle des hasards, à des formules dont l'usage devient impossible, lorsqu'on y substitue des nombres considérables. La solution numérique des Problèmes dont
Mém. 1782. A

comprise entre $\frac{p}{p+q} - \theta$ et $\frac{p}{p+q} + \theta$, a une valeur qui différera d'autant moins de l'unité que p et q sont grands.

Après formulation analytique de cette probabilité, Laplace effectue une application numérique¹ : Entre 1745 et 1770, il est né à Paris, 251 527 garçons et 241 945 filles. Il en déduit que la probabilité que la possibilité de naissances de garçons à Paris, par rapport à celle des filles, soit supérieure ou égale à $1/2$, diffère de l'unité de moins de $1,15 \cdot 10^{-13}$.

12.3. Mémoire sur les approximations des formules qui sont fonctions de très grands nombres

Ce mémoire² a pour objet d'effectuer des transformations de certaines fonctions impliquant des grands nombres tout comme les travaux de De Moivre et de Stirling avaient permis l'évaluation du coefficient central d'un binôme élevé à une grande puissance.

Laplace écrit :

« On est souvent conduit dans l'Analyse, et principalement dans celle des hasards, à des formules dont l'usage devient impossible, lorsqu'on y substitue des nombres considérables. La solution numérique des problèmes dont elles sont la solution analytique, présente alors de grandes difficultés que l'on n'est encore parvenu à vaincre que dans des quelques cas particuliers, dont les deux principaux sont relatifs au produit des nombres naturels 1, 2, 3, 4, etc. et au terme moyen du binôme élevé à une grande puissance. Si l'on suppose cette puissance paire et égale à $2s$, ce terme sera, comme l'on sait, $\frac{2s \cdot (2s-1) \cdot (2s-2) \dots (s+1)}{1 \cdot 2 \cdot 3 \cdot 4 \dots s}$. Quoique cette expression soit

fort simple ; cependant si s est très considérable, par exemple, égal à dix mille, il devient très difficile de le réduire en nombres, à cause de la multiplicité de ses facteurs. M. Stirling est heureusement parvenu à la transformer dans des séries d'autant plus convergentes, que s est un plus grand nombre [...]. Mais la méthode de M. Stirling, fondée sur l'interpolation des suites, et sur quelques théorèmes de Wallis, laisse à désirer une méthode directe qui s'étende à toutes les fonctions composées d'un grand nombre de termes et de facteurs. »

Laplace développe ensuite les outils recherchés. Il développe pour réduire en séries convergentes les fonctions différentielles renfermant des facteurs élevés à de grandes puissances et parvient à déterminer des valeurs approchées de formules complexes.

Voici un commentaire de la partie de ce papier où il introduit, en fait, la fonction caractéristique et la formule d'inversion :

1. *Id.*, p. 277.

2. *Histoire de l'Académie Royale des sciences, année 1782* (parue en 1785), p. 1-88.

Laplace comme de Moivre développe le binôme $(1+1)^{2n}$, et établit que le rapport du terme moyen du binôme $(1+1)^{2n}$, à la somme de tous les termes tend vers $\frac{1}{\sqrt{\pi n}}$ lorsque n est très grand ; puis, au lieu de parler de tirages de boules d'une urne et de la distribution des succès d'un tel tirage, il s'intéresse à une distribution d'erreurs en considérant les deux possibilités d'erreur de taille -1 et $+1$. Ainsi les probabilités binomiales représentent les probabilités d'une erreur totale $-2n, \dots, -2, 0, 2, \dots, 2n$. Le but de Laplace était de trouver une méthode de généralisation pour obtenir une approximation pour la distribution de l'erreur totale lorsque l'erreur peut prendre plusieurs valeurs discrètes. Pour trois valeurs $-1, 0, +1$, il faut développer le trinôme $(1+1+1)^n$, et ainsi de suite pour un nombre de valeurs plus élevé.

La formule de Moivre-Stirling avait donné la solution pour le binôme. Laplace, après avoir approximé le terme central du binôme avec cette formule, ajoute¹ :

« On peut parvenir plus simplement aux résultats précédents de la manière suivante : pour cela nommons y_n le terme moyen du binôme $(1+1)^{2n}$; il est visible que ce terme est égal au terme indépendant de e^{it} dans le développement du binôme $(e^{it} + e^{-it})^{2n}$; or, si nous multiplions ce binôme par dt , et que l'on prenne ensuite l'intégrale entre $t=0$ et $t=\pi$, il est clair que cette intégrale sera égale à πy_n ; on aura donc, en substituant $2 \cos t$ au lieu de $(e^{it} + e^{-it})$,

$$y_n = \frac{2^{2n}}{\pi} \int_0^{\pi} \cos^{2n} t dt .$$

[...]. Cette méthode a l'avantage de s'étendre à la détermination du terme moyen du trinôme $(1+1+1)^n$, de celui du quadrinôme $(1+1+1+1)^{2n}$ et ainsi de suite. Considérons le trinôme $(1+1+1)^n$ et nommons y_n (le y_n sera égal au terme indépendant de e^{it} dans le développement du trinôme $(e^{it} + 1 + e^{-it})^n$, et on aura conséquemment

$$y_n = \frac{1}{\pi} \int_0^{\pi} (2 \cos t + 1)^n dt \quad »$$

Comme l'indique Hald² Laplace a ainsi introduit la fonction caractéristique et utilisé la formule d'inversion pour trouver la probabilité d'une erreur totale nulle.

En effet, soit la fonction caractéristique de la variable aléatoire x définie par

$$\psi(t) = E[e^{itx}] ,$$

(la fonction caractéristique est la transformée de Fourier de la densité).

1. *Id.*, p. 66-67 (nous citons les termes de Laplace, en changeant très légèrement ses notations).
2. *A history of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 304-306

Pour x prenant les valeurs -1 et $+1$, chacune avec une probabilité $1/2$, on a

$$\psi(t) = E[e^{itx}] = \frac{1}{2}(e^{it} + e^{-it}),$$

et pour x prenant les valeurs $-1, 0$, et $+1$, chacune avec une probabilité $1/3$,

$$\text{on a } \psi(t) = E[e^{itx}] = \frac{1}{3}(e^{it} + 1 + e^{-it}),$$

Posant $s_n = x_1 + x_2 + \dots + x_n$, la somme de n variables indépendantes et distribuées identiquement, Laplace cherche la probabilité $P(s_n = 0)$,

il trouve pour le binôme

$$P(s_{2n} = 0) = \frac{1}{\pi} \int_0^\pi \psi^{2n}(t) dt \simeq \frac{1}{\sqrt{\pi n}},$$

et pour le trinôme

$$P(s_n = 0) = \frac{1}{\pi} \int_0^\pi \psi^n(t) dt \simeq \frac{\sqrt{3}}{2\sqrt{\pi n}}.$$

Considérant la généralisation à $2a+1$ erreurs possibles uniformément distribuées sur les entiers entre $-a$ et $+a$, la fonction caractéristique devient

$$\psi(t) = E[e^{itx}] = \frac{1}{2a+1} \left(1 + 2 \sum_{x=1}^a \cos tx \right),$$

$$\text{et } P(s_n = 0) = \frac{1}{\pi} \int_0^\pi \psi^n(t) dt.$$

Laplace écrit alors, sans démonstration

$$P(s_n = 0) \simeq \frac{\sqrt{3}}{\sqrt{2\pi na(a+1)}}.$$

Laplace applique ce qui précède à la théorie des probabilités dans un travail intitulé *Suite du mémoire sur les approximations des formules qui sont fonctions de très grands nombres*¹. Dans cette importante publication, Laplace commence par donner à nouveau des définitions élémentaires relatives aux probabilités, ainsi que celle des événements composés. Il indique ensuite que si x est la possibilité d'un événement simple et $y(x)$ la probabilité de l'observation d'un résultat, la probabilité que x sera compris entre θ et θ' est $\int_\theta^{\theta'} y dx / \int_0^1 y dx$.

Il écrit :

« La possibilité de la plupart des événements simples est inconnue, et considérée *a priori*, elle nous paraît également susceptible de toutes les valeurs depuis zéro jusqu'à l'unité ; mais si l'on a observé un résultat composé de plusieurs de ces événements, la manière dont ils y entrent, rend quelques-unes de ces valeurs plus probables que les autres ; ainsi à mesure que le résultat observé se compose par le développement des événements simples, leur vraie possibilité se fait de plus en plus connaître, et

1. *Histoire de l'Académie Royale des sciences, année 1783* (parue en 1786), p. 423-467.

il devient de plus en plus probable qu'elle tombe dans des limites qui se resserrent sans cesse, et finissent par coïncider, lorsque le nombre des événements simples est infini. Pour déterminer les lois suivant lesquelles cette possibilité se découvre, nous la nommerons x ; la théorie connue des hasards donnera la probabilité du résultat observé, en fonction de x ; soit y cette fonction. Si l'on regarde les différentes valeurs de x , comme autant de causes du résultat observé, la probabilité de x sera égale à une fraction dont le numérateur est y et dont le dénominateur est la somme de toutes les valeurs de y ; en multipliant donc les deux termes de cette fraction par dx , cette probabilité sera $ydx / \int ydx$, l'intégrale du dénominateur étant prise depuis $x=0$ jusqu'à $x=1$. La probabilité que x est compris entre les deux limites $x = \theta$ et $x = \theta'$ est par conséquent égale à $\int ydx / \int ydx$, l'intégrale du numérateur étant prise depuis $x = \theta$ jusqu'à $x = \theta'$, et celle du dénominateur étant prise depuis $x = 0$ jusqu'à $x = 1$.

La valeur de x , la plus probable est celle qui rend y un maximum, nous la désignerons par a . [...].

Si les valeurs de x , considérées indépendamment du résultat observé, ne sont pas toutes également possibles, mais que leur probabilité soit exprimée par une fonction z de x ; il suffira de changer dans les formules précédentes, y dans yz ; ce qui revient à supposer toutes les valeurs de x également possibles, et à considérer le résultat observé, comme étant formé de deux résultats indépendants, dont les probabilités sont y et z . »

12.4. *Mémoire sur les approximations des formules qui sont fonctions des très grands nombres et leur application aux probabilités*

Dans ce mémoire Laplace donne la première démonstration du théorème central limite qui établit que des variables aléatoires indépendantes $x_1, \dots, x_n$, identiquement distribuées avec une densité $f(x)$, une moyenne μ , et une variance σ^2 , ont une somme $s_n = x_1 + \dots + x_n$ qui est asymptotiquement normale avec une moyenne $n\mu$ et une variance $n\sigma^2$.

Ce mémoire¹ de 1810 est essentiellement consacré à la théorie des erreurs. Après une introduction semblable à celle de son mémoire de 1782, Laplace aborde le problème de l'inclinaison de l'écliptique d'orbites de corps célestes et cherche la probabilité que l'inclinaison moyenne de n orbites soit comprise entre deux limites données.

Cette fois, Laplace exprime cette probabilité sous la forme d'une série convergente et suppose que les lois de probabilités (appelées facilités) observées des inclinaisons obéissent à des lois quelconques. Il évalue les différences finies des puissances élevées des variables à l'aide d'une technique qu'il avait publiée l'année précédente² afin de compléter ses travaux antérieurs sur les fonctions

1. *Mémoires de l'Académie des sciences*, 1^{re} série, volume X (1809) paru en 1810, p. 353-415.
2. *Mémoires sur divers points d'analyse*, Journal de l'école polytechnique, 8, cahier 15, 1809, p. 229-265.

SUPPLÉMENT AU MÉMOIRE

Sur les approximations des formules qui sont fonctions de très-grands nombres.

Par M. LAPLACE.

J'AI fait voir dans l'article VI de ce Mémoire, que si l'on suppose dans chaque observation, les erreurs positives et négatives également faciles; la probabilité que l'erreur moyenne d'un nombre n d'observations sera comprise dans les limites $\pm \frac{r}{n}$, est égale à

$$\frac{2}{\sqrt{\pi}} \cdot \sqrt{\frac{k}{2k'}} \cdot \int dr \cdot c^{-\frac{k}{2k'}} \cdot r$$

h est l'intervalle dans lequel les erreurs de chaque observation peuvent s'étendre. Si l'on désigne ensuite par $\phi\left(\frac{x}{h}\right)$ la probabilité de l'erreur $\pm x$, k est l'intégrale $\int dx \cdot \phi\left(\frac{x}{h}\right)$ étendue depuis $x = -\frac{1}{2}h$, jusqu'à $x = \frac{1}{2}h$; k' est l'intégrale $\int \frac{x^2}{h^2} \cdot dx \cdot \phi\left(\frac{x}{h}\right)$, prise dans le même intervalle: π est la demi-circonférence dont le rayon est l'unité, et c est le nombre dont le logarithme hyperbolique est l'unité.

Supposons maintenant qu'un même élément soit donné par n observations d'une première espèce, dans laquelle

génératrices. Il applique ici ce qu'il appelle un « artifice remarquable » : le passage réciproque des résultats réels aux résultats imaginaires qui permet d'obtenir les valeurs de certaines intégrales définies. Laplace arrive au théorème limite central en changeant la question relative à l'inclinaison moyenne des écliptiques en question relative à la probabilité que l'erreur moyenne d'un nombre n d'observations soit comprise dans les limites $\pm \frac{rh}{\sqrt{n}}$, où r est la somme des erreurs. Laplace débute en supposant les erreurs équiprobables dans l'intervalle h , puis suppose toute loi de distribution des erreurs acceptable et obtient pour la probabilité la formule¹

$$\frac{2}{\sqrt{\pi}} \sqrt{\frac{k}{2k'}} \int e^{-\frac{k}{2k'} r^2} dr.$$

Si on désigne par $\Phi\left(\frac{x}{h}\right)$ la probabilité de l'erreur $\pm x$, $k = \int_{-h/2}^{+h/2} \Phi\left(\frac{x}{h}\right) dx$ et $k' = \int_{-h/2}^{+h/2} \frac{x^2}{h^2} \Phi\left(\frac{x}{h}\right) dx$.

Dans la *Suite du Mémoire sur les approximations des formules qui sont fonctions des très grands nombres et leur application aux probabilités*², Laplace rappelle les formules données ci dessus et ajoute :

« Supposons maintenant qu'un même élément soit donné par n observations d'une première espèce, dans laquelle la loi de facilité des erreurs soit la même pour chaque observation, et qu'il soit trouvé égal à A par un milieu entre toutes ces observations. Supposons ensuite qu'il soit trouvé à $A + q$ par n' observations d'une seconde espèce, dans laquelle la loi de facilité des erreurs ne soit pas la même que dans la première espèce ; qu'il soit trouvé égal à $A + q'$ par n'' observations d'une troisième espèce, et ainsi de suite. On demande le milieu qu'il faut choisir entre ces divers résultats. »

Si on suppose que $A + x$ soit le résultat vrai, l'erreur du résultat moyen des observations n sera $-x$ et la probabilité de cette erreur sera par la formule qui précède. Laplace l'appelle $\Psi(-x)$. Pour le second groupe l'erreur est $\Psi(q - x)$, pour le troisième $\Psi(q' - x)$, ... La probabilité que l'erreur du premier résultat sera $-x$, et que les autres résultats s'écarteront du premier respectivement de $q, q', \dots$, sera égale au produit $\Psi(-x) \cdot \Psi(q - x) \cdot \Psi(q' - x) \dots$; donc, si l'on construit une courbe dont l'ordonnée y soit égale à ce produit, les ordonnées de cette courbe seront proportionnelles aux probabilités des abscisses et Laplace l'appelle *courbe des probabilités*. Laplace ajoute :

« Pour déterminer le point de l'axe des abscisses où l'on doit fixer le milieu entre les résultats des observations $n, n', n'', \dots$, nous observerons que ce point est celui où

1. *Mémoires de l'Académie des sciences*, 1^{re} série, volume X (1809) paru en 1810, page 386.
2. *Id.*, p. 559-565.

MÉMOIRE

SUR

LES FONCTIONS GÉNÉRATRICES,
LES INTÉGRALES DÉFINIES,

ET

LEUR APPLICATION AUX PROBABILITÉS,
ET SPÉCIALEMENT A LA RECHERCHE DU MILIEU QU'IL FAUT CHOISIR
ENTRE LES RÉSULTATS DES OBSERVATIONS.

PAR M. LE C^{TE}. LAPLACE.

~~~~~

PARIS.

BAUDOUIN, IMPRIMEUR DE L'INSTITUT IMPÉRIAL DE FRANCE.

Septembre 1811.

l'écart de la vérité que l'on peut craindre, est un *minimum* ; or de même que dans la théorie des probabilités, on évalue la perte à craindre, en multipliant chaque perte que l'on peut éprouver, par sa probabilité, et en faisant une somme de tous ces produits ; de même on aura la valeur de l'écart à craindre, en multipliant chaque écart de la vérité, ou chaque erreur abstraction faite du signe, par sa probabilité et en faisant une somme de tous ces produits. »

Laplace en déduit que la courbe des probabilités présente des aires égales de part et d'autre de l'ordonnée correspondant à l'abscisse  $z$  du point milieu à choisir. Il ajoute :

« Daniel Bernoulli, ensuite Euler et M. Gauss ont pris pour cette ordonnée la plus grande de toutes. Leur résultat coïncide avec le précédent lorsque cette plus grande ordonnée divise l'aire de la courbe en deux parties égales [...] mais, dans le cas général, il me paraît que la manière dont je viens d'envisager la chose résulte de la théorie même des probabilités. »

Laplace développe alors la probabilité  $y$ , cherche son maximum, et trouve qu'il correspond au

« minimum de la somme des carrés des erreurs de chaque résultat, multipliées respectivement par la plus grande ordonnée de la courbe de facilité des erreurs ».

Ainsi, ajoute Laplace :

« cette propriété qui n'est qu'hypothétique, lorsque l'on ne considère que des résultats donnés par une seule observation ou par un petit nombre d'observations, devient nécessaire, lorsque les résultats entre lesquels on doit prendre un milieu, sont donnés chacun par un très grand nombre d'observations, quelles que soient d'ailleurs les lois de facilité des erreurs des erreurs de ces observations. »<sup>1</sup>

## **12.5. Sur les intégrales définies et leur application aux probabilités et spécialement à la recherche du milieu qu'il faut choisir entre les résultats des observations**

Ce mémoire<sup>2</sup> publié en 1811 avait fait l'objet d'un tirage à part, daté de septembre 1811, qui comportait seulement les sept premières sections (sur huit) avec le titre suivant : *Mémoire sur les fonctions génératrices, les intégrales définies et leur application aux probabilités et spécialement à la recherche du milieu qu'il faut choisir entre les résultats des observations.*

Laplace y écrit :

1. Cf., par exemple, Hald A., *A history of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 307-317.
2. *Mémoires de l'Institut National des sciences et des arts ; sciences mathématiques et physiques*, 11, 1811, p. 279-347.

« La théorie des approximations des formules fonctions de très grands nombres est fondée sur l'expression des variables données par des équations aux différences, au moyen d'intégrales définies que l'on intègre par des approximations très convergentes ; et il y a cela de très remarquable, savoir que la quantité sous le signe intégral est la fonction génératrice de la variable exprimée par l'intégrale définie, en sorte que les théories des fonctions génératrices et des approximations des formules fonctions de très grands nombres, peuvent être considérées comme les deux branches d'un même calcul que je désigne par le nom de *calcul des fonctions génératrices*. [...]

Le calcul des fonctions génératrices est le fondement d'une théorie que je me propose de publier bientôt sur les probabilités. Les questions relatives aux événements dus au hasard se ramènent le plus souvent avec facilité à des équations aux différences : la première branche de ce calcul en fournit les solutions les plus générales et les plus simples. Mais, lorsque les événements que l'on considère sont en très grand nombre, les formules auxquelles on est conduit se composent d'une si grande multitude de termes et de facteurs, que leur calcul numérique devient impraticable. Il est alors indispensable d'avoir une méthode qui transforme ces formules en séries convergentes. C'est ce que la seconde branche du calcul des fonctions génératrices fait avec d'autant plus d'avantage que la méthode devient plus nécessaire. Par ce moyen, on peut déterminer avec facilité les limites de la probabilité des résultats et des causes, indiqués par les événements considérés en grand nombre, et les lois suivant lesquelles cette probabilité approche de ses limites, à mesure que les événements se multiplient. »

Laplace applique dans son mémoire les outils mathématiques développés à divers problèmes de calculs de probabilités. Le plus intéressant est celui intitulé *Du milieu qu'il faut choisir entre les résultats des observations* où il met en œuvre la méthode des moindres carrés qu'il présente comme suit :

« Lorsque l'on veut corriger par l'ensemble d'un grand nombre d'observations plusieurs éléments déjà connus à peu près, on s'y prend de la manière suivante. Chaque observation étant une fonction des éléments, on substitue dans cette fonction leurs valeurs approchées, augmentées respectivement de petites corrections qu'il s'agit de connaître. En développant ensuite la fonction en série, par rapport à ces corrections, et négligeant leurs carrés et leurs produits, on égale la série à l'observation, ce qui donne une première équation de condition entre les corrections des éléments. Une seconde observation fournit une équation de condition semblable, et ainsi du reste. Si les observations étaient rigoureuses, il suffirait d'en employer autant qu'il y a d'éléments ; mais, vu les erreurs dont elles sont susceptibles, on en considère un grand nombre, afin de compenser les unes par les autres ces erreurs, dans les valeurs des corrections que l'on déduit de leur ensemble. Mais de quelle manière, faut-il combiner entre elles les équations de condition pour avoir les corrections les plus précises ? C'est ici que l'analyse des probabilités peut être d'un grand secours. Toutes les manières de combiner ces équations se réduisent à les multiplier chacune par un facteur particulier et à faire une somme de tous ces produits : on forme ainsi une première équation finale entre les corrections des éléments. Un second système de facteurs donne une seconde équation finale, et ainsi de suite jusqu'à ce que l'on ait autant d'équations finales que d'éléments dont on déterminera les corrections en résolvant ces équations. Maintenant il est visible qu'il faut choisir les systèmes de

facteurs, de sorte que l'erreur moyenne à craindre *en plus* ou *en moins* sur chaque élément soit un minimum. J'entends par *erreur moyenne* la somme des produits de chaque erreur par sa probabilité. En déterminant les facteurs par cette condition, l'analyse conduit à ce résultat remarquable, savoir que, si l'on prépare chaque équation de condition de manière que son second membre soit zéro, la somme des carrés des premiers membres est un minimum, en y faisant varier successivement chaque correction. Ainsi cette méthode, que MM. Legendre et Gauss ont proposée, et qui, jusqu'à présent, ne présentait que l'avantage de fournir, sans aucun tâtonnement, les équations finales nécessaires pour corriger les éléments, donne en même temps les corrections les plus précises. »

Voici le détail du paragraphe VII. Laplace y écrit :

« Supposons maintenant qu'un même élément soit donné 1° par le résultat moyen de  $s$  observations d'un premier genre et qu'il soit, par ces observations égal à  $A$  ; 2° par le résultat moyen de  $s'$  observations d'un second genre et qu'il soit égal à  $A + q$  ; 3° par le résultat moyen de  $s''$  observations d'un troisième genre et qu'il soit égal à  $A + q'$ , et ainsi du reste. Si on représente par  $A + x$ , l'élément vrai, l'erreur du résultat des observations  $s$ , sera  $-x$  ; en supposant donc  $C$  égal à

$$\sqrt{\frac{k}{k'}} \frac{\sqrt{\sum_i p_i^2}}{2a},$$

si l'on fait usage des moindres carrés des erreurs.

[où  $\psi(x)$  est la probabilité d'une erreur  $x$  dans chaque observation,  $k = \int_0^\infty \psi(x) dx$  et  $k' = \int_0^\infty x^2 \psi(x) dx$  .]

La probabilité de cette erreur sera, en supposant  $s$  un grand nombre,

$$\frac{C}{\sqrt{\pi}} e^{-C^2 x^2} . »$$

Répétant le même calcul pour les autres observations, Laplace écrit que la probabilité que les erreurs soient  $-x$ ,  $q - x$ ,  $q' - x$ , ... est

$$\frac{C}{\sqrt{\pi}} \frac{C'}{\sqrt{\pi}} \frac{C''}{\sqrt{\pi}} \dots e^{-C^2 x^2 - C'^2 (x-q)^2 - C''^2 (x-q')^2 \dots}$$

Il ajoute :

« En la multipliant par  $dx$ , et prenant l'intégrale depuis  $x = -\infty$  jusqu'à  $x = \infty$ , on aura la probabilité que les résultats moyens des observations  $s'$ ,  $s''$ , etc., surpasseront respectivement de  $q$ ,  $q'$ , etc., le résultat moyen des observations  $s$ .

Si l'on prend l'intégrale, dans des limites déterminées on aura la probabilité que la condition précédente étant remplie, l'erreur du premier résultat sera comprise dans ces limites ; en divisant cette probabilité par celle de la condition elle-même, on aura la probabilité que l'erreur du premier résultat sera comprise dans les limites données, lorsqu'on est certain que la condition a effectivement lieu ; cette probabilité est donc

$$\frac{\int e^{-C^2 x^2 - C'^2 (x-q)^2 - C''^2 (x-q')^2 \dots} dx}{\int e^{-C^2 x^2 - C'^2 (x-q)^2 - C''^2 (x-q')^2 \dots} dx} ,$$

# THÉORIE ANALYTIQUE DES PROBABILITÉS;

PAR M. LE COMTE LAPLACE,

Chancelier du Sénat-Conservateur, Grand-Officier de la Légion d'Honneur;  
Membre de l'Institut impérial et du Bureau des Longitudes de France;  
des Sociétés royales de Londres et de Gottingue; des Académies des  
Sciences de Russie, de Danemarck, de Suède, de Prusse, de Hollande,  
d'Italie, etc.

---

PARIS,

M<sup>ME</sup> V<sup>E</sup> COURCIER, Imprimeur-Libraire pour les Mathématiques,  
quai des Augustins, n° 57.

1812.

l'intégrale du numérateur étant prise entre des limites données et celle du dénominateur étant prise depuis  $x = -\infty$  jusqu'à  $x = \infty$  . »

Finalement, Laplace montre que la correction du premier résultat, que donne avec le plus de probabilité, l'ensemble de toutes les observations  $s, s', s'',$  etc.,

$$\text{est } \frac{C'^2 \cdot q + C''^2 \cdot q' + \text{etc.}}{C^2 + C'^2 + C''^2 + \text{etc.}}$$

$$\text{et l'erreur à craindre } \frac{1}{\sqrt{\pi(C^2 + C'^2 + C''^2 + \text{etc})}}$$

dont la moitié est l'erreur à craindre en plus.

Et Laplace conclut :

« Le milieu qu'il faut choisir entre les divers résultats, est donc celui qui rend un *minimum*, la somme des carrés de l'erreur de chaque résultat, multipliée par la plus grande ordonnée de la courbe de probabilité. Ce milieu est le premier résultat A plus sa correction.

Ainsi la loi du minimum des carrés des erreurs devient nécessaire, lorsque l'on doit prendre un milieu entre des résultats donnés chacun, par un grand nombre d'observations. »

Dans l'article VIII du même mémoire Laplace montre que la distribution normale multivariée apparaît comme la distribution limite de formes linéaires des erreurs, pour  $n$  erreurs indépendantes et de lois de distribution identiques, lorsque  $n$  tend vers l'infini.

Il écrit :

« L'analyse exposée dans l'article VI peut être étendue à la correction d'un nombre quelconque d'éléments par les observations. Elle conduit toujours à ce résultat : savoir que la méthode des moindres carrés des erreurs des observations est celle qui donne sur la correction des éléments la plus petite erreur moyenne à craindre. »

Les calculs<sup>1</sup> précédents sont repris, pratiquement à l'identique, au chapitre IV du livre II de la *Théorie analytique des Probabilités*.

## 12.6. La Théorie analytique des Probabilités

### 12.6.1. Rappels

Laplace n'a pas introduit le concept de variable aléatoire<sup>2</sup> dans ses travaux.

1. Ces calculs sont détaillés in : Dale A., *A history of inverse probability*, Springer Verlag, 1991, p. 196-201 ; Sheynin O. B., *Laplace's work on probability*, AHES, vol. 16, 1976, p. 137-187 ; Sheynin O. B., *Laplace's theory of errors*, AHES, vol. 17, 1977, p. 1-61.
2. C'est Poisson qui, le premier, en 1837, a introduit le concept de variable aléatoire en considérant « une chose A d'une nature quelconque, susceptible d'un nombre  $\lambda$  de valeurs, connues ou inconnues, parmi lesquelles une seule devra avoir lieu à chaque épreuve, de

Aujourd'hui, la loi de probabilité continue d'une variable aléatoire  $X$  est représentée par une densité de probabilité  $f$ , la fonction  $f$  étant positive, intégrable, d'intégrale égale à 1, telle que :  $P(X = x) = f(x) dx$  est la probabilité que la variable prenne ses valeurs dans un intervalle élémentaire  $(x, x + dx)$ . Sa primitive est la fonction de répartition. Soit  $X$  une variable aléatoire de fonction de répartition  $F_X(x)$ . La *fonction de densité de probabilité* de  $x$ , est, par définition, la dérivée de la fonction de répartition :

$$f_X(x) = \frac{dF_X(x)}{dx}$$

De cette définition, et du fait que  $F_X(-\infty) = 0$ , on peut trouver la fonction de répartition à partir de la densité de probabilité :

$$F_X(x) = \int_{-\infty}^x f_X(u) du$$

On définit le moment d'ordre  $k$  d'une variable aléatoire  $X$ , que l'on note  $m_k$ , comme :

$$m_k = E[X^k]$$

Les moments sont utiles pour caractériser une variable aléatoire.

Le moment d'ordre un de la variable correspond à l'espérance  $\mu$ .

Le moment d'ordre deux de la variable centrée :

$$m_2 = \sigma^2 = E[(X - \mu)^2] \text{ est appelé variance.}$$

Le moment d'ordre trois de la variable centrée-réduite :

$$m_3 = E\left[\left(\frac{x - \mu}{\sigma}\right)^3\right] \text{ est appelé coefficient d'asymétrie.}$$

Le moment d'ordre quatre de la variable centrée-réduite :

$$m_4 = E\left[\left(\frac{x - \mu}{\sigma}\right)^4\right] \text{ est appelé kurtosis.}$$

La fonction génératrice est une transformée de la densité de probabilité (variable continue) ou distribution de probabilité (variable discrète). (Cette fonction a également été appelée, pour la première fois, *fonction caractéristique* par Poincaré en 1912.)<sup>1</sup> Elle se définit en trois étapes :

1. On introduit un paramètre, traditionnellement noté  $t$

2. On crée la variable aléatoire  $e^{tX}$

---

sorte que celle qui sera arrivée ou qui arrivera sera, dans cette question, l'événement observé ou l'événement futur. »

1. H. Poincaré, *Calcul des Probabilités*, 2<sup>e</sup> éd., Gauthier-Villars, 1912, p. 206.

3. On calcule l'espérance de  $e^{tX}$ , soit  $E[e^{tX}]$ . Par exemple, si  $X$  est continue, on a :

$$G_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} p(x) dx$$

où  $p(x)$  est la densité de probabilité de  $X$ .

L'intégrale portant sur  $x$ , la variable  $x$  ne figure plus dans le résultat, qui est donc une fonction de  $t$  seulement. Cette fonction est, par définition, la Fonction Génératrice des Moments de la variable  $X$  (ou de sa ddp  $p(x)$ ). C'est la transformée de Laplace de  $p(x)$ . Le moment d'ordre  $n$  de la variable aléatoire  $X$  est égal à la valeur de la dérivée d'ordre  $n$  de sa fonction génératrice pour  $t = 0$ . La fonction génératrice des moments d'une somme de variables aléatoires est le produit de leurs fonctions génératrices.

Si la variable aléatoire entière  $X$  est binomiale de paramètre  $(n; p)$  on a

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

On remarque que les nombres  $p_X(k) = P(X = k)$  sont les coefficients du polynôme de la variable  $t$  suivant :

$$G_X(t) = ((1-p) + pt)^n = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} t^k = \sum_{k=0}^n p_X(k) t^k = E(t^X).$$

On généralise l'expression comme suit :

$$G_X(t) = \sum_{k=0}^{\infty} \binom{n}{k} p_X(k) t^k = E(t^X).$$

On peut aussi travailler avec la fonction caractéristique  $E(e^{itx})$ , au lieu de la fonction génératrice.

La fonction caractéristique  $\varphi(t)$  est la valeur probable de l'exponentielle imaginaire  $e^{itx}$ ,  $x$  étant la variable étudiée et  $t$  une variable auxiliaire. Pour une loi continue où chaque intervalle  $dx$  a la probabilité  $f(x)dx$  on a

$$\varphi(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx.$$

La formule d'inversion de Fourier permet de calculer  $f(x)$  connaissant  $\varphi(t)$  :

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \varphi(t) dt. ]$$



## 12.6.2. Les premières éditions de la Théorie analytique des Probabilités

Avant d'analyser l'ouvrage fondamental de Laplace<sup>1</sup>, qui reprend souvent, mot à mot, les mémoires analysés précédemment, nous pensons qu'il est utile d'en décrire les premières éditions.

La *Théorie analytique des Probabilités* a fait l'objet de trois éditions, l'une originale, deux en partie originales, datées 1812, 1814 et 1820. De fait, le livre a été disponible sans interruption à partir de 1812 jusqu'à ce que le tirage initial ait été épuisé. En effet, les pages 4 à 444 sont identiques dans les trois éditions et n'ont fait l'objet que d'un seul tirage. La première édition comportait des erreurs pages 25, 27, 28, 37, 38, 147, 148, 304, 359, 360, 391, 392. Ces pages ont été remplacées, dès la deuxième édition, par des cartons signés 4\*, 5\*, 19\*, 22\*, 23\*, 38\*, 45\*, 49\*.

La première édition comportait une dédicace à Napoléon le Grand. Cette dernière a été remplacée, à partir de la seconde édition de 1814, par un avertissement. Laplace avait connu Napoléon à l'École Militaire, en septembre 1785, alors qu'il était lui-même examinateur en mathématiques et que l'artilleur Buonaparte passait l'examen de sortie. Le 25 décembre 1797, Napoléon fut élu à l'Institut, dans la section de mécanique, à la place de Lazare Carnot, en exil suite au coup d'état de fructidor. Laplace accompagna Berthollet dans l'escorte qui installa Bonaparte dans son siège. En octobre 1799, trois semaines avant le coup d'état du 18 brumaire, Laplace lui offrit une copie des deux premiers volumes de la *Mécanique céleste*. Le remerciement de Bonaparte est célèbre : il promet de les lire dans les six premiers mois dont il disposerait librement et invita Laplace et son épouse à dîner le lendemain « s'ils n'avaient rien de mieux à faire »<sup>2</sup>. À l'Institut, Laplace et Bonaparte servirent dans la même commission avec Lacroix afin de rédiger un rapport sur un mémoire de Biot. Laplace ne fut jamais, comme Monge et Berthollet un favori de Bonaparte, mais, en 1807 et 1808, sa sœur Élisabeth, qui avait été élevée au rang de princesse, s'attacha les services de Madame Laplace à sa cour de Lucca. Après le coup d'état du 18 brumaire (9 novembre 1799) Bonaparte avait nommé Laplace ministre de l'Intérieur mais le remplaça au bout de quelques mois, par son propre frère Lucien. « Laplace, dit Napoléon, ne saisissait aucune question sous son véritable point de vue ; il cherchait des subtilités partout, n'avait que des idées problématiques, et portait enfin

1. De très nombreuses analyses de la *Théorie analytique des Probabilités* ont été publiées. Par exemple : Todhunter I., *History of the Theory of probability*, Cambridge 1865 ; Sheynin O. B., *Laplace's work on probability*, AHES, 1976, vol. 16, p. 137-187 ; Dale A., *A history of inverse probability*, Springer Verlag, 1991 ; Hald A., *A history of mathematical statistics from 1750 to 1930*, 1998.

2. Correspondance de Napoléon I<sup>er</sup>, 27 vendémiaire an VIII (19 octobre 1799), n° 4384, 6, 1861, p. 1.

l'esprit des infiniment petits jusque dans l'administration »<sup>1</sup>. Désireux de garder Laplace à son service, Napoléon le nomma sénateur et même vice-président du Sénat en 1803. Il le fit Comte d'Empire en 1806. En 1813, il le nomma dans l'ordre de La Réunion. Mais, souple et servile en politique, Laplace vota la déchéance de Napoléon début avril 1814. Prudent, il s'absenta de Paris durant les Cent jours. Il fut élu à l'Académie française en 1816 et, en 1817, Louis XVIII le fit marquis et pair de France.

La deuxième édition du livre comporte en introduction l'*Essai sur les probabilités* également publié séparément en 1814. La troisième édition reproduit en introduction la quatrième édition du même essai. Quatre suppléments parurent après la première édition ; Todhunter et Pearson indiquent que le premier et le second supplément parurent entre 1812 et 1820. le troisième en 1820 et le quatrième après 1820, car les premiers exemplaires de la troisième édition ne le comportaient pas. De fait, le livre resta en vente en permanence. La première édition comporte toujours l'indication « par M. le Comte Laplace » et la dédicace à Napoléon. La seconde édition, vendue de 1814 à 1820, se trouve toujours avec un avertissement à la place de la dédicace. Selon l'année de la mise en vente, les exemplaires de seconde édition comportent le premier ou les deux premiers suppléments. La troisième édition, datée 1820, porte, au titre, l'indication : « par M. le Marquis de Laplace ». Comme nous l'avons indiqué, seule l'édition originale comporte la dédicace à Napoléon. Nous la reproduisons ici :

« À Napoléon-le-Grand

Sire, La bienveillance avec laquelle Votre Majesté a daigné accueillir l'hommage de mon *Traité de Mécanique Céleste*, m'a inspiré le désir de Lui dédier cet Ouvrage sur le Calcul des Probabilités. Ce calcul délicat s'étend aux questions les plus importantes de la vie, qui ne sont en effet, pour la plupart, que des problèmes de probabilité. Il doit, sous ce rapport, intéresser Votre Majesté dont le génie sait si bien apprécier et si dignement encourager tout ce qui peut contribuer au progrès des lumières et de la prospérité publique. J'ose La supplier d'agréer ce nouvel hommage dicté par la plus vive reconnaissance, et par les sentimens profonds d'admiration et de respect, avec lesquels je suis, Sire, de Votre Majesté, le très humble et très obéissant serviteur et fidèle sujet, Laplace. »

### 12.6.3. L'Essai sur les probabilités

L'introduction de la Théorie analytique, à partir de la seconde édition est constituée, comme nous l'avons déjà dit, par « l'Essai sur les probabilités » également publié séparément en 1814.

1. Correspondance de Napoléon I<sup>er</sup>, 30, 1870, p. 330.

**ESSAI PHILOSOPHIQUE**  
**sur**  
**LES PROBABILITÉS;**  
**PAR M. LE COMTE LAPLACE,**

Chancelier du Sénat-Conservateur, Grand-Officier de la Légion d'Honneur;  
Grand-Croix de l'Ordre de la Réunion; Membre de l'Institut impérial et  
du Bureau des Longitudes de France; des Sociétés royales de Londres  
et de Gottingue; des Académies des Sciences de Russie, de Danemarck,  
de Suède, de Prusse, d'Italie, etc.

---

**PARIS,**  
**M<sup>re</sup> V<sup>e</sup> COURCIER, Imprimeur-Libraire pour les Mathématiques,**  
**quai des Augustins, n° 57.**

**1814.**

Ce texte contient, dès les premières pages, l'affirmation du déterminisme universel dans toute sa rigueur. Laplace l'énonce comme suit, son déterminisme étant synonyme de prédictibilité :

« Nous devons envisager l'état présent de l'univers comme l'effet de son état antérieur, et comme la cause de celui qui va suivre. Une intelligence qui, pour un instant donné, connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome : rien ne serait incertain pour elle, et l'avenir, comme le passé, serait présent à ses yeux. L'esprit humain offre, dans la perfection qu'il a su donner à l'astronomie, une faible esquisse de cette intelligence. Ses découvertes en mécanique et en géométrie, jointes à celles de la pesanteur universelle, l'ont mis à portée de comprendre dans les mêmes expressions analytiques les états passés et futurs du système du monde. En appliquant la même méthode à quelques autres objets de ses connaissances, il est parvenu à ramener à des lois générales les phénomènes observés, et à prévoir ceux que les circonstances données doivent faire éclore. »

Laplace définit, plus loin, les principes généraux des Probabilités :

Le premier de ces principes définit la probabilité comme le rapport du nombre des cas favorables à celui de tous les cas possibles supposés tous également possibles.

Le second principe indique que si tous les cas ne sont pas équiprobables, on déterminera d'abord leurs possibilités respectives. Alors la probabilité sera la somme des possibilités de chaque cas favorable.

Le troisième principe est le suivant :

« Si les événements sont indépendants les uns des autres, la probabilité de l'existence de leur ensemble est le produit de leurs probabilités particulières. »

Le quatrième indique :

« Quand deux événements dépendent l'un de l'autre ; la probabilité de l'événement composé est le produit de la probabilité du premier événement, par la probabilité que cet événement étant arrivé, l'autre aura lieu. »

Le cinquième principe dit que :

« Si l'on calcule *a priori*, la probabilité de l'événement arrivé, et la probabilité d'un événement composé de celui-ci et d'un autre qu'on attend ; la seconde probabilité divisée par la première sera la probabilité de l'événement attendu, tirée de l'événement observé. »

Le sixième observe :

« Chacune des causes, auxquelles un événement observé peut être attribué, est indiquée avec d'autant plus de vraisemblance, qu'il est plus probable que, cette cause étant supposée exister, l'événement aura lieu ; la probabilité de l'existence d'une quelconque de ces causes est donc une fraction dont le numérateur est la probabilité de l'événement, résultante de cette cause, et dont le dénominateur est la somme des probabilités semblables relatives à toutes les causes : si ces diverses causes considérées *a priori* sont inégalement probables, il faut au lieu de la probabilité de l'événement, résultante de chaque cause, employer le produit de cette probabilité, par la possibilité de la cause elle-même. C'est le principe fondamental de cette branche de l'Analyse des hasards, qui consiste à remonter des événements aux causes. »

Le septième :

« La probabilité d'un événement futur est la somme des produits de la probabilité de chaque cause, tirée de l'événement observé, par la probabilité que, cette cause existant, l'événement futur aura lieu. »

Laplace ajoute :

« Quand la probabilité d'un événement simple est inconnue, on peut lui supposer également toutes les valeurs depuis zéro jusqu'à l'unité. La probabilité de chacune de ces hypothèses, tirée de l'événement observé, est par le sixième principe une fraction dont le numérateur est la probabilité de l'événement dans cette hypothèse, et dont le dénominateur est la somme des probabilités semblables relatives à toutes les hypothèses. Ainsi, la probabilité que la possibilité de l'événement est comprise dans des limites données, est la somme des fractions comprises dans ces limites. Maintenant, si l'on multiplie chaque fraction par la probabilité de l'événement futur, déterminée dans l'hypothèse correspondante, la somme des produits relatifs à toutes les hypothèses sera par le septième principe la probabilité de l'événement futur, tirée de l'événement observé. »

Laplace définit ensuite l'espérance mathématique par « le produit de la somme espérée, par la probabilité de l'obtenir. » Il ajoute dans son huitième principe que lorsque l'espérance mathématique dépend de plusieurs événements, on l'obtient en prenant la somme des produits de chaque événement par le bien attaché à son arrivée.

Cette même introduction s'achève par une *notice historique sur les probabilités*. Après avoir évoqué les travaux de Pascal, Fermat, Huygens, Bernoulli, Montmort, de Moivre et Bayes, Laplace décrit le contenu de son ouvrage :

« En généralisant l'Analyse et les problèmes concernant les probabilités, je fus conduit au Calcul des différences finies partielles, que Lagrange a traité depuis par une méthode fort simple et dont il a fait d'élégantes applications à ce genre de problèmes. La théorie des fonctions génératrices, que je donnai vers le même temps, comprend ces objets parmi ceux qu'elle embrasse, et s'adapte d'elle-même et avec la plus grande généralité aux questions de probabilité les plus difficiles. Elle détermine encore, par des approximations très convergentes, les valeurs des fonctions composées d'un grand nombre de termes et de facteurs, et, en faisant voir que la racine carrée

du rapport de la circonférence au rayon entre le plus souvent dans ces valeurs, elle montre qu'une infinité d'autres transcendentes peuvent s'y introduire.

On a encore soumis au calcul la probabilité des témoignages, les votes et les décisions des assemblées électorales et délibérantes et les jugements des tribunaux. Tant de passions, d'intérêts divers et de circonstances compliquent les questions relatives à ces objets qu'elles sont presque toujours insolubles. Mais la solution de problèmes plus simples et qui ont avec elles beaucoup d'analogie peut souvent répandre sur ces questions difficiles et importantes de grandes lumières, que la sûreté du calcul rend toujours préférables aux raisonnements les plus spécieux.

L'une des plus intéressantes applications du Calcul des Probabilités concerne les milieux qu'il faut choisir entre les résultats des observations. Plusieurs géomètres s'en sont occupés, et Lagrange a publié dans les *Mémoires de Turin* une belle méthode pour déterminer ces milieux, quand la loi des erreurs des observations est connue. J'ai donné pour le même objet une méthode fondée sur un artifice singulier, qui peut être employé avec avantage dans d'autres questions d'Analyse, et qui, en permettant d'étendre indéfiniment dans tout le cours d'un long calcul des fonctions qui doivent être limitées par la nature du problème, indique les modifications que chaque terme du résultat final doit recevoir en vertu de ces limitations. On a vu précédemment que chaque observation fournit une équation de condition, du premier degré, qui peut toujours être disposée de manière que tous ses termes soient dans le premier membre, le second étant zéro. L'usage de ces équations est une des causes principales de la grande précision de nos Tables astronomiques, parce que l'on a pu ainsi faire concourir un nombre immense d'excellentes observations à la fixation de leurs éléments. Lorsqu'il n'y a qu'un seul élément à déterminer, Cotes avait prescrit de préparer les équations de condition de sorte que le coefficient de l'élément inconnu fût positif dans chacune d'elles, et d'ajouter ensuite toutes ces équations, pour former une équation finale d'où l'on tire la valeur de cet élément. La règle de Cotes fut suivie par tous les calculateurs. Mais quand il fallait déterminer plusieurs éléments, on n'avait aucune règle fixe pour combiner les équations de condition de manière à obtenir les équations finales nécessaires : seulement, on choisissait pour chaque élément les observations les plus propres à le déterminer. Ce fut pour obvier à ces tâtonnements que Legendre et Gauss imaginèrent d'ajouter les carrés des premiers membres des équations de condition, et d'en rendre la somme un minimum, en y faisant varier chaque élément inconnu : par ce moyen, on obtient directement autant d'équations finales qu'il y a d'éléments. Mais les valeurs déterminées par ces équations méritent-elles la préférence sur toutes celles que l'on peut obtenir par d'autres moyens ? C'est ce que le Calcul des Probabilités pouvait seul apprendre. Je l'appliquai donc à cet objet important, et je parvins, par une analyse délicate, à une règle qui renferme la précédente, et qui réunit, à l'avantage de donner par un procédé régulier les éléments cherchés, celui de les faire sortir avec le plus d'évidence de l'ensemble des observations, et d'en déterminer les valeurs qui ne laissent à craindre que les plus petites erreurs possibles.

On n'a cependant encore qu'une connaissance imparfaite des résultats obtenus, tant que la loi des erreurs dont ils sont susceptibles n'est pas connue ; il faut pouvoir assigner la probabilité que ces erreurs sont contenues dans des limites données, ce qui revient à déterminer ce que j'ai nommé *poids* d'un résultat. L'Analyse conduit à des formules générales et simples pour cet objet. J'ai appliqué cette analyse aux

résultats des observations géodésiques. Le problème général consiste à déterminer les probabilités que les valeurs d'une ou de plusieurs fonctions linéaires des erreurs d'un très grand nombre d'observations sont renfermées dans des limites quelconques. La loi de possibilité des erreurs des observations introduit dans les expressions de ces probabilités une constante, dont la valeur semble exiger la connaissance de cette loi presque toujours inconnue. Heureusement, cette constante peut être déterminée par les observations mêmes. [...]

J'ai rassemblé toutes ces méthodes dans ma *Théorie analytique des Probabilités*, ou je me suis proposé d'exposer, de la manière la plus générale, les principes et l'analyse du Calcul des Probabilités, ainsi que les solutions des problèmes les plus intéressants et les plus difficiles que ce Calcul présente.

On voit, par cet Essai, que la théorie des probabilités n'est, au fond, que le bon sens réduit au calcul ; elle fait apprécier avec exactitude ce que les esprits justes sentent par une sorte d'instinct, sans qu'ils puissent souvent s'en rendre compte. Si l'on considère les méthodes analytiques auxquelles cette théorie a donné naissance, la vérité des principes qui lui servent de base, la logique fine et délicate qu'exige leur emploi dans la solution des problèmes, les établissements d'utilité publique qui s'appuient sur elle, et l'extension qu'elle a reçue et qu'elle peut recevoir encore par son application aux questions les plus importantes de la Philosophie naturelle et des Sciences morales ; si l'on observe ensuite que, dans les choses mêmes qui ne peuvent être soumises au calcul, elle donne les aperçus les plus sûrs qui puissent nous guider dans nos jugements, et qu'elle apprend à se garantir des illusions qui souvent nous égarent, on verra qu'il n'est point de science plus digne de nos méditations et qu'il soit plus utile de faire entrer dans le système de l'instruction publique. »

#### **12.6.4. Le livre I de la Théorie analytique des Probabilités**

Ce livre premier traite des fonctions génératrices à une et deux variables et leurs applications à la solution d'équations aux différences et à l'approximation de diverses intégrales définies. Laplace reproduit ici le mémoire publié dans l'Histoire de l'Académie Royale des Sciences pour 1779, paru en 1782 sous le titre *Mémoire sur les suites*. Le traité est ensuite consacré aux approximations des formules fonctions des grands nombres déjà exposées en 1782 et parues en 1785.

#### **12.6.5. Le livre II de la Théorie analytique des Probabilités**

Le livre II a pour titre *Théorie générale des probabilités*.

Le chapitre I du livre II donne les théorèmes d'addition et de multiplication des probabilités pour les événements indépendants et introduit le théorème de Bayes, considéré comme un principe.<sup>1</sup>

1. Il est à noter que le mathématicien Paolo Ruffini (1765-1822) a publié en 1821 un ouvrage intitulé *Riflessioni critiche sopra il saggio filosofico intorno alle probabilità del Sig. Laplace* dans laquelle il estime que Laplace utilise parfois d'une façon fautive le théorème de Bayes. Ruffini considère, par ailleurs, dans ce livre, la possibilité que la vie résulte du hasard, anticipant ainsi des idées modernes.

[Comme nous l'avons déjà indiqué, le théorème de Bayes est une conséquence de la loi de composition des probabilités (qui est un des axiomes fondamentaux de la théorie des probabilités). Si  $A$  et  $B$  sont deux événements, cette loi de composition des probabilités indique que la probabilité  $P(AB)$  d'observer à la fois  $A$  et  $B$  est donnée par :

$$P(AB) = P(A) P(B|A) = P(B) P(A|B)$$

où  $P(A|B)$  est la probabilité d'observer  $A$  sachant que  $B$  s'est réalisé. On en déduit :

$$P(B|A) = P(B) P(A|B)/P(A)$$

qui est le théorème de Bayes. Ce théorème se généralise sans peine au cas de plusieurs événements.]

Laplace définit également, dans ce premier chapitre du livre II, les espérances mathématique et morale. Sur ce dernier point, il reprend, mot pour mot, son texte publié dans son mémoire *Recherches sur l'intégration des équations différentielles aux différences finies, et sur leur usage dans la théorie des hasards* de 1776 et que nous avons reproduit plus haut. On notera que son concept d'espérance morale est identique à celui proposé par Daniel Bernoulli.

Le chapitre II présente quelques problèmes de probabilités d'événements composés, dont l'un est relatif à un échantillonnage avec remplacement.

Le chapitre III est relatif au théorème Moivre-Laplace :

Les premières lignes du chapitre indiquent ce que Laplace entend par théorème-limite :

« À mesure que les événements se multiplient, leurs probabilités respectives se développent de plus en plus ; leurs résultats moyens et les bénéfices ou les pertes qui en dépendent convergent vers des limites dont ils approchent avec des probabilités toujours croissantes. La détermination de ces accroissements et de ces limites est une des parties les plus intéressantes et les plus délicates de l'analyse des hasards. »

De Moivre avait développé son théorème en développant le binôme  $(a + b)^n$  ou plutôt  $(1 + 1)^n$ . L'extension des résultats faite par Laplace porte quelquefois le nom de théorème de Moivre-Laplace :

Pour une variable aléatoire de Bernoulli  $X_n$  suivant une loi binomiale de probabilité de succès  $p$  en  $n$  tirages, la loi appelée *Moivre-Laplace* s'écrit pour  $a < b$  :

$$\lim_n P\left(a < \frac{X_n - np}{\sqrt{np(1-p)}} < b\right) = \Phi(b) - \Phi(a) \text{ avec}$$

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du.$$



Laplace traite ensuite dans le chapitre IV la probabilité des erreurs des résultats moyens d'un grand nombre d'observations, et de la méthode des moindres carrés. Le but est « de rechercher le résultat moyen que des observations nombreuses et non faites encore, doivent indiquer avec le plus d'avantage et la loi de probabilité des erreurs de ce résultat. » Il prouve d'une façon non rigoureuse le théorème de la limite centrale pour les sommes d'erreurs indépendantes, distribuées uniformément ainsi que pour leurs carrés.

Voici, avec des notations actuelles, quelques détails du calcul de Laplace<sup>1</sup> :

Soient  $x_1, \dots, x_n$  des variables aléatoires indépendantes identiquement distribuées avec la fréquence  $f(x)$ , la moyenne  $\mu$ , et la variance  $\sigma^2$ ,  $0 < \sigma^2 < \infty$ .

Le théorème central limite indique que la somme  $s_n = x_1 + \dots + x_n$  est asymptotiquement normale avec la moyenne  $n\mu$ , et la variance  $n\sigma^2$ .

Pour la distribution de la somme de variables aléatoires discrètes uniformément distribuées sur les entiers de  $-a$  à  $+a$ , on a la fonction caractéristique de la variable aléatoire  $x$  définie par

$$\psi(t) = E[e^{itx}],$$

(la fonction caractéristique est la transformée de Fourier de la densité) :

$$\begin{aligned} \psi(t) &= \frac{1}{2a+1} \sum_{x=-a}^a e^{itx} = \frac{1}{2a+1} \left( 1 + 2 \sum_{x=1}^a \cos tx \right) = \\ &= \frac{1}{2a+1} \left( 1 + 2 \cos \left[ \frac{1}{2}(a+1)t \right] \frac{\sin \left( \frac{1}{2}at \right)}{\sin \left( \frac{1}{2}t \right)} \right) = \\ &= \frac{1}{2a+1} \frac{\sin \left[ \frac{1}{2}(2a+1)t \right]}{\sin \left( \frac{1}{2}t \right)}. \end{aligned}$$

Ce résultat se démontrant à l'aide de la formule

$$2 \cos \alpha \sin \beta = \sin(\alpha + \beta) - \sin(\alpha - \beta).$$

Laplace indique alors que la probabilité  $P(s_n = s) = c_s$  est le coefficient de  $e^{its}$  dans le développement de  $\psi^n(t)$ , donc

$$\psi^n(t) = \sum_{k=-na}^{na} c_k e^{itk}.$$

1. *Théorie analytique des probabilités*, 1812, Livre II, chapitre IV, p. 304.

Le coefficient  $c_s$  peut aussi être trouvé comme le terme indépendant de  $t$  dans le développement de  $e^{-its}\psi^n(t)$  ou comme le terme indépendant de  $t$  dans le développement de

$$\frac{1}{2}(e^{its} + e^{-its})\psi^n(t) = \psi^n(t) \cos st.$$

Comme  $c_{-s} = c_s$ , on peut écrire

$$\frac{1}{2}(e^{its} + e^{-its})\psi^n(t) = \frac{1}{2} \sum_k c_k (e^{it(s+k)} + e^{-it(s-k)}) ,$$

Laplace, utilisant le fait que  $\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{itr} dt$  est égal à 0 lorsque  $r \neq 0$  et est égal à 1 lorsque

$r = 0$ , obtient la formule d'inversion

$$c_s = \frac{1}{4\pi} \int_{-\pi}^{\pi} (e^{its} + e^{-its})\psi^n(t) = \frac{1}{\pi} \int_{-\pi}^{\pi} \psi^n(t) \cos st dt.$$

Il démontre ensuite que le terme principal de  $c_s$  est approximativement égal à  $\frac{1}{\sigma\sqrt{2\pi n}} e^{-s^2/n\sigma^2}$ , les termes omis étant de l'ordre de  $1/n$  fois le terme principal. Cela signifie que  $s_n$  est asymptotiquement normal avec une moyenne nulle et une variance  $n\sigma^2 = \frac{na(a+1)}{3}$ .

Les résultats contenus dans les paragraphes VI à VIII de *Sur les intégrales définies et leur application aux probabilités et spécialement à la recherche du milieu qu'il faut choisir entre les résultats des observations*<sup>1</sup> sont repris, pratiquement à l'identique, dans ce chapitre IV du livre II de la *Théorie Analytique*.

Le chapitre V est intitulé *De l'application des probabilités à la recherche des phénomènes et de leurs causes*, Laplace y applique les résultats de la théorie à la variation diurne des baromètres et à la rotation de la terre. Il remarque que la théorie peut être appliquée à l'économie et à la médecine. Il discute également le problème de Buffon relatif à la chute d'une aiguille de longueur donnée tombant sur un parquet divisé en petits rectangles et à la détermination de la probabilité que l'aiguille tombe sur un joint des rectangles.

Le chapitre VI traite de la *probabilité des causes et des événements futurs tirée des événements observés* en utilisant les méthodes de Bayes. Il reprend le texte de son exposé de 1782 (*Mémoire sur les approximations des formules qui sont fonctions de très grands nombres*) et indique que si  $x$  est la possibilité d'un événement simple et  $y(x)$  la probabilité de l'observation d'un résultat, la probabilité que

1. *Mémoires de l'Institut National des sciences et des arts ; sciences mathématiques et physiques*, 11, 1811, p. 279-347.

$x$  sera compris entre  $\theta$  et  $\theta'$  est  $\int_{\theta}^{\theta'} y dx / \int_0^1 y dx$ . Il traite ensuite le cas d'un événement composé de deux événements indépendants<sup>1</sup>.

Laplace aborde ensuite le problème du rapport du nombre des naissances des garçons et des filles que nous avons exposé plus haut dans un paragraphe spécifique.

Dans son chapitre VII, Laplace étudie *l'influence des inégalités inconnues* entre des probabilités supposées égales.

Le chapitre VIII utilise les statistiques sociales et traite des durées moyennes de la vie, des mariages et des associations, de l'espérance morale et de la probabilité des témoignages.

Le chapitre IX intitulé *Des bénéfices dépendants de la probabilité des événements futurs*, Laplace considère des problèmes liés à des annuités.

Dans le chapitre X il définit, à nouveau, son concept d'espérance morale et introduit un principe selon lequel, «  $x$  étant la fortune physique d'un individu, l'accroissement  $dx$  qu'elle reçoit, produit à l'individu un bien moral réciproque à cette fortune ; l'accroissement de sa fortune morale peut donc être exprimé par  $\frac{k dx}{x}$ ,  $k$  étant une constante. Ainsi en désignant par  $y$  la fortune morale correspondante à la fortune physique  $x$ , on aura

$$y = k \cdot \log x + \log h,$$

$h$  étant une constante. »

Après analyse, Laplace démontre qu'« il vaut mieux exposer sa fortune par parties, à des dangers indépendants les uns des autres, que de l'exposer tout entière au même danger. »

L'édition originale de *la Théorie Analytique* de 1812 ne comporte que ces dix chapitres pour le livre II.

Le chapitre XI, ajouté à partir de la deuxième édition de 1814 traite de la probabilité des témoignages.

À la page 460 de cette édition, on peut lire :

« On peut assimiler le jugement d'un tribunal qui prononce entre deux opinions contradictoires, au résultat des témoignages de plusieurs témoins de l'extraction d'un numéro d'une urne qui ne contient que deux numéros. En exprimant par  $p$  la

1. Todhunter dans son *History of the Theory of probability* (Cambridge, 1865), a donné une analyse détaillée des calculs de Laplace, ainsi que A. Dale dans *A history of inverse probability* (Springer Verlag, 1991), de même que beaucoup d'autres commentateurs.

probabilité que le juge prononce la vérité ; la probabilité de la bonté d'un jugement rendu à l'unanimité sera

$$\frac{p^r}{p^r + (1-p)^r}$$

$r$  étant le nombre des juges. On peut déterminer  $p$  par l'observation du rapport des jugements rendus à l'unanimité par le tribunal, au nombre total des jugements. Lorsque ce nombre est très grand ; en le désignant par  $n$ , et par  $i$  le nombre de jugements rendus à l'unanimité ; on aura à fort peu près

$$p^r + (1-p)^r = \frac{i}{n} ;$$

la résolution de cette équation donnera la véracité  $p$  des juges »

Après quelques calculs complémentaires, Laplace conclut, sans hésiter :

« Il y a donc un grand avantage à former des tribunaux d'appel composés d'un grand nombre de juges choisis parmi les personnes les plus éclairées. »

Laplace a complété ses travaux sur ce sujet dans son *Essai philosophique des probabilités*<sup>1</sup>, mais seulement, semble-t-il à partir des éditions de cet essai qui sont postérieures à 1816. Cet essai sert d'introduction à la *Théorie Analytique* à partir de 1814. Dans la deuxième édition de la *Théorie Analytique* de 1814, l'introduction comporte *cvi* pages. Elle en comporte *cxlj* dans la troisième édition de la *Théorie Analytique* de 1820, et Laplace y indique, pour la première fois, ses opinions sur le problème. On comprend qu'il préfère protéger la société contre les dangers sociaux que de protéger les droits des citoyens. Dans le premier supplément à la *Théorie Analytique*, intitulé *Sur l'application du calcul des probabilités à la philosophie naturelle* Laplace a rédigé un paragraphe relatif à *La probabilité des jugements des tribunaux*.

En fait, l'on trouve dans la troisième édition de la *Théorie Analytique* de 1820, des phrases totalement identiques dans l'Introduction et dans le premier supplément.

On peut y lire à la page 26 du supplément et à la page *lxxxvij* de l'introduction :

« Nous devrions nous abstenir de juger, s'il nous fallait attendre l'évidence mathématique. Mais lorsque les preuves ont une force telle que le produit de l'erreur à craindre, par sa faible probabilité, soit inférieure au danger qui résulterait de l'impunité au crime ; le jugement est commandé par l'intérêt de la société. »

Pour le reste, Laplace entreprend d'examiner la probabilité de l'exactitude d'un témoignage puis celle de la valeur (qu'il appelle bonté) d'un jugement à partir du nombre de juges, et de la probabilité qu'un juge prononce la vérité.

Il écrit pages *lxxxvii* et *lxxxix* de l'introduction :

1. L'édition originale de l'*Essai philosophique des probabilités* date de 1814. Elle comporte 97 pages dans le format in-quarto.

« La probabilité que l'opinion de chaque juge est juste, entre comme élément principal dans ce calcul [...]. Je suppose ainsi qu'elle peut varier depuis  $1/2$  jusqu'à l'unité, mais qu'elle ne peut être en dessous de  $1/2$ . Si cela n'était pas, la décision du tribunal serait insignifiante comme le sort : elle n'a de valeur, qu'autant que l'opinion du juge a plus de tendance à la vérité qu'à l'erreur. C'est ensuite par le rapport des nombres de voix favorables et contraires à l'accusé, que je détermine la probabilité de cette opinion. »

À partir de cette hypothèse, par ailleurs fort discutable, Laplace effectue quelques calculs<sup>1</sup> :

« Dans les tribunaux spéciaux où cinq voix sur huit, suffisent pour la condamnation de l'accusé, la probabilité de l'erreur à craindre sur la bonté du jugement est de  $65/256$  [...]. Dans un jury de douze membres ; si la pluralité exigée pour la condamnation est de huit voix sur douze ; la probabilité de l'erreur à craindre est de  $1\ 093/8\ 192$  [...] elle est à peu près de  $1/22$ , si cette pluralité est de neuf voix. Dans le cas de l'unanimité, la probabilité de l'erreur à craindre est de  $1/8192$ , c'est-à-dire mille fois moins que dans nos jurys. »

La conclusion de Laplace, dans son introduction<sup>2</sup>, est celle-ci :

« La probabilité des décisions est trop faible dans nos jurys, et je pense que pour donner une garantie suffisante à l'innocence, on doit exiger au moins, la pluralité de neuf voix sur douze. »

Laplace traite aussi dans ce premier supplément des meilleures estimations linéaires lorsque le problème de la minimisation des erreurs est multiparamétrique. Il écrit :

*« Les phénomènes de la nature sont le plus souvent enveloppés de tant de circonstances étrangères, un si grand nombre de causes perturbatrices y mêlent leur influence, qu'il est très difficile de les reconnaître. On ne peut y parvenir qu'en multipliant les observations ou les expériences, afin que les effets étrangers venant à se détruire réciproquement, les résultats moyens mettent en évidence ces phénomènes et leurs éléments divers. Plus les observations sont nombreuses et moins elles s'écartent entre elles, plus leurs résultats approchent de la vérité. On remplit cette dernière condition par le choix des méthodes, par la précision des instruments et par le soin que l'on met à bien observer. Ensuite on détermine par la théorie des probabilités les résultats moyens les plus avantageux ou ceux qui donnent le moins de prise à l'erreur. Mais cela ne suffit pas ; il est de plus nécessaire d'apprécier la probabilité que les erreurs de ces résultats sont comprises dans des limites données. Sans cela, on n'a qu'une connaissance imparfaite du degré d'exactitude obtenu. Des formules propres à cet objet sont donc un vrai perfectionnement de la méthode des sciences, et qu'il est bien important d'ajouter à cette méthode. L'analyse qu'elles exigent est la plus délicate et la plus difficile de la théorie des probabilités. C'est une des choses que j'ai eue principalement en vue dans mon Ouvrage, dans lequel je suis parvenu à*

1. 1<sup>er</sup> supplément, pages 30 et 31.

2. *Théorie Analytique* de 1820, page xcj.

des formules de ce genre, qui ont l'avantage remarquable d'être indépendantes de la loi de probabilité des erreurs et de ne renfermer que des quantités données par les observations mêmes et par leurs expressions. Je vais en rappeler ici les principes. Chaque observation a pour expression analytique une fonction des éléments que l'on veut déterminer ; et, si ces éléments sont à peu près connus, cette fonction devient une fonction linéaire de leurs corrections. En l'égalant à l'observation même, on forme ce que l'on nomme une *équation de condition*. Si l'on a un grand nombre d'équations semblables, on les combine, de manière à obtenir autant d'équations finales qu'il y a d'éléments dont on détermine ensuite les corrections, en résolvant ces équations. Mais quelle est la manière la plus avantageuse de combiner les équations de condition pour obtenir les équations finales ? Quelle est la loi des erreurs dont les éléments que l'on en tire sont encore susceptibles ? C'est ce que la théorie des probabilités fait connaître. La formation d'une équation finale, au moyen des équations de condition, revient à multiplier chacune de celles-ci par un facteur indéterminé et à réunir ces produits ; mais il faut choisir le système de facteurs qui donne la plus petite erreur à craindre. Or il est visible que, si l'on multiplie les erreurs possibles d'un élément par leurs probabilités respectives, le système le plus avantageux sera celui dans lequel la somme de ces produits, tous pris positivement est un minimum ; car une erreur positive ou négative doit être considérée comme une perte. En formant donc cette somme de produits, la condition du minimum déterminera le système de facteurs qu'il faut choisir. On trouve ainsi que ce système est celui des coefficients des éléments dans chaque équation de condition, en sorte que l'on forme une première équation finale en multipliant respectivement chaque équation de condition par son coefficient du premier élément et en réunissant toutes ces équations ainsi multipliées. On forme une seconde équation finale en employant de même les coefficients du second élément, et ainsi de suite. De cette manière les éléments et les lois des phénomènes, renfermés dans le recueil d'un grand nombre d'observations, se développent avec le plus d'évidence. J'ai donné, dans le n° 21 du livre II de ma *Théorie analytique des probabilités* l'expression de l'erreur moyenne à craindre sur chaque élément. Cette expression donne la probabilité des erreurs dont l'élément est encore susceptible, et qui est proportionnelle au nombre dont le logarithme hyperbolique est l'unité, élevé à une puissance égale au carré de l'erreur pris en moins et divisé par le carré du double de celle expression et par le rapport de la circonférence au diamètre. Le coefficient du carré négatif de l'erreur dans cet exposant peut donc être considéré comme le module de la probabilité des erreurs, puisque l'erreur restant la même, la probabilité décroît avec rapidité quand il augmente, en sorte que le résultat obtenu pèse, si je puis ainsi dire, vers la vérité, d'autant plus que ce module est plus grand. Je nommerai, par cette raison, ce module *poids* du résultat. Par une analogie remarquable de ces poids avec ceux des corps comparés à leur centre commun de gravité, il arrive que, si un même élément est donné par divers systèmes, composés chacun d'un grand nombre d'observations, le résultat moyen le plus avantageux de leur ensemble est la somme des produits de chaque résultat partiel par son poids, cette somme étant divisée par la somme de tous les poids. De plus, le poids total des divers systèmes est la somme de leurs poids partiels, en sorte que la probabilité du résultat moyen de leur ensemble est proportionnelle au nombre qui a l'unité pour logarithme hyperbolique, élevé à une puissance égale au carré de l'erreur, pris en moins et multiplié par la somme de tous les poids. Chaque poids dépend, à la vérité,

de la loi de probabilité des erreurs dans chaque système, et presque toujours cette loi est inconnue ; mais je suis heureusement parvenu à éliminer le facteur qui la renferme, au moyen de la somme des carrés des écarts des observations du système, de leur résultat moyen. Il serait donc à désirer, pour compléter nos connaissances sur les résultats obtenus par l'ensemble d'un grand nombre d'observations, qu'on écrivît à côté de chaque résultat le poids qui lui correspond. Pour faciliter le calcul de ce poids, je développe son expression analytique, lorsque l'on n'a pas plus de trois éléments à déterminer. Mais, cette expression devenant de plus en plus compliquée à mesure que le nombre des éléments augmente, je donne un moyen fort simple pour déterminer le poids d'un résultat, quel que soit le nombre des éléments. Quand on a ainsi obtenu l'exponentielle qui représente la loi de probabilité des erreurs, on aura la probabilité que l'erreur du résultat est comprise dans des limites données, en prenant, dans ces limites, l'intégrale du produit de cette exponentielle par la différentielle de l'erreur et en la multipliant par la racine carrée du poids du résultat divisé par la circonférence dont le diamètre est l'unité. De là il suit que, pour une même probabilité, les erreurs des résultats sont réciproques aux racines carrées de leurs poids, ce qui peut servir à comparer leur précision respective. »

Le but de Laplace, dans ce supplément, est d'établir une méthode d'estimation indépendante des lois de distribution des erreurs (qu'il considère quelconques mais possédant une variance finie) et non pas limitée à des distributions normales. Pour  $m$  variables, il élimine  $m - 1$  paramètres dans ses équations de condition et applique la théorie de l'estimation pour le paramètre restant.

Les calculs de Laplace comportent une erreur. Czuber<sup>1</sup> a donné en 1891, une analyse de ces calculs de Laplace et des observations faites par divers auteurs, dont Bienaymé<sup>2</sup>.

La troisième édition comporte quatre suppléments qui ont pour objet :

- l'application des probabilités à la philosophie naturelle, dont il a été question, ci-dessus,
- l'application du calcul des probabilités aux opérations géodésiques<sup>3</sup>,
- l'application à la détermination de la méridienne de France,
- l'application des fonctions génératrices à quelques problèmes de jeux.

Le deuxième supplément, daté de février 1818, comporte une analyse détaillée de la probabilité des résultats déduits, par des procédés quelconques d'un grand nombre d'observations. Il adopte, au départ, dans ce mémoire, la loi

---

1. Czuber E., *Théorie der Beobachtungsfehler*, Leipzig, 1891, p. 252.  
 2. Bienaymé I. J., *Mémoire sur la probabilité des erreurs d'après la méthode des moindres carrés*, Journal de Liouville 1, 17, 1852, p. 33-78. Cf. Hald A., *A history of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 410-418.  
 3. Cf. Hald A., *A history of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 418-443.

normale comme loi de distribution des erreurs de mesures elles-mêmes et non comme loi de leurs moyennes, puis il montre que ses calculs restent valables quelle que soit la loi de distribution des erreurs élémentaires et que l'on obtient, au total, une loi de distribution normale de l'erreur sur les quantités inconnues.

Le troisième supplément donne des applications numériques des formules du second supplément. Il se termine par une *Méthode générale du calcul des probabilités, lorsqu'il y a plusieurs sources d'erreurs*. Laplace traite ici le cas de plusieurs types d'erreurs qui ne doivent pas être supposées avoir la même loi de probabilité.

Le résultat le plus important obtenu par Laplace, au total, dans son ouvrage est appelé le théorème limite central (ou Théorème de la limite centrale)<sup>1</sup> où *central* signifie *fondamental*<sup>2</sup>. Voici comment Laplace commente ce résultat à la page XLIX de l'introduction de la 3<sup>e</sup> édition de sa *Théorie analytique des probabilités* :

« Les phénomènes qui semblent le plus dépendre du hasard, présentent donc en se multipliant, une tendance à se rapprocher sans cesse, de rapports fixes ; de manière que si l'on conçoit de part et d'autre de chacun de ces rapports, un intervalle aussi petit que l'on voudra, la probabilité que le résultat moyen des observations tombe dans cet intervalle, finira par ne différer de la certitude, que d'une quantité au-dessous de toute grandeur assignable. On peut ainsi par le calcul des probabilités, appliqué

1. Le terme *Central limit theorem* ou *Théorème de la limite centrale* ou *Théorème limite central* a été appelé *Fundamentalsätze der Wahrscheinlichkeitsrechnung* en 1919 dans un papier portant le même titre (Math.Z.4, p. 1-97) par R. von Mises. George Polya a introduit le terme en 1920 dans son papier *Ueber den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung* (Math.Z.15)
2. Lucien Le Cam a commenté comme suit, en 1986, l'histoire du Théorème de la limite centrale [*The central limit theorem around 1935*, Statistical science, volume 1, 1986, p.78-96.] (traduction page suivante) :  
 « Au commencement, il y eut De Moivre, Laplace, et de nombreux Bernouillis et ils engendrèrent des théorèmes de la limite, et les hommes avisés dirent que cela était bien et ils donnèrent au théorème le nom de Gauss. Puis vinrent de nouvelles générations qui trouvèrent que le théorème avait une validité expérimentale mais manquait de rigueur. Alors vinrent Chebychev, Liapunov et Markov et ils prouvèrent le théorème et Polya voyant l'usage des moments déclara que le nom *Théorème de la limite centrale* devait être utilisé. Puis vint Lindeberg et il dit que cela était élémentaire car Taylor avait développé ce qui demandait à être développé, et il le dit deux fois, mais Lévy avait vu que les séries de Fourier sont des fonctions caractéristiques et il déclara "laissez-les se multiplier et donner naissance à des théorèmes de la limite et des lois stables." Et cela était bon, stable et suffisant, mais on demanda "Est-ce nécessaire ?" Lévy répondit, "Je dois vous dire, à la vérité, que cela n'était pas nécessaire, mais le moment viendra où le nom de Gauss ne sera présent que dans l'image de Gauss lui-même et alors, cela deviendra nécessaire." Ce fut une prophétie et quand Cramer annonça que le temps était venu, il y eut de nombreuses réjouissances et Lévy déclara qu'il fallait l'écrire dans les bibles, et ainsi fut fait, et il y eut de nombreux théorèmes de la limite et beaucoup furent centraux et ils saturèrent les chroniques. Voilà l'histoire du théorème de la limite centrale. »



à un grand nombre d'observations, reconnaître l'existence de ces rapports. Mais avant que d'en rechercher les causes, il est nécessaire, pour ne point s'égarer dans de vaines spéculations, de s'assurer qu'ils sont indiqués avec une probabilité qui ne permet point de les regarder comme des anomalies dues au hasard. La théorie des fonctions génératrices donne une expression très simple de cette probabilité, que l'on obtient en intégrant le produit de la différentielle de la quantité dont le résultat déduit d'un grand nombre d'observations s'écarte de la vérité, par une constante moindre que l'unité, dépendante de la nature du problème, et élevée à une puissance dont l'exposant est le rapport du carré de cet écart au nombre des observations. L'intégrale prise entre des limites données, et divisée par la même intégrale étendue à l'infini positif et négatif, exprimera la probabilité que l'écart de la vérité est compris entre ces limites. »

Nous avons indiqué plus haut que la démonstration du théorème limite central donnée par Laplace en 1810, se trouve dans son mémoire intitulé *Sur les approximations des formules qui sont fonctions de très grands nombres, et sur leurs applications aux probabilités* et dans la suite de ce mémoire.

La démonstration du théorème limite central est donnée, ici, pour la première fois par Laplace<sup>1</sup>. Comme cas particulier de ce théorème, on retrouve également la convergence d'une suite de loi binomiale vers la loi normale (théorème de Bernoulli). Ce théorème justifie l'utilisation de la loi normale lorsqu'il y a répétition d'expériences identiques. On voit apparaître dans ce supplément (page 559) la loi normale (appelée aujourd'hui loi de Laplace-Gauss) pour une variable aléatoire de valeur moyenne  $\mu$  et d'écart type  $\sigma$  :

(Laplace prend le soin de préciser que  $\pi$  est la *demi circonférence dont le rayon est l'unité*, et  $e$  le nombre dont le *logarithme hyperbolique est l'unité*.)

La loi de paramètre  $\mu = 0$  peut être dérivée comme une limite de la loi binomiale de paramètres  $n$  et  $p$  quand  $n \rightarrow \infty$  :

$$P(a < \frac{x - np}{\sqrt{n}} < b) \rightarrow \int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}x^2} dx$$

où  $\sigma^2 = p(1 - p)$ .

La démonstration de Laplace n'est pas parfaite. L'approche directe des probabilités suivie par Bayes, Bernoulli, Montmort, de Moivre et par Laplace dans son traité, ne sera remplacée qu'en 1933, par l'approche axiomatique de Andrei Kolmogorov réalisée dans le cadre de la théorie de la mesure.

1. Bienaymé a écrit en 1853 (*Considérations à l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés*, C.r. Acad. Roy. Sci. Paris, t. 37, p. 311-312.) : « De 1770 à 1809, pendant près de quarante ans, Laplace avait donné des Mémoires nombreux sur les probabilités : mais, quelque intérêt qu'il y eût dans ces Mémoires, il n'avait pas voulu les rédiger en théorie générale. Aussitôt qu'il a reconnu la propriété des fonctions de probabilités [le théorème de la limite centrale], il voit clairement que c'est un principe qui régit presque toutes les applications, et il compose sa théorie. »

Notons que Stiegler<sup>1</sup> pense que c'est après avoir vu les travaux de Gauss publiés en 1809, que Laplace a compris la connexion entre le théorème limite central et la technique d'estimation linéaire et publié sa suite du mémoire de 1810. Nous reviendrons sur ce point dans le chapitre consacré à Gauss.

Oscar Sheynin<sup>2</sup> estime, pour sa part, que Laplace a réuni dans sa *Théorie analytique des probabilités*, ses mémoires séparés publiés plus de vingt ans plus tôt, mais que l'ensemble n'est pas homogène, Laplace ayant négligé de résoudre des problèmes similaires par des méthodes identiques. Sheynin trouve que la lecture du traité de Laplace est difficile et cite Bowditch, le traducteur en Anglais du *Traité de Mécanique céleste* qui déclare que lorsque Laplace écrit « il suit d'une façon évidente que... », il faut des heures ou des jours de travail pour découvrir l'évidence.

Sheynin ajoute que l'*Essai philosophique des probabilités* de Laplace eut peu de succès à cause de « la parution des contributions superficielles de Quetelet, écrites dans un bon style ». Cela nous semble peu probable, l'essai de Laplace ayant connu un tirage au moins cinq fois supérieur au petit livre de vulgarisation de Quetelet<sup>3</sup>.

Pour le reste, Sheynin donne une excellente analyse des travaux de Laplace.

Pour résumer le présent chapitre, on peut dire que le résultat le plus important obtenu par Laplace est consigné dans le court mémoire intitulé *Suite du Mémoire sur les approximations des formules qui sont fonctions des très grands nombres et leur application aux probabilités*, publié en 1810. Ce mémoire donne le théorème limite central énoncé plus haut.

On peut ajouter les remarques suivantes :

Avec des notations modernes, si  $(X_n)$  est une suite de variables aléatoires indépendantes ( $n \in \mathbb{N}^*$ , ensemble des entiers naturels privé de l'élément zéro) et si ces variables aléatoires sont de même loi, d'espérance  $\mu$  et d'écart-type  $\sigma$  finis, en posant

$$\forall n \in \mathbb{N}^*, \overline{X_n} = \frac{X_1 + X_2 + \dots + X_n}{n},$$

la quantité  $\overline{X_n} = \frac{X_1 + X_2 + \dots + X_n}{n}$  est appelée moyenne empirique des mesures de la variable aléatoire, moyenne qui tend, lorsque le nombre de mesures  $n$  augmente, vers l'espérance  $\mu$ .

1. Stiegler, *The history of statistics*, Belknap Press, 1986, p. 143.

2. *Theory of Probability. A historical Essay*, Berlin, 2005, ISBN 3-938417-15-3, NG Verlag, Design Viatcheslav Demidov.

3. A. Quetelet, *Théorie des Probabilités*, Bruxelles, A. Jamar, (1853).

Lorsque  $n$  est grand, la moyenne empirique est comprise entre  $\mu - \frac{3\sigma}{\sqrt{n}}$  et  $\mu + \frac{3\sigma}{\sqrt{n}}$ .

Gauss<sup>1</sup>, en 1809, avait démontré la loi normale, en considérant comme un axiome l'hypothèse que si une quantité a été obtenue par plusieurs observations, faites avec le même soin dans des circonstances semblables, la moyenne arithmétique des valeurs observées sera la valeur la plus probable de cette quantité. Laplace, dans la suite de son mémoire de 1810, démontre que même en ignorant la distribution des erreurs sur chaque groupe de mesures, on peut, lorsque le nombre d'observations est très grand, remplacer les mesures par leur moyenne empirique et obtenir pour l'ensemble des observations une loi normale. Comme dans le mémoire de Gauss, Laplace obtient le meilleur résultat en minimisant la somme des carrés des erreurs de chaque résultat, multipliées respectivement par la plus grande ordonnée de la courbe de probabilité des erreurs, méthode appelée *des moindres carrés*. Il faut, de plus, souligner que les résultats de Laplace sont obtenus, comme ceux de Gauss, à l'aide d'un raisonnement probabiliste.

Comme on le verra au chapitre consacré à Gauss, l'établissement, en 1809 et 1810, de la *loi normale*, dont l'usage se révélera fondamental en statistique, s'est fait par deux voies : celle de Gauss relative à la théorie des erreurs aboutissant à la méthode des moindres carrés, et celle de Laplace qui a énoncé la première version du théorème limite central.

La synthèse entre ces deux approches impose donc, dès 1810, la loi normale comme loi quasi universelle. En effet, même si la distribution individuelle des erreurs ne suit pas une loi normale, celle des moyennes des erreurs suit sous certaines conditions (indépendance, lois identiques), une loi normale.

Comme nous l'avons indiqué par ailleurs, la fonction erreur est une intégrale numérique qui a été introduite par Gauss en 1816. On la trouve tabulée dans tous les manuels et elle est définie par

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\zeta^2} d\zeta.$$

La probabilité pour qu'une variable normale centrée réduite<sup>2</sup> prenne une valeur dans l'intervalle  $[-z, z]$  est  $\operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)$ .

1. *Theoria motus corporum coelestium*, Hamburg, Perthes et Besser, 1809.

2. Une variable aléatoire centrée réduite a une espérance mathématique  $E(X)$  nulle et un écart type égal à 1. Par rapport à une variable aléatoire quelconque  $X$ , la variable centrée réduite est une variable translatée et homothétique  $\frac{X-E(X)}{\sigma(X)}$ .

Si  $F(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\zeta^2/2} d\zeta$  est la fonction de répartition de la loi normale centrée réduite, on a

$$F(x) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{x}{\sqrt{2}} \right).$$

## 12.7. Biographie de Laplace (1749-1827)

Pierre Simon de Laplace naît d'une famille modeste en Normandie, à Beaumont-en-Auge (Calvados), le 23 mars 1749. Il est élève du collège de Beaumont et se révèle doué pour les études. Il y acquiert en particulier des bases d'algèbre et de géométrie conventionnelles, mais solides. Sa mère meurt peu après son huitième anniversaire, ce que curieusement il n'évoque jamais dans ses écrits. À l'âge de seize ans il entre à l'université de Caen. Le 1<sup>er</sup> juillet 1769, Laplace qui avait d'abord pensé se destiner aux ordres, passe sa maîtrise ès arts. Mais le goût des mathématiques est plus fort, il veut tenter sa chance à Paris. Il arrive dans la capitale, âgé de vingt ans, avec sur lui une lettre de recommandations de ses professeurs et un essai de mathématique. Il essaie de rencontrer d'Alembert mais ce dernier, harcelé par les solliciteurs, ne le reçoit pas. Laplace ne désarme pas, il transmet un essai de quatre pages sur les principes généraux de mécanique, en particulier touchant au problème de l'inertie, à l'auteur du *Traité de dynamique*. On prétend que d'Alembert, impressionné par la profondeur de vue du jeune provincial, le reçoit en lui disant : « Monsieur, vous avez remarqué que je n'ai pas fait cas de vos recommandations : vous n'en avez pas besoin, vous vous introduisez bien mieux tout seul. » Ces mots, venant de l'un des savants les plus influents d'Europe sont plus qu'un encouragement. Laplace se porte candidat à un poste de professeur de mathématiques à l'École militaire<sup>1</sup>. Après un concours où il se classe deuxième, il est recruté et reçoit un salaire confortable. Le jeune professeur, que d'Alembert appelle « l'abbé Laplace », doit porter la robe et le col ecclésiastiques. Il passe beaucoup de temps dans la riche bibliothèque de l'École militaire et dévore les ouvrages d'Euler et de Lagrange. Laplace rêve d'entrer à l'Académie des sciences pour rejoindre l'élite de la science. Dès le 28 mars 1770, il tente une première fois sa chance et soumet une communication. Ses communications successives (treize entre 1770 et 1773), de grande valeur<sup>2</sup>, le font connaître et apprécier des membres de l'institution. Finalement Laplace est élu à l'Académie des sciences comme « adjoint mécanicien » le 4 avril 1773. Il vient tout juste d'avoir vingt-quatre ans et emporte l'élection devant, entre autres, Monge

1. L'École militaire a été fondée en 1753. En 1769 elle triple ses effectifs et recrute de nouveaux professeurs.
2. Condorcet écrira plus tard : « Cette compagnie qui s'est empressée de récompenser ses travaux et ses talents, n'avait encore vu personne aussi jeune, lui présenter en si peu de temps, tant de mémoires importants, et sur des matières si diverses et si importantes ».

et Legendre ! Pour consacrer tout son temps aux mathématiques, il souhaite obtenir un poste à Saint-Petersbourg (où se trouve Euler) ou à Berlin (où se trouve Lagrange). Malgré le soutien de d'Alembert il n'y parvient pas et reste à l'École militaire. Toutefois en 1776, il est dégagé de ses obligations d'enseignement et peut se consacrer entièrement à la recherche. Laplace axe alors ses travaux sur ce qui deviendra, suivant sa propre appellation, la mécanique céleste. Mais en 1777, il interrompt son programme pour rejoindre celui de son collègue chimiste Lavoisier. Il écrit à Lagrange en 1783 qu'il est incapable de « refuser aux instances de [son] confrère M. de Lavoisier, qui met dans ce travail commun toute la complaisance et toute la sagacité que [je puis] désirer ». Ce travail commun allait conduire au *Mémoire sur la chaleur*. Cependant, Laplace écrit à un ami genevois, M. Deluc, « je ne sais comment je me suis laissé entraîner par Monsieur de Lavoisier à travailler avec lui sur la chaleur. [...] Un penchant irrésistible pour les mathématiques, et une paresse qui m'est naturelle, m'empêcheront toujours de contribuer aux progrès de la physique. » Mais en fait son intérêt pour les sciences physiques ne s'est pas estompé, comme en témoignent ses rencontres hebdomadaires dans sa maison d'Arcueil. La collaboration de Laplace et de Lavoisier continuera par intermittence jusqu'à l'exécution du chimiste en 1794.

En 1785, Laplace est promu au rang de membre de l'Académie des sciences<sup>1</sup>. La même année, en septembre, il examine la candidature d'un cadet à l'École d'artillerie de Brieenne : le jeune qui a seize ans s'appelle Buonaparte. Laplace le classe 42<sup>e</sup> sur 58, ce qui permet à l'élève officier une affectation dans le corps de l'artillerie. Ensuite, le jeune officier ira à l'École d'artillerie d'Auxonne et suivra des cours de mathématiques. Ainsi débuta la prodigieuse carrière de Napoléon Bonaparte (1769-1821), qui par ailleurs joua un rôle important dans la vie de Laplace.

En 1788, notre mathématicien a trente-neuf ans et est un célibataire endurci. Mais en mars de cette même année il épouse une jeune fille de dix-huit ans, Marie Anne Charlotte Courty de Romange, issue d'une bonne famille de Besançon. Berthollet, l'ami de la Société d'Arcueil, est le premier témoin de Laplace. Le couple aura deux enfants, Charles Émile Pierre Joseph, né le 5 avril 1789 et Sophie Suzanne, née le 18 avril 1792. Comme Lagrange, Laplace traverse sans heurt la période de la révolution et de la terreur. Il a la prudence de n'afficher aucune position politique. Par ailleurs, il prend part aux commissions scientifiques, comme celle devant rendre toutes les unités à base décimale. Après la Terreur, Laplace est, avec Monge et Fourcroy, l'un des rares à participer à la reconstruction de la vie scientifique. Comme son collègue Lagrange il est appelé auprès des deux nouvelles institutions qui viennent d'être créées : l'École Normale de l'an III et l'École Polytechnique.

1. Laplace est depuis 1783 « associé mécanicien » et devient le 23 avril 1785 « pensionnaire de la Classe de mécanique ».

À l'École Normale il enseigne les mathématiques, aux côtés de Monge et de Lagrange. Il prononce sa leçon inaugurale le 20 janvier 1795, dans l'amphithéâtre Verniquet du Jardin des Plantes, devant une foule de sept cents étudiants adultes. Il est par ailleurs examinateur à la nouvelle École Polytechnique<sup>1</sup> et retrouve là une charge semblable à celle qu'il a perdue en 1792 à l'École d'artillerie. Laplace prend également part à la création du Bureau des Longitudes. Il en est le concepteur officieux et, suite au décret de la Convention nationale du 26 juin 1795, en devient l'un des membres fondateurs. Le Bureau supervise les deux observatoires parisiens et correspond avec diverses institutions en province et à l'étranger. Quelques mois plus tard, le gouvernement fonde l'Institut national des sciences et des arts (le décret fondant l'Institut date du 26 octobre 1795) et Laplace participe à son organisation. À ce titre, c'est lui qui présente le premier rapport annuel de l'Institut aux députés. L'Institut fait renaître l'Académie des sciences (qui avait été fermée en août 1793) où Laplace est nommé à la section de mathématiques, ce qui lui confère le statut d'électeur.

En même temps qu'il prépare ses cours de l'École Normale, Laplace rédige un ouvrage destiné au grand public qui aura un grand retentissement : *l'Exposition du système du monde* (1796).

Le 25 décembre 1797 (5 nivôse an VI), le général Bonaparte est élu à l'Institut, dans la section de mécanique, à la place de Lazare Carnot, en exil suite au coup d'état de fructidor. Laplace accompagne alors Berthollet dans l'escorte qui installe Bonaparte dans son siège. On peut se demander la raison de cette élection. En fait elle récompensait l'intérêt qu'avait manifesté le jeune général pour les sciences et les arts lors de la campagne d'Italie. C'est ainsi qu'en 1796, dans une lettre ouverte à Oriani, astronome piémontais et ami de Laplace, Bonaparte écrivait : « Les sciences, qui honorent l'esprit humain, les arts qui embellissent la vie et transmettent les grandes actions à la postérité, doivent être spécialement honorés dans les gouvernements libres ». Un tel programme ne pouvait laisser indifférents les membres de l'Académie. Bonaparte assiste donc aux séances de l'Académie des sciences au côté de Monge, Berthollet, Lagrange, Laplace<sup>2</sup>, Legendre, Delambre, Lalande, Méchain, Haüy, Lamarck, de Jussieu, Daubenton, Cuvier... Lorsque le Directoire décide, le 12 avril 1798, la conquête militaire de l'Égypte et en confie le commandement au jeune général (Bonaparte n'a alors que 26 ans), ce dernier décide d'emmener avec lui des savants et dit à Monge : « Nous aurons avec nous un tiers de l'Institut et des instruments de toutes espèces ». Même si l'on est loin du « tiers de l'Institut », Bonaparte embarque

- 
1. Laplace deviendra également membre du Conseil d'Administration de cette École.
  2. À l'Institut, Laplace et Bonaparte servirent dans la même commission avec Lacroix afin de rédiger un rapport sur un mémoire de Biot.

le 19 mai 1798 avec 167 savants<sup>1</sup>. Il reviendra en France le 9 octobre 1799. Sa campagne d'Égypte est un échec militaire, mais une réussite scientifique marquée par la publication de la fameuse *Description de l'Égypte* dont l'introduction (un tome entier) est rédigée par Joseph Fourier. On connaît la suite, le 9 novembre 1799 (18 brumaire), Bonaparte réalise son coup d'état et devient deux jours plus tard Premier Consul. Le lendemain, à sa grande surprise, Laplace est nommé ministre de l'Intérieur. Il ne restera à ce poste que six semaines. Il considère son bureau comme le « ministère des Sciences et des Arts », en quelque sorte une extension exécutive de l'Institut. Il est rapidement remplacé par Lucien Bonaparte, le frère du Premier Consul, qui avait aidé ce dernier à préparer son coup d'État<sup>2</sup>. Désireux de garder Laplace à son service, Napoléon I<sup>er</sup> nomme, le 24 décembre 1799, au Sénat, poste qu'il juge moins sensible. Laplace en devient successivement le secrétaire, le président, le vice-président puis, en 1803 le chancelier. Le couple Laplace fréquente alors les proches de Bonaparte et la nouvelle élite gouvernante. La sœur du Premier Consul, Élisa, élevée au rang de princesse, s'attache même les services de Madame Laplace à sa cour de Lucca. Quant à Pierre Simon, il est fait, en 1806, Comte d'Empire. De plus Napoléon I<sup>er</sup> nomme, en 1813, dans l'ordre de La Réunion. Le dernier grand traité de Laplace, la *Théorie analytique des Probabilités*, dont la rédaction a débuté en 1795, paraît en 1812 et comporte une dédicace à Napoléon.

En 1813, Laplace accueille, comme beaucoup d'autres français, la nouvelle de la débâcle de la campagne de Russie avec une profonde tristesse. Sa fille bien-aimée meurt en septembre et il craint pour la vie de son fils Émile, posté sur le front oriental avec l'Empereur. Napoléon met la nation en danger, et la loyauté de Laplace envers lui commence à flancher. D'autant qu'à son retour de la déroute de Leipzig, on prétend que l'Empereur aborde notre mathématicien et lui dit : « Ah ! vous avez bien maigri. – Sire, j'ai perdu ma fille. – Oh ! il n'y a pas de quoi maigrir. Vous êtes géomètre ; soumettez cet événement au calcul, et vous verrez que tout cela égal à zéro »<sup>3</sup>.

Lorsque le Sénat vote la destitution de Napoléon, le 2 avril 1814, Laplace, par prudence, est à Lisieux chez son gendre (il trouva préférable cette « retraite » pendant l'épisode des Cent Jours). À son retour, le 12 avril, il se rend au

1. Bonaparte dira plus tard : « Le temps que j'ai passé en Égypte a été le plus beau de ma vie car il a été le plus idéal ».
2. Les critiques du Premier Consul vis-à-vis de son ministre de l'Intérieur ne viendront que bien plus tard. « Laplace, dit Napoléon, ne saisissait aucune question sous son véritable point de vue ; il cherchait des subtilités partout, n'avait que des idées problématiques, et portait enfin l'esprit des infiniment petits jusque dans l'administration » (Correspondance de Napoléon I<sup>er</sup>, 30, 1870, p. 330). Mais il convient de nuancer ce texte, écrit à Sainte-Hélène, après la déchéance de l'Empereur par le Sénat en général et par Laplace en particulier !
3. Ce dialogue est rapporté par Claude Chaptal dans *Mes souvenirs sur Napoléon*, Paris, 1813, p. 342.

Sénat et signe le registre du 2 avril, ratifiant ainsi le vote de destitution. En sa qualité de Chancelier du Sénat, Laplace est choisi pour diriger le comité d'accueil de Louis XVIII. Le nouveau monarque nomme Laplace (et plusieurs autres sénateurs) à la nouvelle Chambre des pairs. Le fils de Laplace, Émile, qui avait été l'aide de camp de l'Empereur, devient le secrétaire général de la compagnie des dragons de la famille d'Orléans. Pierre Simon est élu le 11 avril 1816 à l'Académie française. En 1817, Louis XVIII le fait marquis et Pair de France. Jusqu'à ses soixante-dix ans (1819), malgré une santé chancelante, il poursuit ses travaux et ses publications. En avril 1823, on fête le cinquantième anniversaire de son élection à l'Académie des sciences. C'est vers cette époque que Laplace entreprend des recherches sur la vie de Newton. Peut-être souhaite-t-il comparer sa carrière à celle de son illustre prédécesseur anglais, mais surtout il s'interroge sur le rôle que Newton fait jouer à Dieu dans l'Univers. Laplace, grand admirateur et, en quelque sorte, successeur de Newton (il sera surnommé le « Newton français »), ne partage pas ses convictions religieuses. Il est connu que lorsque l'Empereur Napoléon Bonaparte lui reprocha de ne pas citer Dieu dans son traité de mécanique céleste, il répondit : « Sire, je n'ai pas eu besoin de cette hypothèse ». Par un curieux hasard, Laplace meurt le 9 mars 1827, cent ans presque jour pour jour après Newton. Fourier rapporte, dans son éloge funèbre de Laplace, que sur son lit de mort ce dernier aurait dit : « Ce que nous connaissons est peu de chose, ce que nous ignorons est immense »<sup>1</sup>.

## 12.8. Les publications de Laplace

Les mémoires de Laplace, au nombre de cent trente environ, ont été publiés dans dix-sept revues ou journaux différents. À cela il faut ajouter sept éditions originales de livres et des rééditions ou remises en vente multiples de ces livres. De nombreux papiers de Laplace ont été insérés, simultanément, dans deux revues à la fois. C'est particulièrement le cas pour les documents édités dans l'important journal *Annales de Chimie et de Physique*, à partir de 1816, que l'on retrouve souvent dans *Connaissance des Temps*. Enfin il faut noter que nombre d'articles ont été recyclés en suppléments à la *Mécanique céleste* et à la *Théorie des probabilités*. Voici la liste, les dates et quelques précisions concernant les livres de Laplace :

- *Théorie du mouvement et de la figure elliptique des planètes*, Paris, 1784
- *Théorie des attractions des sphéroïdes et de la figure des planètes*, Paris, 1785. Ce second ouvrage reprend les calculs du précédent, les simplifie et les généralise. On retrouvera le tout, bien sûr, un peu plus tard dans la *Mécanique céleste*.

1. Dans un récit plus détaillé, Maurice, auteur de *Maladie et derniers moments de Mr de la Place*, indique qu'il dit plutôt : « Ah ! Nous courons après des chimères ».



- *Exposition du système du monde*, 2 volumes, Paris, an IV (1796). Ce livre décrit l'origine et la formation du système solaire sans faire appel aux mathématiques.
- *Traité de mécanique céleste*, 5 volumes et divers suppléments. Paris, vol. I et II : an VII (1799) ; vol. III : an XI (1802) ; vol. IV, 1805 ; vol. V : 1823-1825. Ce grand traité, publié sur 26 ans est resté le grand classique de la mécanique newtonienne jusqu'aux travaux de Poincaré.
- *Théorie analytique des probabilités*, Paris, 1812.
- *Essai philosophique sur les probabilités*, Paris 1814. Cet essai a servi ensuite d'introduction au *Traité analytique des probabilités*.
- *Précis de l'histoire de l'astronomie*, Paris, 1821. Il s'agit là de remise en vente, sous forme séparée, du livre V de la cinquième édition de l'*Exposition du système du monde*.

## 12.9. Références

- C.C. Gillispie, *Pierre-Simon Laplace. 1749-1827. A life in exact science*, Princeton, NJ, 1997.
- J. Grattan-Guinness, *Dictionary of Scientific Biography*, Charles Scribner's sons, New York, vol. 15, p. 273-403, 1970.
- E.T. Bell, *Men of mathematics*, Simon et Schuster, New York, 1986.
- M.P. Corsland, *The Society of Arcueil: A View of French Science at the Time of Napoleon I*, 1967.

# Condorcet

## 13.1. Les travaux publiés dans les *Mémoires de l'Académie royale des sciences de Paris*, entre 1781 et 1784

Les premières publications de Condorcet, relatives aux probabilités, ont paru dans les *Mémoires de l'Académie royale des sciences de Paris*, entre 1781 et 1784.

- Partie 1. Réflexions sur la règle générale qui prescrit de prendre pour valeur d'un événement incertain la probabilité de cet événement multipliée par la valeur de cet événement en lui-même (1781)<sup>1</sup> ;
- Partie 2. Application de l'analyse à cette question : déterminer la probabilité qu'un arrangement régulier est l'effet d'une intention de le produire (1781)<sup>2</sup> ;
- Partie 3. Sur l'évaluation des droits éventuels (1782)<sup>3</sup> ;
- Partie 4. Réflexions sur la méthode de déterminer la probabilité des événements futurs d'après l'observation des événements passés (1783)<sup>4</sup> ;
- Partie 5. Sur la probabilité des faits extraordinaires (1783)<sup>5</sup> ;
- Partie 6. Application de l'article précédent à quelques questions de critique (1784)<sup>6</sup>.

Le premier mémoire justifie comme suit la règle générale qui prescrit de prendre pour valeur d'un événement incertain la probabilité de cet événement multipliée par la valeur de cet événement en lui-même : Si deux joueurs  $A$  et  $B$  jouent ensemble et si la probabilité que  $A$  gagne une partie est  $p$  et celle de  $B$  est  $q$ , alors selon la règle générale, si la mise de  $A$  est  $kp$ , la mise de  $B$  doit être  $kq$ . En effet souligne Condorcet, selon le théorème de Jacques

1. *Histoire de l'Académie royale des sciences*, 1781, p. 707-728.

2. Publié avec la première partie.

3. *Histoire de l'Académie royale des sciences*, 1782, p. 674-691.

4. *Histoire de l'Académie royale des sciences*, 1783, p. 539-553.

5. *Id.*, p. 553-559.

6. *Histoire de l'Académie royale des sciences*, 1784, p. 454-468.

**M É M O I R E**  
**SUR LE CALCUL DES PROBABILITÉS.**  
**QUATRIÈME PARTIE.**

*Réflexions sur la méthode de déterminer la Probabilité  
 des évènements futurs, d'après l'Observation  
 des évènements passés.*

Par M. le Marquis DE CONDORCET.

CETTE partie de l'Analyse qui enseigne à déterminer la probabilité des évènements futurs, d'après l'ordre qu'ont suivi les évènements passés du même genre que l'on a observés, est susceptible d'un grand nombre d'applications utiles & curieuses : j'ai cru en conséquence qu'il pourroit n'être pas inutile d'examiner les principes sur lesquels cette Analyse est fondée; tel est l'objet des Réflexions suivantes.

I.

SOIENT deux évènements  $A$  &  $N$  que je suppose simplement contradictoires, c'est-à-dire, ne pouvant subsister ensemble & ayant nécessairement lieu l'un ou l'autre. Si  $A$  a eu lieu  $m$  fois &  $N$   $n$  fois, & qu'on demande la probabilité d'avoir en  $p + q$  évènements,  $p$  évènements  $A$  &  $q$  évènements  $N$ , la probabilité demandée sera

$$\frac{p + q \dots \dots q + 1}{1.2 \dots \dots q} \cdot \frac{\int x^{m+p} (1-x)^{n+q} dx}{\int x^m (1-x)^n dx},$$

cette intégrale étant prise depuis  $x = 1$  jusqu'à  $x = 0$ ; telle est la règle générale.

I I.

ON voit d'abord que cette loi n'exprime réellement la probabilité que dans les deux hypothèses suivantes,

Y y y ij

Bernoulli, si les joueurs jouent un très grand nombre de parties, le rapport des gains des joueurs sera très proche du rapport de  $p$  à  $q$ . En conséquence, la règle générale est correcte vis-à-vis des deux joueurs alors que toute autre règle avantagerait l'un d'eux.

Abordant ensuite de problème dit de Saint-Pétersbourg, Condorcet souligne (avec raison) que si les deux joueurs acceptent de jouer un très grand nombre de parties, seule la règle générale est équitable.

Le second mémoire est basé sur des hypothèses arbitraires et le troisième mémoire est peu clair.

Les premières phrases du quatrième papier indiquent son but et énoncent le théorème de Bayes généralisé :

« Cette partie de l'Analyse qui enseigne à déterminer la probabilité des événements futurs, d'après l'ordre qu'ont suivi les événements passés du même genre que l'on a observés, est susceptible d'un grand nombre d'applications utiles et curieuses ; j'ai cru en conséquence qu'il pourrait n'être pas inutile d'examiner les principes sur lesquels cette Analyse est fondée ; tel est l'objet des Réflexions suivantes. Soient deux événements  $A$  et  $N$  que je suppose simplement contradictoires, c'est-à-dire, ne pouvant subsister ensemble et ayant nécessairement lieu l'un ou l'autre. Si  $A$  a eu lieu  $m$  fois, et  $N$   $n$  fois, et qu'on demande la probabilité d'avoir en  $p + q$  événements,  $p$  événements  $A$  et  $q$  événements  $N$ , la probabilité sera

$$\frac{(p+q) \dots (q+1)}{1.2 \dots q} \cdot \frac{\int_0^1 x^{m+p}(1-x)^{n+q} dx}{\int_0^1 x^m(1-x)^n dx}$$

telle est la règle générale.

On voit d'abord que cette loi n'exprime réellement la probabilité que dans les deux hypothèses suivantes :

1° Si la probabilité des événements  $A$  et  $N$  reste la même dans toute la suite des événements ; cela est évident par la formule même qui exprime la loi.

2° Dans le cas où cette même probabilité est variable, mais où l'on supposerait en même temps que la valeur de la probabilité, quoique pouvant être différente pour chaque événement, est cependant prise au hasard pour chacun, d'après une certaine probabilité générale  $x$  pour  $A$ , et  $1 - x$  pour  $N$ . »

Condorcet souligne alors que Bayes suppose une loi de probabilité *a priori* constante bien que l'on puisse rencontrer des lois qui ne le sont pas. Appelant  $x$  la probabilité en faveur de  $A$ , il écrit :

« Nous considérerons ici l'hypothèse où l'on suppose  $x$  variable dans les deux cas, 1.° d'une suite d'événements qui ne sont liés entr'eux par aucune loi relative au temps et à l'ordre de leur production ; 2.° d'une suite d'événements liés entr'eux par une loi relative à cet ordre. Dans le premier cas, si on connoissait la loi de la probabilité de ces événements, la formule qui l'exprimerait ne serait pas une fonction du temps ou de la place qu'occupe l'événement ; elle le serait dans le second.

Si, par exemple, je suppose des paquets de cartes rouges et noires, dont le nombre soit  $N = m + n + p + q$ , que j'aie tiré de  $(m + n)$  de ces paquets  $m$  cartes rouges, et  $n$  noires, et que je cherche la probabilité de tirer des  $(p + q)$  paquets restants,  $p$  cartes rouges et  $q$  noires. Si je ne sais pas que ces paquets ont été formés en tirant au hasard des cartes d'un tas donné de cartes rouges et noires, il est clair que je n'ai aucune raison de supposer ni que le rapport du nombre des cartes rouges à celui des cartes noires, soit constant pour tous les paquets, ni que la valeur moyenne de ce rapport supposé variable, soit la même pour tous ; et en même temps je n'ai aucune raison de supposer que ce rapport varie suivant l'ordre dans lequel on range les paquets, ou suivant celui où l'on tire une carte de chacun.

Dans une suite au contraire d'événements naturels, il se présente un grand nombre de cas où l'on peut supposer que l'ordre du temps influe sur la production des événements. »

La loi variable conduit, selon Condorcet, à la probabilité suivante :

$$\frac{(p+q)! \iiint \dots \left[ \sum \frac{x_i}{N} \right]^{m+p} \left[ 1 - \sum \frac{x_i}{N} \right]^{n+q} dx_1 dx_2 \dots dx_N}{p!q! \iiint \dots \left[ \sum \frac{x_i}{N} \right]^m \left[ 1 - \sum \frac{x_i}{N} \right]^n dx_1 dx_2 \dots dx_N}$$

où  $x_i$  est la probabilité *a priori* assignée au  $i$ ème paquet de cartes et les intégrales étant prises entre 0 et 1.

Cette formule est, à l'évidence, trop complexe pour être utilisable.

Le cinquième mémoire *Sur la probabilité des faits extraordinaires* suppose que la probabilité d'un événement est  $p$  et que  $t$  est la probabilité qu'un certain témoignage est vrai. On demande, en supposant que le témoin déclare que l'événement s'est produit quelle est la probabilité que l'événement a eu lieu et celle qu'il n'a pas eu lieu. Condorcet donne avec peu d'explications les résultats respectifs suivants :

$$\frac{pt}{pt + (1-p)(1-t)} \text{ et } \frac{(1-p)(1-t)}{pt + (1-p)(1-t)}.$$

Ces résultats sont ensuite commentés par Condorcet dans son sixième mémoire intitulé *Application de l'article précédent à quelques questions de critique*. Il écrit :

« J'ai exposé dans l'article précédent, une méthode d'exprimer la probabilité des faits extraordinaires, en ayant égard à celle des témoignages qui les attestent, et à la probabilité propre de ces faits [...].

J'ai observé en même temps qu'il ne fallait pas dans ce cas entendre, par la probabilité propre d'un fait, le rapport du nombre des combinaisons où il a lieu, avec le nombre total des combinaisons. Par exemple, si d'un jeu de dix cartes on en a tiré une, et qu'un témoin me dise que c'est telle carte en particulier, la

probabilité propre de ce fait, qu'il s'agit de comparer avec la probabilité qui naît du témoignage, n'est pas la probabilité de tirer cette carte, qui serait  $1/10$ , mais la probabilité d'amener cette carte plutôt que telle autre carte déterminée en particulier ; et comme toutes ces probabilités sont égales, la probabilité propre est ici  $1/2$ .

Cette distinction était nécessaire et elle suffit pour expliquer la contrariété d'opinions entre deux classes de philosophes. Les uns ne peuvent se persuader que les mêmes témoignages puissent produire, pour un fait extraordinaire, une probabilité égale à celle qu'ils produisent pour un fait ordinaire ; et que, par exemple, si je crois un homme de bon sens qui me dit qu'une femme est accouchée d'un garçon, je dusse le croire également s'il me disait qu'elle est accouchée de douze.

Les autres au contraire sont convaincus que les témoignages conservent toute leur force, pour les faits extraordinaires et très peu probables, et ils font frapper de cette observation, que si on tire une loterie de 100 000 billets, et qu'un homme, digne de foi, dise que le numéro 256, par exemple, a eu le premier lot, personne ne doutera de son témoignage, quoiqu'il y ait 99 999 à parier contre 1 que cet événement n'est pas arrivé.

Or, au moyen de l'observation précédente, on voit que dans le second cas la probabilité propre du fait étant  $1/2$ , le témoignage conserve toute sa force, au lieu que dans le premier, cette probabilité étant très petite, réduit presque à rien celle du témoignage.

J'ai proposé ensuite de prendre, pour la probabilité propre du fait, le rapport du nombre de combinaisons qui donnent ce fait, ou un fait semblable au nombre total des combinaisons.

Ainsi, par exemple, dans le cas où on tire une carte d'un jeu de dix cartes, le nombre des combinaisons où l'on tire une carte déterminée quelconque est un ; celui des combinaisons où l'on tire une autre carte déterminée est aussi un ; donc  $1/2$  exprimera la probabilité propre.

Si on me dit qu'on a tiré deux fois de suite la même carte, alors on trouvera qu'il n'y a que dix combinaisons qui donnent deux fois une même carte, et quatre-vingt-dix qui donnent deux cartes différentes : la probabilité propre du fait n'est donc que  $1/10$ , et celle du témoignage commence à devenir plus faible.

Mais je crois devoir abandonner cette manière de considérer la question, 1.° parce qu'elle me paraît trop hypothétique ; 2.° parce que souvent cette comparaison d'événements semblables serait difficile à faire, ou, ce qui est encore pis, ne se ferait que d'après des suppositions arbitraires ; 3.° parce qu'en l'appliquant à des exemples, elle conduit à des résultats trop éloignés de ceux que donnerait la raison commune.

J'en ai donc cherché une autre, et il m'a paru plus exact de prendre, pour probabilité propre d'un événement, le rapport de la probabilité de cet événement prise dans le sens ordinaire, avec la probabilité moyenne de tous les autres événements. »

Ces mémoires de Condorcet ont été analysés et commentés en détail par Todhunter<sup>1</sup> qui jugeait, cependant, que lesdits écrits étaient « obscurs, pleins

1. Todhunter I., *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace*, Macmillan, London, 1865, chapitre XVIII.

de contradictions » et très difficiles à lire, d'autant que Condorcet utilisait des notations mathématiques qui lui étaient propres.

### **13.2. *L'Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*<sup>1</sup>**

Ce livre comporte un discours préliminaire suivi de l'essai proprement dit. On trouve, dans les premières pages, l'affirmation par Condorcet du fait qu'il considère être le premier à traiter le problème mentionné dans le titre. Puis il donne sa définition de la probabilité et indique que la mesure mathématique de la probabilité est une mesure précise de notre degré d'appréciation d'une situation. Il énonce :

« [...] Si on se borne à entendre par probabilité d'un événement le nombre de combinaisons où il a lieu, divisé par le nombre total des combinaisons possibles, le principe n'est qu'une vérité de définition, et qu'ainsi le calcul, dont il est la base, devient une vérité rigoureuse.

Mais on ne se borne point à ce seul sens.

On entend de plus, 1.° que si on connaît le nombre des combinaisons qui amènent un événement, et le nombre des combinaisons qui ne l'amènent pas, et que le premier surpasse le second, il y a lieu de croire que l'événement arrivera plutôt que de croire qu'il n'arrivera pas.

2.° Que ce motif de croire augmente en même temps que le rapport du nombre des combinaisons favorables avec le nombre total.

3.° Qu'il croît proportionnellement à ce même rapport.

La vérité de cette dernière proposition dépend de celle de la seconde et de la première. En effet, si le motif de croire devient plus fort lorsque le nombre des combinaisons augmente, on peut démontrer que si on répète un certain nombre de fois le jugement conforme à cette opinion, c'est-à-dire, que l'événement qui a plus de combinaisons en sa faveur arrivera plutôt que l'autre, la combinaison la plus probable est celle où le nombre des jugements vrais serait au nombre total des jugements, comme le nombre des combinaisons favorables à l'événement à leur nombre total, c'est-à-dire, que la combinaison la plus probable et celle où le rapport du nombre des jugements vrais au nombre total des jugements, sera égal à ce que nous appelons la probabilité de l'événement.

On peut démontrer également que plus on multiplier les jugements, plus il deviendra probable que ces deux rapports s'écarteront très peu l'un de l'autre. Ainsi admettre qu'une probabilité plus grande (ce mot étant pris dans le sens abstrait de la définition) est un motif plus grand de croire, c'est admettre en même temps que ces motifs sont proportionnels aux probabilités.

Cela posé, du moment qu'on admet que, dès que le nombre des combinaisons qui amènent un événement, est plus grand que le nombre des combinaisons qui ne l'amènent pas, on a un motif de croire que l'événement arrivera ; on doit admettre que si la probabilité d'un autre événement est plus grande, le motif de croire sera plus grand aussi. »

---

1. Condorcet, *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Imprimerie Royale, Paris, 1785.

Condorcet a été le premier à tenter d'appliquer le calcul des probabilités à l'évaluation des décisions judiciaires rendues à la majorité des voix. Son ouvrage date de 1785. Il a ainsi précédé Laplace qui reprendra le problème en 1814 dans son *Essai philosophique sur les probabilités* et Poisson qui publia ses *Recherches sur la probabilité des jugements* en 1837.

Il justifie comme suit l'utilisation du calcul des probabilités dans ce domaine :

« En cherchant, d'après la raison seule, quelle confiance plus ou moins grande mérite le jugement d'assemblées plus ou moins nombreuses, assujetties à une pluralité plus ou moins forte, partagées en plusieurs Corps différents ou réunies en un seul, formées d'hommes plus ou moins éclairés ; on sent qu'on ne parviendrait qu'à des résultats vagues, et souvent assez vagues pour devenir incertains, et pour nous induire en erreur si nous les admettions sans les avoir soumis au calcul.

Ainsi, par exemple, on sentiroit aisément qu'en exigeant d'un Tribunal une pluralité plus grande pour condamner un accusé, on acquiert une sûreté aussi plus grande qu'un innocent ne sera pas envoyé au supplice : mais la raison sans calcul ne vous apprendra ni jusqu'à quelles bornes il peut être utile de porter cette sûreté, ni comment on peut la concilier avec la condition de ne pas laisser échapper trop de coupables.

La raison, avec un peu de réflexion, fera sentir la nécessité de constituer un Tribunal de manière qu'il soit presque impossible qu'un seul innocent soit condamné, même dans un long espace de temps ; mais elle n'apprendra ni dans quelles limites on peut renfermer cette probabilité, ni comment y parvenir, sans multiplier le nombre des Juges au-delà des bornes qu'il n'est guère possible de passer.

Ces exemples suffisent pour faire apercevoir l'utilité et, j'oserais presque dire, la nécessité d'appliquer le calcul à ces questions. »

Dans un article intitulé *Les Recherches sur la probabilité des jugements de Simon-Denis Poisson*<sup>1</sup>, Évelyne Barbin et Yannick Marec écrivent :

« Les motivations et les pratiques mathématiques de ces trois auteurs [Condorcet, Laplace, Poisson] sont différentes. Condorcet cherche surtout à protéger l'individu contre une mauvaise décision tandis que Laplace demande au juré de prendre en compte les intérêts de l'individu et ceux de la société. Quant à Poisson, il pense d'abord à garantir l'ordre public et il ne se place plus à l'échelle d'un tribunal, mais à celle d'un pays. »

L'Essai proprement dit de Condorcet analyse onze hypothèses puis la probabilité qu'un jugement soit correct, trois paramètres étant supposés connus : le nombre de votants  $n$ , la majorité requise  $q$  et la probabilité  $v$  de la justesse du jugement de chacun des votants.

Selon Condorcet la probabilité qu'une décision collective ayant obtenu  $q$  votes soit correcte est :

$$\frac{v^q}{v^q + (1-v)^q}.$$

1. *Histoire et Mesure*, 1987, II-2, p. 39-58.



Si l'on exige une majorité de  $q$  votants parmi  $n$ , la probabilité que la majorité sera atteinte et que la décision sera correcte est donnée par la somme :

$$v^n + \sum_{i=1}^{n-q} \binom{n}{i} v^{n-i} (1-v)^i.$$

Condorcet poursuit en examinant l'influence d'un leader sur les votants. Il utilise la loi de Bayes pour établir la probabilité *a posteriori* d'un vote, connaissant le taux de partage de l'opinion du leader dans des scrutins passés.

### 13.3. Les considérations relatives aux élections

Dans un journal révolutionnaire appelé *Journal d'Instruction Sociale*, créé en 1793 par lui-même et par Sieyès et Duhamel, Condorcet a publié un mémoire intitulé *Tableau général de la science qui a pour objet l'application du calcul aux sciences politiques et morales* où il a donné sa célèbre définition de la *Mathématique sociale* qu'il entend créer et qui vise à la compréhension des comportements humains à l'aide d'outils mathématiques. Il projette, en fait, d'élaborer une description statistique de la société, une théorie de l'économie politique et une approche combinatoire des processus mentaux.

Il écrit<sup>1</sup> :

« J'ai cru que le nom de *mathématique sociale* était celui qui convenait le mieux à cette science. Je préfère le mot *mathématique*, quoiqu'actuellement hors d'usage au singulier [...] parce qu'il s'agit d'applications dans lesquelles toutes les méthodes peuvent être employées. [...] Je préfère le mot *sociale* à ceux *morale* ou *politique*, parce que le sens de ces derniers mots est moins étendu et moins précis ».

Dans le numéro du 1<sup>er</sup> juin 1793 du même journal, Condorcet commente ses propres propos relatifs à sa théorie des élections inclus dans l'Essai de 1785. Il déclare :

« Dans une société libre et fondée sur l'égalité, la prospérité publique, la sûreté de l'État, la conservation même des principes de cette société dépend de la bonté des choix populaires.

Si la masse des représentants et des fonctionnaires est composée d'hommes animés de l'esprit public, éclairés, honnêtes, les vices des formes constitutionnelles ne peuvent avoir une influence dangereuse ; et on aura le temps de les réformer avant qu'ils aient pu nuire.

Si, au contraire, des hommes corrompus, des ignorants présomptueux, sont portés aux places, alors les meilleures lois deviennent un faible rempart contre l'ambition et l'intrigue ; et le peuple qui juge sainement, mais qui ne juge que les résultats, se dégoûte, non de la liberté, mais de la tyrannie anarchique à laquelle on en a donné le nom.

1. Cf. Feldman J., *Condorcet et la mathématique sociale. Enthousiasmes et bémols*, Math et Sc. hum., 43<sup>e</sup> année, n° 172, 2005, p. 7-41.

Malgré tous les vices de la constitution anglaise, la force inhérente à toute représentation nationale, dans un pays où la presse est libre, met le ministère dans l'impossibilité de résister à la majorité de la chambre des communes, surtout si, après une convocation nouvelle, elle se trouve animée du même esprit. [...]

J'ai examiné les questions les plus importantes relatives au mode d'élire, dans un ouvrage de calcul, publié en 1785 ; j'ai présenté de nouvelles vues sur ces mêmes questions dans plusieurs écrits imprimés, soit avant, soit depuis la Révolution. [...]

On ne peut faire, en ce genre, des pas utiles, si l'on n'appuie l'analyse métaphysique, les observations morales, les résultats de l'expérience, sur le calcul des combinaisons et la théorie mathématique des probabilités. Toute autre méthode ne conduirait qu'à des conséquences incertaines et vagues. [...]

### **Théorie des élections.**

Pour un homme qui choisirait seul, mais qui voudrait s'astreindre à une marche régulière, une élection est le résultat d'une suite de jugements sur tous les candidats, comparés deux à deux. Les candidats sont, pour lui, ceux qu'il juge dignes de la place, et entre lesquels il doit fixer son choix.

Un vote d'élection est de même le résultat de ces jugements rendus à la majorité. Un homme seul pèse, dans chaque jugement, les motifs de préférence entre deux candidats. Dans une élection, les voix de chacun, pour ou contre, représentent ces motifs, et alors on les compte au lieu de les peser. Un homme qui, comparant deux individus entre eux, préfère le second au premier, puis, comparant ce second à un troisième, donne la préférence à ce dernier, ne peut, sans se contredire, ne pas le préférer aussi au premier. Cependant, s'il comparait ensuite immédiatement le premier et le troisième, peut-être trouverait-il des motifs de préférer le premier ; et alors il serait obligé d'examiner ce jugement, d'en balancer les motifs avec ceux des deux autres qui ne peuvent subsister en même temps, et de sacrifier celui qu'il juge le moins probable.

Dans le cas d'une élection entre trois candidats, il est possible que les trois jugements de la majorité sur ces concurrents, comparés deux à deux, ne puissent subsister ensemble, quoique le résultat des jugements de chaque votant ne renferme aucune contradiction. La possibilité peut aisément s'en prouver par des exemples, et s'expliquer par cette seule observation, que la majorité en faveur de chacune des propositions admises n'est pas formée des mêmes individus. Alors il faut abandonner la proposition qui a une moindre majorité, et s'en tenir aux deux autres. C'est ainsi que, sur tous les objets où l'on ne peut atteindre qu'à une probabilité plus ou moins grande, on doit rejeter une proposition probable en elle-même, si elle en exclut une qui l'est davantage.

La bonté du résultat d'une élection, quoiqu'il renferme une proposition contraire à un des jugements de la majorité, peut rester très probable. En effet, cela n'arrive que dans une combinaison où il est certain que la majorité s'est trompée au moins une fois. Alors la probabilité de la bonté du résultat est celle que la majorité ne s'est trompée qu'une fois, et dans une telle proposition.

Il est inutile d'avertir que cette contradiction entre les jugements de la majorité, qui se présente dans le cas de trois candidats, doit se présenter bien plus aisément quand ils sont en plus grand nombre, et qu'alors plusieurs de ces jugements peuvent être en

contradiction avec les autres ; d'où il suit que, dans ce cas, la majorité a dû nécessairement se tromper plus d'une fois ; mais les conséquences sont les mêmes.

Pour que, dans une assemblée d'électeurs, chacun puisse donner son vœu en entier, c'est-à-dire comparer deux à deux les mêmes candidats, ou, ce qui revient au même, en former la liste suivant l'ordre de mérite, il faut que cette liste soit déterminée, c'est-à-dire que chacun sache ceux sur lesquels, les autres votants peuvent émettre une opinion.

Mais il n'est pas nécessaire que chacun fasse toutes ces comparaisons, compose une liste complète ; il peut en regarder un certain nombre comme égaux entre eux, soit qu'il les juge tels d'après un examen, soit que, ne les connaissant pas, il ne puisse ou ne veuille pas les juger.

Cette condition ne gêne en aucune manière la liberté, puisqu'elle exige seulement que chacun arrête, une fois définitivement, quels sont ceux entre lesquels il croit devoir choisir. La liste de tous ceux qui seraient ainsi présentés offrirait alors à chacun tous les candidats entre lesquels les autres votants pourraient, aussi vouloir juger ; ensuite il déciderait, avec une entière liberté, comment il peut prendre part à ce jugement, quels sont ceux entre lesquels il ne veut pas rester indécis, ou ceux qu'il veut rejeter le plus fortement, en les plaçant après tous les autres.

Tous les modes d'élection où chacun ne donne qu'un vote incomplet, conduisent à des résultats contraires au vœu qu'aurait eu la majorité, si on avait recueilli des votes complets.

La bonté des résultats de ces votes incomplets peut, sans doute, avoir quelque probabilité ; mais elle ressemble à celle d'une proposition qui n'a été examinée qu'à demi ; et nous ne devons adhérer à une proposition probable que dans le cas où l'on aurait découvert l'impossibilité de combiner de nouvelles données, et tant que dure cette impossibilité.

Mais on manquerait également le but, si l'on forçait chaque votant d'émettre, non le vœu complet qu'il forme réellement, mais un vœu complet dans un sens absolu, c'est-à-dire de fixer un ordre de préférence entre tous les candidats, même entre ceux qu'il ne connaît pas. En effet, il est clair qu'alors il rangerait ceux-là au hasard, et qu'ainsi son jugement pourrait faire préférer celui qui ne l'aurait pas été, s'il n'avait pas jugé. Dans le premier cas, on ne compte point des jugements qui auraient dû l'être ; dans le second, on en compte qui n'ont point été rendus. Dans le premier, on agit comme si on avait exclu au hasard une portion des votants ; dans le second, comme si on donnait au hasard une double voix à quelques-uns d'entre eux.

Voici donc quelle serait, dans la théorie, la marche d'une élection : après avoir déterminé ceux entre lesquels on est convenu de choisir, chacun émettrait son vœu complet, soit de préférence, soit d'indifférence.

On formerait le tableau des jugements de la majorité entre les candidats pris deux à deux, et on en tirerait le résultat, c'est-à-dire l'ordre de mérite dans lequel la majorité les a placés. Si ces jugements ne peuvent subsister ensemble, on abandonnerait ceux qui ont obtenu la majorité la plus faible.

Cette marche est précisément la même que suivrait un individu qui voudrait faire un choix avec scrupule, en suivant une méthode générale, régulière, uniforme dans tous les cas.

En général, comme la méthode d'atteindre la vérité est une, il faut que les procédés d'une assemblée délibérante se rapprochent, autant qu'il est possible, de ceux que suit l'esprit d'un seul individu, dans l'examen d'une question.

Ce principe peut avoir d'autres applications importantes ; il nous conduit ici à développer un mode d'élection en quelque sorte naturel, ayant toute la perfection que permet la nature même des choses.

C'est à ce mode qu'il faut comparer tous ceux que l'on peut proposer d'y substituer par des considérations tirées du temps que l'on peut employer aux élections, du degré de lumière de ceux auxquels on les confie, de la difficulté d'écarter d'eux l'influence de l'intrigue, de la nécessité de conserver entre eux une véritable égalité, et de les intéresser à remplir cette fonction avec zèle, avec scrupule. »

### **Paradoxe de Condorcet**

Parmi un même électorat, et lors d'une même élection, constate Condorcet, il est possible qu'une majorité préfère  $A$  à  $B$ , qu'une autre majorité préfère  $B$  à  $C$ , et qu'une troisième majorité préfère  $C$  à  $A$ . Les décisions prises à une majorité populaire par ce mode de scrutin seraient donc *incohérentes* par rapport à celles que prendrait un individu rationnel. Condorcet a précisé lui-même, dans le texte que nous avons cité ci-dessus, comment lever ce paradoxe lié à l'intransitivité, dans certains cas de la majorité.

Dans son Essai de 1785, Condorcet donne l'exemple suivant :

Soit une assemblée de 60 votants ayant le choix entre trois propositions  $A, B$  et  $C$ . Supposons que les votes se répartissent comme suit (en notant  $A > B$ , le fait que  $A$  est préféré à  $B$ ) :

23 votants préfèrent :  $A > B > C$

17 votants préfèrent :  $B > C > A$

2 votants préfèrent :  $B > A > C$

10 votants préfèrent :  $C > A > B$

8 votants préfèrent :  $C > B > A$

Dans les comparaisons binaires, on obtient :

33 préfèrent  $A > B$  contre 27 pour  $B > A$

42 préfèrent  $B > C$  contre 18 pour  $C > B$

35 préfèrent  $C > A$  contre 25 pour  $A > C$

La conclusion  $A > B > C > A$  est incohérente.

Dans le texte cité plus haut, Condorcet propose d'éliminer le candidat le moins performant (ici  $A$  car  $A > B$  remporte le plus faible score) et d'organiser un duel entre  $B$  et  $C$  qui sera remporté par  $B$ .

### 13.4. Biographie de Condorcet (1743-1794)

Marie Jean Antoine-Nicolas Caritat de Condorcet naît le 17 septembre 1743 à Ribemont, petite ville de Picardie. Il est issu d'une famille de vieille, mais petite noblesse provinciale ; son père, Antoine, chevalier de Condorcet, est militaire, mais ne dépassera pas le grade de capitaine ; sa mère, Marie Madeleine est une femme dévote et anxieuse, à la santé fragile. Cinq semaines après la naissance de Nicolas, un drame survient, le chevalier de Condorcet est tué à Neuf-Brisach lors d'une manœuvre d'entraînement. Le jeune Nicolas est alors élevé « avec autant de raison que de douceur » par sa mère, veuve pour la seconde fois. Elle s'occupe seule de son éducation jusqu'à l'âge de neuf ans. Ensuite il est pris en charge pendant deux ans par un précepteur à domicile. Un oncle évêque intervient alors pour l'envoyer en tant qu'interne chez les jésuites de Reims. L'enseignement laisse beaucoup à désirer, quant à la discipline elle est fondée sur les châtiments corporels<sup>1</sup>. Heureusement à l'âge de quinze ans, Condorcet est envoyé au Collège de Navarre à Paris, l'un des plus prestigieux établissements du royaume. La physique y est enseignée par l'abbé Nollet, célèbre pour ses retentissantes expériences d'électricité. Tandis que la première année est consacrée à la philosophie, matière qui n'attire pas le jeune Nicolas, la deuxième année est dédiée aux mathématiques et à la physique. C'est une véritable révélation, Condorcet va se plonger littéralement dans ces matières. Il est reçu bachelier en 1759 et passe un mois plus tard une thèse sur « l'art de résoudre des problèmes » en présence de d'Alembert, Bézout et Grandjean de Fouchy.

Mais Condorcet doit revenir à Ribemont où son avenir, selon les désirs de sa mère et son oncle, est tout tracé : il deviendra officier, comme son père. Il surmonte sa timidité naturelle et s'oppose fermement à ce projet. Entre 1760 et 1762 il reste à Ribemont où il passe le plus clair de son temps à faire des mathématiques dans sa chambre. Il lit les travaux d'Euler, de d'Alembert et des Bernoulli. Il présente, dès 1761, un premier mémoire à l'Académie des sciences, mais son style n'est pas encore très affirmé et les résultats qu'il présente sont déjà connus. Qu'importe, il replonge dans la littérature mathématique de son temps et poursuit ses réflexions. Il obtient néanmoins une victoire : sa famille renonce à faire de lui un officier, il est autorisé à regagner Paris pour y faire une carrière de mathématicien. En janvier 1764 il transmet un second mémoire à l'Académie sur la généralisation et la simplification

1. Cet enseignement laissera des traces indélébiles. Condorcet écrira par exemple à Voltaire en 1774 : « Ne trouvez-vous pas comme moi que la race d'hommes la plus méprisable et la plus odieuse est celles des prêtres catholiques ? »

d'une équation d'Euler. Son travail est, cette fois, très apprécié et le fait connaître. Dès l'année suivante il publie un ouvrage, intitulé *Du calcul intégral*, dont d'Alembert et Bezout soulignent « l'élégance et la profondeur de vue ». Deux ans plus tard, en 1767, il publie un second ouvrage intitulé *Du problème des trois corps*, et y expose une méthode d'approximation relative à ce très délicat problème de mécanique céleste. D'Alembert est subjugué, il n'aura alors de cesse que de faire entrer son jeune protégé à l'Académie des sciences. Cela sera chose faite le 25 février 1769. Condorcet, qui n'a alors que vingt-six ans, est reçu adjoint à la section de mécanique en remplacement de Bezout promu associé<sup>1</sup>. Condorcet ne gagne toutefois pas encore sa vie, il est toujours dépendant de la petite pension que lui verse sa mère. Sa vie ne change guère : il est toujours plongé dans ses études mathématiques. Il sort peu, il préfère lire à écouter, écrire à parler, s'adresser à un petit groupe plutôt qu'à la foule. Il est entouré de l'affection de Julie de Lespinasse, l'amie de d'Alembert<sup>2</sup>. Celle-ci, son aînée de onze ans, est pour lui une véritable seconde mère. Condorcet entretiendra aussi tout au long de sa vie une grande amitié avec Amélie Suard, jeune sœur du célèbre éditeur Panckoucke<sup>3</sup> et qui sera sa confidente. Condorcet a également été entouré de trois pères spirituels : d'Alembert, bien sûr, mais également Turgot et Voltaire. Selon E. et R. Badinter « Le premier lui légua l'amour de la vérité ; le second la passion du bien public ; le troisième, le refus de l'injustice »<sup>4</sup>. Julie de Lespinasse a écrit un jour à Condorcet : « D'Alembert vous aime par goût, par choix et par l'analogie qu'il a avec vos vertus et vos talents ». Quant à Arago, il écrira au sujet des liens qui unissaient Condorcet et Turgot : « Leurs idées, leurs espérances, leurs sentiments s'étaient complètement identifiés. Il serait vraiment impossible de citer un seul point sur lequel Turgot et Condorcet aient différé, même par d'imperceptibles nuances ». Condorcet est par ailleurs un admirateur de Voltaire. En septembre 1770, d'Alembert et Condorcet se rendent à Ferney, pour y rendre visite à Voltaire. Ils seront son hôte du 25 septembre au 10 octobre. Une grande estime réciproque naît, Voltaire écrira suite à cette visite : « Il est très peu de gens de ces temps-là et même de ces temps-ci qu'on puisse comparer à M. d'Alembert et à M. Condorcet. Ils m'ont fait oublier tous mes maux ; je n'ai pu malheureusement les retenir plus longtemps. Les voici partis... »

1. L'Académie est alors divisée en « classes » très hiérarchisées : adjoints, puis associés, puis pensionnaires. Seuls ces derniers sont payés.
2. Julie de Lespinasse et d'Alembert se sont rencontrés en 1754 et petit à petit leur amitié se transforme en liaison (probablement aux alentours de 1760). D'Alembert refusera l'offre du roi de Prusse de présider l'Académie de Berlin pour ne pas s'éloigner de Julie. C'est vers 1765 que d'Alembert se met en ménage avec Julie.
3. Le nom de Panckoucke est resté attaché à la seconde édition (complète) de *la Description de l'Égypte* (« Édition de Panckoucke », 1820) [la première édition (incomplète) dite « Édition Impériale » date de 1809].
4. *Condorcet, un intellectuel en politique*, Élisabeth et Robert Badinter, Fayard, 1988 et 2001.

En 1771, Condorcet tombe éperdument amoureux d'une jeune femme charmante, la jeune madame de Meulan, qu'il a connue par l'intermédiaire de Julie de Lespinasse. Cet amour n'est malheureusement pas réciproque et en 1774 il se résigne à ne plus rien attendre d'elle. Condorcet trouve refuge dans l'étude des mathématiques. Ses recherches concernent essentiellement le calcul intégral, mais il commence aussi à s'intéresser au calcul des probabilités et confie en 1771 à Turgot qu'il « s'amuse à calculer les probabilités et fera un petit livre sur cet objet. » Pendant ce temps, d'Alembert commence à faire campagne pour faire élire Condorcet Secrétaire Perpétuel de l'Académie des sciences, à la place de Grandjean de Fourny jugé lent et inefficace. Il écrit dans ce sens à Lagrange : « Je travaille à faire retomber à notre ami Condorcet [la place de Secrétaire Perpétuel] qui la remplira supérieurement. » Condorcet commence à écrire, à la demande de d'Alembert qui souhaite le faire connaître, des notices biographiques de tous les membres de l'Académie morts entre 1666 et 1669. Il poursuivra par celles d'académiciens contemporains et au total, entre 1773 et 1790, il aura rédigé soixante et un éloges ! Finalement Condorcet sera élu « adjoint avec survivance au Secrétaire perpétuel Grandjean de Fouchy » le 8 mars 1773.

La mort de Louis XV, le 10 mai 1774, marque un tournant. Maurepas, qui fera office de premier ministre de Louis XVI, nomme Turgot ministre de la marine. Condorcet s'écriera : « Du ciel enfin sur nous la bonté se déclare, M. Turgot est nommé ministre de la Marine. » Turgot engage des actions mais ne garde pas longtemps ce portefeuille ; dès le 24 août il est nommé contrôleur général des finances et ministre d'État. Il appelle son ami Condorcet à ses côtés et le charge de concevoir un vaste projet de réforme de l'activité scientifique. Il mène une action d'envergure sur les canaux en France et fait développer la science de l'hydraulique. L'année suivante, en 1775, Turgot fait nommer Condorcet inspecteur des Monnaies. Ce dernier commence à s'atteler à l'immense travail d'unification des Poids et Mesures lorsque Turgot perd sa place de ministre en 1776. L'unification des Poids et Mesure devra, en fait, attendre la Constituante pour être de nouveau à l'ordre du jour avec la mise en place d'une commission réunissant Laplace, Monge et Condorcet. Toutefois une autre charge attend notre mathématicien, il est élu par un vote unanime Secrétaire perpétuel de l'Académie des sciences le 10 août 1776. L'exigence du poste l'aide probablement à surmonter son chagrin profond lié à la mort, en mai 1776, de Julie de Lespinasse, sa si proche confidente. À cette même époque, Panckoucke décide de réaliser une nouvelle édition à l'*Encyclopédie* de Diderot et d'Alembert, et Condorcet est chargé de la rédaction d'articles sur l'analyse mathématique qui ne figuraient pas dans l'édition originale. D'Alembert, abattu par la mort de sa compagne Julie, se repose entièrement sur Condorcet pour ce lourd travail. L'année 1781 marquera un nouveau tournant dans la vie de notre mathématicien, il se voit obligé de vendre la maison de son enfance à Ribemont et son ami Turgot meurt le 18 mars.

Condorcet s'engage alors dans des actions très chères à son défunt ami : la lutte contre l'esclavage des noirs dans les colonies, la défense des droits des protestants puis de ceux des juifs. Il s'implique de nouveau fortement dans une nouvelle édition (par matières) de l'*Encyclopédie* et introduit des sujets nouveaux comme l'*Arithmétique politique* et les *Probabilités*. Fort de ses éloges académiques et du soutien inconditionnel de d'Alembert, Condorcet est élu, le 21 janvier 1782 membre de l'Académie française (encore surnommée « la grande Académie »). D'Alembert ne cache pas sa joie de cette réussite, il se déclare « plus content d'avoir gagné cette victoire, [qu'il] ne le serait d'avoir trouvé la quadrature du cercle. » Ainsi il aura placé son protégé aux postes clefs des Académies : Secrétaire perpétuel de l'Académie des sciences et membre de l'Académie française, la relève est bien assurée.

En septembre 1782 Condorcet a trente-neuf ans et tombe de nouveau éperdument amoureux, cette fois de la jeune Mlle de Boubers (elle a dix-huit ans), la nièce de sa confidente Amélie Suard. Condorcet souhaite l'épouser mais les familles s'y opposent.

L'année suivante une prouesse technique enchante notre mathématicien : le 27 août 1783 il assiste sur le champ de mars à la première ascension en ballon<sup>1</sup>. La même année, le 29 octobre, d'Alembert meurt. Ce dernier a toujours considéré Condorcet comme son fils spirituel et il lui dit en voyant approcher sa fin : « Mon ami, vous ferez mon éloge dans les deux Académies, vous n'avez pas de temps à perdre pour cette double besogne »<sup>2</sup>. En prononçant son éloge, Condorcet n'a pu retenir ses larmes en disant ces mots : « Honoré par lui, dès ma jeunesse, d'une tendresse vraiment paternelle, personne n'a plus à le regretter que moi. »

Après près de treize ans de gestation, Condorcet fait paraître, en 1785 son *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Dans son introduction, Condorcet rend hommage à Turgot pour l'avoir encouragé dans cette voie.

C'est probablement durant l'été 1786, que Condorcet fait la connaissance de Sophie de Grouchy. Pour la troisième fois de sa vie il tombe très amoureux. Cette fois rien n'empêche le mariage et ils s'unissent le 28 décembre. Sophie est jeune, belle, intelligente, partage les idées de son mari et tient salon dans la capitale. Notre mathématicien est au faite du bonheur. Ses préoccupations sont surtout d'ordre scientifique, il réfléchit à l'application du calcul aux sciences politiques et se pose les questions suivantes : Comment

1. Peu de temps après, le 21 novembre 1783, au château de la Muette, à Paris, Pilâtre de Rozier et le marquis d'Arlandes effectuent la première ascension à bord d'un *ballon à air chaud* conçu par les frères Montgolfier. Le 1<sup>er</sup> décembre le physicien français Jacques Charles (1746-1823) et le papetier Nicolas Robert (1761-1828) quittent les Tuileries, à Paris, à bord du ballon qu'ils ont construit et gonflé à l'hydrogène.
2. Par ailleurs d'Alembert fait de Condorcet son légataire universel.



rationaliser les sciences sociales ? Comment fonder une science de l'homme ? Les résultats de ses réflexions ne seront publiés qu'en 1793 dans le *Tableau général de la science*.

Lorsque la Révolution Française éclate, Condorcet joue un rôle majeur et sa vie devient essentiellement politique. Il est en effet le défenseur de nombreuses causes libérales et espère participer activement à une reconstruction rationaliste de la société. Il vit les journées décisives du samedi 11 au vendredi 17 juillet 1789 comme des milliers de spectateurs ou d'acteurs anonymes. La déclaration des droits de l'homme et du citoyen du 26 août 1789 ne peut que le renforcer dans ses motivations politiques. Son engagement prend une nouvelle forme avec son élection au conseil municipal de Paris. Côté familial, il a également la joie de voir naître sa fille, Éliza, le 24 avril 1790. Côté scientifique, il poursuit ses travaux sur le choix des unités de mesure avec Laplace, Lagrange, Borda et Monge. Le 21 mars 1791, l'Académie le charge de présenter le rapport correspondant à l'Assemblée nationale. Peu de temps après, le 8 avril, il est nommé parmi les six administrateurs de la Trésorerie (Lavoisier fait également partie de ces administrateurs).

Condorcet est l'un des premiers à demander l'établissement de la République : le 8 juillet 1791 il prononce un discours sur le thème : *De la République, ou un roi est-il nécessaire à l'établissement de la liberté ?* Il est l'un des seuls, parmi les grandes voix de son temps, à prôner publiquement et solennellement la déposition du roi et l'instauration de la république. Le 26 septembre de la même année il est élu représentant de Paris au sein de l'Assemblée législative. Il devient tour à tour l'un des six secrétaires de cette Assemblée puis vice-président (25 janvier 1792) et même président (7 février 1792). La tâche n'est pas facile dans une Assemblée ayant en majorité foi en la monarchie constitutionnelle. Dans cette Assemblée, Condorcet siège avec les Girondins. Toutefois, se sachant plus écrivain qu'orateur, Condorcet se met également au journalisme et, à partir d'octobre 1791, il assure la chronique parlementaire du *Journal de Paris*.

L'Assemblée adopte, pour le système éducatif de la nation, la structure que propose Condorcet qui introduit la notion fondamentale de laïcité de l'enseignement<sup>1</sup>. Condorcet propose même un brouillon de constitution pour la nouvelle France. De plus, il se prononce pour le droit de vote des femmes. Mais bientôt Condorcet se trouve en mauvaise posture. Deux courants de pensée s'affrontent quant à la manière de réformer l'état français : les Girondins sont favorables à une reconstruction pacifique du pays et les Jacobins, dirigés par Robespierre prônent une purge radicale du passé monarchique français. Condorcet, qui est Girondin et surtout farouchement opposé à la

1. Condorcet présente à l'Assemblée son *Rapport sur l'instruction publique* le 20 avril 1792. Le même jour, devant la même Assemblée, le Roi propose la déclaration de guerre au « roi de Bohême et de Hongrie ».

peine de mort, vote contre l'exécution de Louis XVI et pour les travaux forcés à perpétuité. En 1793, les Girondins perdent le contrôle de l'Assemblée en faveur des Jacobins et le Jacobin Hérault de Séchelles propose une nouvelle constitution, très différente de celle proposée par Condorcet. Ce dernier la critique, ce qui le fait condamner pour trahison. Le 3 octobre 1793, un mandat d'arrêt est délivré contre lui. Condorcet est contraint de se cacher. Il trouve refuge pendant cinq mois dans la demeure de Mme Vernet, rue Servandoni, à Paris. Il en profite pour écrire son *Esquisse d'un tableau historique des progrès de l'esprit humain* qui sera publié après sa mort, en 1795. Le 25 mars 1794, il quitte sa cachette, car il pense ne plus y être en sécurité, et tente de fuir Paris. Il est arrêté à Clamart deux jours plus tard, et mis en prison à Bourg-Egalité, l'actuelle Bourg-la-Reine. On le retrouve deux jours plus tard mort, dans sa cellule. Les circonstances de sa mort restent mystérieuses. La version la plus répandue est que son ami Cabanis lui aurait remis un poison qu'il aurait absorbé dans sa cellule le 29 mars 1794.

### 13.5. Ouvrages scientifiques de Condorcet

- « Mémoire sur le calcul des probabilités », in *Mémoires de l'Académie royale des sciences*, 1781-1784.
- *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris, Impr. royale, 1785.
- *Nouvelles expériences sur la résistance des fluides* (avec d'Alembert et Bossut), Paris, Jombert, 1777.
- *Éléments du calcul des probabilités, et son application aux jeux de hasard, à la loterie, et aux jugemens des hommes par feu M. de Condorcet, avec un discours sur les avantages des mathématiques sociales et une notice sur M. de Condorcet...*, Paris, Royez, an XIII-1805.
- *Esquisse d'un tableau historique des progrès de l'esprit humain*, Paris, An III.

### 13.6. Références

- G. Granger, *Condorcet*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 3, p. 383-388.
- G. Granger, *La Mathématique sociale du marquis de Condorcet*, 1956.
- S. F. Lacroix, *Notice historique sur la vie et les ouvrages de Condorcet*, Paris, 1813.

- B. Bru, *Condorcet, mathématique sociale et vérité*, Math. Inform. Sci. Humaines, 128, 1994.
- P. Crépel, *Condorcet, un mathématicien du social*, La Recherche, 207 (1989), 248-249.
- P. Crépel, « Condorcet, la théorie des probabilités et les calculs financiers », in *Sciences à l'époque de la Révolution Française*, Paris, 1988, 267-325.
- C. Gilain, « Condorcet et le calcul intégral », in *Sciences à l'époque de la Révolution Française*, Paris, 1988, 87-147.
- E. et R. Badinter, *Condorcet, un intellectuel en politique*, Fayard, 1988.

# Lagrange

## 14.1. Le mémoire de Turin sur les probabilités

La principale publication de Lagrange concernant les probabilités est intitulée *Mémoire sur l'utilité de la méthode de prendre le milieu entre les résultats de diverses observations, dans lequel on examine les avantages de cette méthode par le calcul des probabilités et où on résout différents problèmes relatifs à cette matière*. Elle est incluse<sup>1</sup> dans les *Miscellanea Taurinensia*, Tome V, années 1770-1773, publiées en 1776.

Dans ce mémoire, Lagrange traite essentiellement de la moyenne arithmétique et de sa loi de distribution de probabilité<sup>2</sup>. La technique qu'il met en œuvre est une extension des fonctions génératrices telles que les avait utilisées de Moivre. Simpson avait déjà traité le cas des erreurs possibles  $-1, 0, +1$ , Lagrange reprend ce cas puis élargit le problème à d'autres erreurs discrètes puis quelques cas d'erreurs distribuées continûment. Mais si Lagrange fait un usage systématique de ce que nous appelons aujourd'hui la transformée de Laplace pour trouver la distribution de l'erreur moyenne, il ne dispose pas de formule d'inversion pour remonter de la puissance nième de la transformée à la fonction génératrice sauf si cette puissance nième est à l'évidence elle-même une transformée.

Ce travail de Lagrange, paru en 1776, a incité Laplace à reconsidérer sa propre approche sur le principe de probabilité inverse. Laplace formalisera en 1785 et en 1810, la technique des transformations qui portent son nom, développera une formule d'inversion et utilisera ces outils pour démontrer le théorème central limite.

Rappel : [Une fonction  $f(t)$  définie pour  $0 \leq t < \infty$ , a pour transformée de Laplace

$$F(s) = \int_0^{\infty} e^{-st} f(t) dt .$$

1. *Miscellanea Taurinensia*, Tome V, années 1770-1773 : Mélanges de Philosophie et de mathématique de la société royale de Turin classe mathématique p. 167-232.
2. Cf., par exemple : Todhunter L., *A History of the mathematical theory of probability*, Cambridge, 1865, chapitre XV ; Dale A. I., *A history of inverse probability*, Springer Verlag, 1991, p. 82-86 ; Hald A., *A history of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 40-50.

L'inversion de la transformée de Laplace peut être effectuée par la formule de Post :

Une fonction  $f(t)$  continue sur l'intervalle  $[0, \infty)$  qui pour  $b \in \mathbb{R}$  vérifie l'inégalité

$\sup_{t>0} \frac{|f(t)|}{e^{bt}} < \infty$ , existe et est différentiable par rapport à  $s$ , pour tout  $s > b$ ,

peut être retrouvée à partir de sa transformée de Laplace  $F(s) = \int_0^\infty e^{-st} f(t) dt$  par la formule

$$f(t) = \lim_{k \rightarrow \infty} \frac{(-1)^k}{k!} \left( \frac{k}{t} \right)^{k+1} F^{(k)} \left( \frac{k}{t} \right)$$

pour  $t > 0$ ,  $F^{(k)}$  étant la  $k$ ème dérivée de  $F$ .]

Lagrange, dans son mémoire présente tout d'abord l'objet de ses travaux :

« Quand on a plusieurs observations d'un même phénomène dont les résultats ne sont pas tout à fait d'accord, on est sûr que ces observations sont toutes, ou au moins en partie, peu exactes, de quelque source que l'erreur puisse provenir ; alors on a coutume de prendre le milieu entre tous les résultats, parce que de cette manière, les différentes erreurs se répartissant également dans toutes les observations, l'erreur qui peut se trouver dans le résultat moyen devient aussi moyenne entre toutes les erreurs. Or, quoique tout le monde reconnaisse l'utilité de cette pratique pour diminuer, autant qu'il est possible, l'incertitude qui naît de l'imperfection des instruments et des erreurs inévitables des observations, j'ai cru cependant qu'il serait bon d'examiner et d'apprécier par le calcul les avantages qu'on peut espérer retirer d'une semblable méthode ; c'est l'objet que je me suis proposé dans ce Mémoire. Je commencerai par supposer que les erreurs qui peuvent se glisser dans chaque observation soient données, et qu'on connaisse aussi le nombre des cas qui peuvent donner ces erreurs, c'est-à-dire la facilité de chaque erreur ; je supposerai ensuite que l'on connaisse seulement les limites entre lesquelles toutes les erreurs possibles doivent être renfermées avec la loi de leur facilité, et je chercherai dans l'une et dans l'autre de ces hypothèses quelle est la probabilité que l'erreur du résultat moyen soit nulle, ou égale à une quantité donnée, ou seulement comprise entre des limites données. Je ferai voir en même temps comment on peut déterminer, *a posteriori*, la loi même de la facilité des erreurs, et quelle est la probabilité que dans cette détermination on ne se trompera pas d'une quantité donnée. [...]

Au reste, je suivrai dans toutes ces recherches la règle ordinaire du calcul des probabilités, suivant laquelle on estime la probabilité d'un événement par le nombre des cas favorables, divisé par le nombre de tous les cas possibles. »

Lagrange traite ensuite onze problèmes, le neuvième étant plutôt l'énoncé de divers lemmes.

Voici l'énoncé par Lagrange du premier problème :

« On suppose que dans chaque observation on peut se tromper d'une unité, tant en plus qu'en moins, mais que le nombre des cas qui peuvent donner un résultat exact

est au nombre des cas qui peuvent donner une erreur d'une unité comme  $a : 2b$  ; on demande quelle est la probabilité d'avoir un résultat exact en prenant le milieu entre les résultats particuliers d'un nombre  $n$  d'observations. »

Lagrange poursuit :

« Puisqu'il y a  $a$  cas qui donnent zéro d'erreur, et  $2b$  cas qui donnent  $+1$  et  $-1$ , c'est-à-dire  $b$  cas qui donnent  $+1$  et  $b$  cas qui donnent  $-1$  d'erreur, il est clair par les règles ordinaires des probabilités que la probabilité que l'erreur soit nulle dans chaque observation particulière sera exprimée par  $\frac{a}{a+2b}$  : voyons donc quelle sera la probabilité que l'erreur soit aussi nulle en prenant le milieu entre  $n$  observations. Il est facile de montrer que cette question se réduit à celle-ci : Ayant  $n$  dés dont chacun ait  $a$  faces marquées d'un zéro,  $b$  faces marquées d'une unité positive, et  $b$  faces marquées d'une unité négative, en sorte que le nombre total des faces soit  $a + 2b$ , trouver la probabilité qu'il y a d'amener zéro en jetant tous ces dés au hasard. »

Or, Simpson avait abordé en 1755 le problème des observations astronomiques en se focalisant sur les erreurs de positionnement du corps observé, avait considéré l'erreur comme une variable aléatoire distribuée uniformément sur les entiers entre  $-a$  et  $+a$ , et avait utilisé un résultat établi par de Moivre, qui, parmi les différents problèmes liés aux jeux de dés, avait étudié la distribution de la somme des faces de plusieurs dés<sup>1</sup>. Le résultat indiquait que cette distribution de la somme des faces de  $n$  dés ayant chacun  $f$  faces numérotées 1 à  $f$  est donnée par les coefficients du développement de

$$(1 + r + r^2 + r^3 + \dots + r^{f-1})^n = \frac{(1 - r^f)^n}{(1 - r)^n}.$$

Lagrange développe donc une solution identique à celle de Simpson, qu'il ne connaissait peut-être pas.

La contribution principale du mémoire est développée dans le dixième problème.

Lagrange donne dans les douze dernières pages de son papier une méthode directe pour traiter une distribution de probabilité des erreurs quelconque et continue.

Il écrit :

« On suppose que chaque observation soit sujette à toutes les erreurs possibles comprises entre ces deux limites,  $p$  et  $-q$ , et que la facilité de chaque erreur  $x$ , c'est-à-dire le nombre des cas où elle peut avoir lieu divisé par le nombre total des cas, soit représentée par une fonction quelconque de  $x$  désignée par  $y$  ; on demande la probabilité que l'erreur moyenne de  $n$  observations soit comprise entre les limites  $r$  et  $-s$ .

1. *Doctrine of Chances*, 1738, p. 35-37. Cf. notre § 6.2.

On commencera d'abord par chercher la probabilité que l'erreur moyenne soit  $z$ , et cette probabilité étant représentée par une fonction de  $z$ , il n'y aura qu'à en prendre l'intégrale depuis  $z = r$  jusqu'à  $z = s$ ; ce sera la probabilité cherchée.

Maintenant, pour avoir la probabilité que l'erreur moyenne de  $n$  observations soit  $z$ , il faudra considérer le polynôme qui est représenté par l'intégrale de  $ya^x dx$ , en supposant cette intégrale prise de manière qu'elle s'étende depuis  $x = p$  jusqu'à  $x = -q$ ; on élèvera ce polynôme à la puissance  $n$ , et l'on cherchera le coefficient de puissance  $z$  de  $a$ , [...] ce coefficient, qui sera une fonction de  $z$ , exprimera la probabilité que l'erreur moyenne soit  $z$ . »

En fait, Lagrange procède par analogie (mais sans démonstration) avec le cas d'une distribution de probabilité discrète  $y(x)$ ,  $x = -q, -q + 1, \dots, p$ , pour lequel la fonction génératrice est

$$g(a) = \sum_{x=-q}^p a^x \cdot y(x), \quad a \geq 0,$$

à partir de laquelle, la fonction génératrice pour la distribution de la somme  $z$  est

$$g^n(a) = \left( \sum_{x=-q}^p a^x \cdot y(x) \right)^n = \sum_{x=-nq}^{np} a^z \cdot y_n(z).$$

En remplaçant les sommes par des intégrales, il définit la fonction génératrice pour la distribution  $y(x)$ ,  $-q \leq x \leq p$ , d'une variable aléatoire continue, soit

$$g(a) = \int_{-q}^p a^x \cdot y(x) dx, \quad a \geq 0,$$

$$\text{et } g^n(a) = \left( \int_{-q}^p a^x \cdot y(x) dx \right)^n = \int_{-nq}^{np} a^z \cdot y_n(z) dz.$$

Pour le cas discret  $g^n(a)$  est un polynôme en  $a$  et  $y_n(z)$  est le coefficient de  $a^z$  dans le polynôme. Pour le cas continu, Lagrange considère que  $\left( \int_{-q}^p a^x \cdot y(x) dx \right)^n$  doit être transformé en un polynôme en  $a$ , l'intégrale étant la limite d'une somme en utilisant de petits intervalles. À nouveau,  $y_n(z)$  est le coefficient de  $a^z$  dans ce polynôme. Ce coefficient sera une fonction de  $z$  et devra être intégré de  $-s$  à  $r$  pour obtenir la probabilité que l'erreur moyenne de  $n$  observations soit comprise entre ces limites.

Lagrange donne à la fin de son mémoire quelques exemples d'applications de sa méthode.

## 14.2. Biographie de Lagrange (1736-1813)

Joseph Louis Lagrange<sup>1</sup> naît à Turin le 25 janvier 1736. Il possède une double origine, française (sa famille paternelle est tourangelle) et italienne. Son père,

1. Giuseppe Luigi Lagrangia.

trésorier du roi de Sardaigne, a épousé Marie-Thérèse Gros qui lui a donné onze enfants, dont seul le plus jeune, notre futur mathématicien, survécut au-delà de l'enfance. Les parents de Lagrange sont riches. Mais le père est un incorrigible spéculateur et Joseph-Louis ne recevra rien de sa fortune, ce qui est peut-être une chance car, comme il le disait lui-même : « Si j'avais hérité d'une fortune, je n'aurais jamais consacré ma vie aux mathématiques ». Au reste, Lagrange ne montre pas immédiatement d'intérêt pour ce domaine. À l'école, ses études sont classiques et en mathématiques il apprend la géométrie d'Euclide et d'Archimède. Mais le hasard le conduit à lire, vers l'âge de dix-sept ans, un livre d'Edmond Halley qui le marque profondément. Il est littéralement subjugué par la supériorité des calculs conduits par Halley sur ceux des Grecs anciens. Il décide d'apprendre par lui-même ces nouvelles techniques de mathématiques si puissantes pour la résolution des problèmes d'astronomie. Son « auto-apprentissage » est fulgurant : Delambre indique (dans son éloge de Lagrange), qu'il devient professeur de mathématiques à l'École Royale d'Artillerie de Turin à dix-neuf ans. Ses élèves sont plus âgés que lui. La même année (1755) il développe la méthode des maxima et minima (*i.e.* le calcul des variations<sup>1</sup>) et commence les travaux qui vont le conduire à son chef-d'œuvre : *la Mécanique analytique*<sup>2</sup>. Malgré son jeune âge il transmet à Euler, dont la réputation est immense, ses premiers travaux. Ce dernier reconnaît immédiatement les talents du jeune Lagrange et l'encourage à poursuivre. Quatre ans plus tard, en 1759, notre mathématicien envoie à Euler de nouveaux résultats. L'illustre mathématicien allemand lui répond : « Votre solution du problème des isopérimètres ne laisse rien à désirer, et je me réjouis que ce sujet, dont je m'étais presque seul occupé depuis les premières tentatives, ait été porté par vous au plus haut degré de perfection ». Euler décide alors de retarder sa propre publication sur le sujet pour « ne pas priver [Lagrange] d'aucune part de gloire qui lui est due ». Ensuite, Euler publie un article indiquant que beaucoup de difficultés avaient pu être surmontées grâce aux indications de Lagrange.

Grâce à Euler, Lagrange est élu, le 2 octobre 1759, membre étranger de l'Académie de Berlin. Cette reconnaissance est, compte tenu de son jeune âge (23 ans), exceptionnelle. Elle va lui être utile pour se faire connaître dans son propre pays. Toujours en 1759, Lagrange applique le calcul différentiel à la théorie des probabilités et développe une théorie du son qui dépasse celle de Newton. Il travaille également sur une théorie de la libration de la lune<sup>3</sup> qui lui vaut de recevoir le grand prix de l'Académie des sciences de Paris en 1764. Encouragée par ce brillant succès, l'Académie met en concours un

1. Le terme « calcul des variations » a d'ailleurs été introduit par Lagrange.

2. Hamilton qualifiera ce traité de « poème scientifique ».

3. La libration est le balancement apparent de la face visible de la lune de part et d'autre de sa position moyenne. Ce balancement est dû à l'ellipticité de l'orbite, à l'inclinaison de l'axe de ses pôles et à la rotation de la terre.



sujet encore plus difficile relatif au mouvement des satellites de Jupiter. En 1766, Lagrange emporte à nouveau le prix de l'Académie. La même année, le roi de Sardaigne, ébloui par ces travaux, décide de lui offrir un voyage pour Paris et Londres. Des ennuis de santé obligent Lagrange à rester à Paris. Il y rencontre un grand nombre d'intellectuels. Notre mathématicien aime Paris, mais le roi de Prusse, Frédéric II, le fait venir à Berlin en disant : « Le plus grand roi d'Europe doit avoir à sa cour le plus grand mathématicien ». Lagrange arrive à Berlin en novembre 1766 et devient directeur de la division de mathématiques de l'Académie de Berlin. Il restera vingt ans dans cette ville et y rédigera un grand nombre de mémoires. En 1767, Lagrange se marie avec une de ses cousines. Le couple ne semble pas très heureux et n'aura pas d'enfant. Lui est sujet à des périodes de dépression et elle est atteinte d'une lente maladie, dont elle mourra en 1783. Notre mathématicien se réfugie dans le travail, « mon occupation, dit-il, est réduite à faire des mathématiques, tranquillement et en silence ». À Berlin ses travaux portent essentiellement sur la mécanique céleste, mais il aborde également d'autres thèmes. C'est ainsi qu'il indique en 1768 à son ami d'Alembert<sup>1</sup> « qu'il diversifie ses études dans des problèmes d'arithmétique ». Il résout également plusieurs conjectures difficiles dues à Fermat. En 1772, Lagrange emporte un troisième prix de l'Académie de Paris pour son essai sur le problème des trois corps, essai qui ne sera dépassé qu'un siècle plus tard. C'est également à cette époque qu'il introduit le potentiel de gravitation (1772) et étudie le mouvement de rotation d'un solide (1773). Dans une lettre à Laplace en date du 15 septembre 1782, Lagrange indique qu'il a « presque complètement terminé son traité sur la mécanique analytique », mais le livre (dont l'édition sera assurée grâce à Legendre), ne paraîtra qu'en 1788. Lagrange est maintenant à Paris et y restera jusqu'à sa mort en 1813. Lorsque Frédéric II meurt le 17 août 1786, le climat à Berlin devient inconfortable pour les étrangers, de plus Louis XVI invite Lagrange à poursuivre ses travaux à l'Académie des sciences de Paris. À son arrivée à Paris, en 1787, il est reçu avec grand respect tant par la famille royale que par l'Académie<sup>2</sup>. Il logera au Louvre jusqu'à la révolution et sera l'un des favoris de Marie-Antoinette. Lavoisier (1743-1794) devient l'un de ses amis et notre mathématicien déclare que grâce à ce dernier la chimie est devenue « aussi simple que l'algèbre ». Sa pondération et son sens de la diplomatie lui permettent de traverser sans heurt la période de la révolution, et surtout de la terreur, malgré ses attaches avec la famille royale. Ainsi, pour contourner le décret du 16 octobre 1793 forçant les étrangers (dont faisait partie Lagrange) à quitter la France, Guyton de Morveau fait prendre un

1. D'Alembert encourage toute sa vie Lagrange, de 20 ans son cadet. Un mois avant sa mort il lui écrit : « Je vous dis au revoir, peut-être pour la dernière fois. Garder un peu le souvenir de la personne qui, de tous les gens du monde, vous chérissait et vous admirait le plus ».
2. Lagrange est nommé « pensionnaire vétérane » le 29 juillet 1787. Il avait été admis à l'Académie de Paris comme « associé étranger » dès 1772.

arrêté par le Comité de Salut Public pour faire rester Lagrange en France afin qu'il puisse continuer ses calculs sur la théorie du mouvement des projectiles. On sait que malheureusement, au nom du slogan « la république n'a pas besoin de savants », son ami Lavoisier sera guillotiné, ce qui fera dire à Lagrange « Seul un instant a été nécessaire pour couper cette tête, et il faudra probablement plus d'un siècle pour en faire une semblable ».

En 1792, alors qu'il a cinquante-six ans, Lagrange traverse une phase dépressive. Il parle peu, est distrait et profondément mélancolique. La fille de l'astronome Lemonnier se désole de le voir dans cet état. Malgré son jeune âge, elle lui propose de l'épouser. Lagrange accepte et... retrouve la joie de vivre auprès de cette jeune femme qui lui est toute dévouée. Il retrouve son activité et sa production scientifique. Il est président de la commission (créée par un décret du 8 mai 1790 de l'Assemblée Constituante) chargée d'établir un *Système général et uniforme de poids et mesure*. Entre 1790 et 1799, il fait également partie de divers comités chargés d'élaborer le système métrique, que la révolution a décidé de mettre en place. Il est, de même, appelé à être professeur dans les deux nouvelles institutions qui viennent d'être créées : l'École Normale et l'École Polytechnique. L'École Normale est fondée le 30 octobre 1794. Les premiers cours ont lieu au mois de décembre. Lagrange y enseigne les mathématiques. À la demande de Monge, il est nommé professeur à Polytechnique<sup>1</sup> dès ses débuts (le premier cours a lieu le 21 décembre 1794). Ses leçons sont « si savantes, dit-on, qu'il ne se trouve peut-être pas en France trente personnes en état de les entendre ». Lagrange devient également, le 4 décembre 1794, le premier président du Conseil d'Administration de cette École. Il développe la théorie des fonctions et publie, en 1797, son important traité intitulé : *La théorie des fonctions analytiques*. Il y étudie de nombreux cas d'intégration d'équations différentielles aux dérivées partielles et on lui doit aussi la technique des « multiplicateurs de Lagrange » (que l'on trouve déjà dans sa mécanique analytique) et celle de la « variation des constantes ».

Lagrange est apprécié et a le sens de la diplomatie : après avoir servi à la cour du roi de Prusse et à celle de Louis XVI, il traverse sans encombre la révolution française et la terreur et finalement sera admiré par Napoléon Bonaparte. Ce dernier le couvre d'honneurs : Sénateur, Comte d'Empire, grand officier de la Légion d'Honneur<sup>2</sup>.

Lagrange meurt le 14 avril 1813, à 77 ans « n'ayant haï personne et n'ayant point fait de mal » (Hamilton), et entre au Panthéon trois jours plus tard.

- 
1. Il s'agit alors de l'École Centrale des Travaux Publics qui sera appelée École Polytechnique à partir du 1<sup>er</sup> septembre 1795.
  2. Lagrange reçoit cette distinction le 3 avril 1813, une semaine avant sa mort.

Son œuvre est immense. Une grande part se trouve dans le domaine des mathématiques pures (calcul des variations, résolution algébrique des équations, théorie des fonctions...). Hamilton le baptise « le Shakespeare des mathématiques ». Son apport à la mécanique céleste et à la physique n'en est pas moins essentiel. L'œuvre de Newton l'a profondément marqué dès sa jeunesse : « Il n'y a qu'une seule loi de l'Univers, dit Lagrange, et c'est Newton qui l'a découverte ».

En développant la mécanique analytique, Lagrange a offert à la physique l'un de ses plus beaux bijoux qui se révélera d'une fécondité presque intemporelle.

Le génie de Lagrange est traduit par les paroles de Laplace : « Il possédait au plus haut point ce tact heureux qui, faisant discerner dans les objets les principes généraux qu'ils recèlent, constitue le véritable génie des sciences, dont le but est la découverte de ces principes, ce tact, joint à une rare élégance dans l'exposition des théories les plus abstraites, caractérise Lagrange ».

### 14.3. Principales publications de Lagrange

Différents mémoires de Lagrange ont été publiés dans : *Mélanges de philosophie et de mathématique de la Société royale de Turin*<sup>1</sup>, *Mémoires de l'Académie royale des sciences et belles lettres de Berlin*, *Nouveaux mémoires de l'Académie de Berlin*, *Journal de l'École Polytechnique*, *Mémoires de l'Académie des sciences de l'Institut de France*, *Mémoires de la classe de mathématiques et physiques de l'Institut de France*.

Les éditions originales principaux livres sont les suivantes :

- *Leçons sur le calcul des fonctions* in *Séances des Écoles Normales*, 1795-1796.
- *Lettera di Luigi de La Grange Tournier, Torinese all' illustrissimo signor conte Giulio Carlo da Fagnano... Contenante una nuova serie per differenziali, ed integrali di qualsivoglia grado corrispondente alla Newtoniana per le potestà et le radici*, Turin, Stamparia Reale, 1754<sup>2</sup>.

1. En 1757, Beccaria et certains de ses étudiants, dont Lagrange, le comte Saluzze et le médecin Giovanni Cigna créent la *Societas Privata Taurinensis* qui devient deux ans plus tard la *Société royale des sciences de Turin*. La société publie un seul volume intitulé *Miscellanae philosophico-mathematica societatis privatae Taurinensis*. Les années suivantes quatre volumes successifs paraissent sous l'appellation *Mélanges de philosophie et de mathématique de la Société royale de Turin*. Dans le volume de 1762 (p. 173-195) on trouve le mémoire de Lagrange : *Essai d'une nouvelle méthode pour déterminer les maxima, et les minima des formules intégrales indéfinies*.
2. Première publication de Lagrange où il développe un calcul basé sur l'analogie entre la théorie binomiale de Newton et les différentiations successives de deux fonctions. Lagrange découvrit plus tard qu'une méthode de ce type avait été déjà développée par Leibniz et Jean Bernoulli.

- *Mécanique analytique*, Paris, Veuve Desaint, 1788.
- *Théorie des fonctions analytiques contenant les principes du calcul différentiel, dégagés de toute considération d'infiniment petits et d'évanouissans de limites ou de fluxions, et réduits à l'analyse algébrique des quantités finies*, Paris, Imprimerie de la République, 1797.
- *De la résolution des équations numériques de tous les degrés*, Paris, Duprat, 1798.

## 14.4. Références

- J. Itard, *Lagrange*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 7, p. 559-573, 1970.
- E. T. Bell, *Men of mathematics*, Simon et Schuster, New York, 1986.
- J. B. Delambre, *Notice sur la vie et les ouvrages de M. le Comte J.-L. Lagrange*, Mémoires de la classe des sciences mathématique de l'Institut 1812, Paris, 1816.
- C. B. Boyer, *A history of mathematics*, New York, 1968.



# Legendre

## 15.1. Nouvelles méthodes pour la détermination des orbites des comètes

Dans cet ouvrage, paraît la première publication de la méthode des moindres carrés qui est le fondement de l'analyse statistique. Gauss, dans son traité *Theoria motus corporum coelestium* publié en 1809, indique qu'il utilisait cette technique depuis 1795, mais de fait la méthode a été publiée en premier par Legendre<sup>1</sup>. L'affirmation de Gauss irrita Legendre<sup>2</sup>.

1. Gauss écrit en 1809 : « Au reste, ce principe, dont nous avons fait usage depuis 1795, a été donné dernièrement par Legendre dans ses *Nouvelles méthodes pour la détermination des orbites des comètes*, Paris, 1806 ; on trouvera dans cet ouvrage plusieurs conséquences que le désir d'abrégier nous a fait omettre ».

2. Legendre a écrit à Gauss, en 1809 (Gauss, Werke, Band. 10/1, p. 380) :

« Je ne vous dissimulerai donc pas, Monsieur, que j'ai éprouvé quelque regret de voir qu'en citant mon mémoire..., vous dites *principum nostrum*... Il n'est aucune découverte qu'on ne puisse s'attribuer en disant qu'on avait trouvé la même chose quelques années auparavant ; mais si on n'en fournit pas la preuve en citant le lieu où on l'a publiée, cette assertion devient sans objet et n'est plus qu'une chose désobligeante pour le véritable auteur de la découverte. En Mathématiques il arrive très souvent qu'on trouve les mêmes choses qui ont été trouvées par d'autres et qui sont bien connues ; c'est ce qui m'est arrivé nombre de fois, mais je n'en ai point fait mention et je n'ai jamais appelé *principum nostrum* un principe qu'un autre avait publié avant moi.

Vous êtes assez riche de [votre] fonds, Monsieur, pour n'avoir rien à envier à personne ; et [je suis] bien persuadé au reste que j'ai à me plaindre de l'expression seulement et nullement de l'intention... »

Dans une lettre à Laplace datée de 1812 (Werke, Band 10/1 p. 373-374) Gauss déclare, pour sa part :

« J'ai fait usage de la méthode des moindres carrés depuis l'an 1795 [...] Cependant mes applications fréquentes de cette méthode ne datent que de l'année 1802, depuis ce temps j'en fais usage pour ainsi dire tous les jours dans mes calculs astronomiques sur les nouvelles planètes. [...] Au reste j'avais déjà communiqué cette même méthode, beaucoup avant la publication de l'ouvrage de M. Legendre, à plusieurs personnes, entre autres à Mr. Olbers en 1803. Ainsi, pouvais-je dans ma théorie parler de la méthode des moindres carrés, dont j'avais fait depuis sept ans mille et mille applications [...]. Pouvais-je parler de ce principe, – que j'avais annoncé à plusieurs de mes amis déjà en 1803 comme devant faire partie de l'ouvrage que je préparais, – comme d'une méthode empruntée de Mr. Legendre ? Je n'avais pas l'idée, que Mr. Legendre pouvait attacher tant de prix à une idée aussi simple, qu'on doit plutôt s'étonner qu'on ne l'a pas eue depuis 100 ans, pour se fâcher que je raconte, que je m'en suis servi avant lui ?... Mais j'ai cru que tous ceux qui me connaissent le croiraient même sur ma parole, ainsi que je l'aurais cru de tout mon cœur si Mr. Legendre avait avancé, qu'il avait possédé la méthode déjà avant 1795. »

**NOUVELLES MÉTHODES**  
**POUR LA DÉTERMINATION**  
**DES**  
**ORBITES DES COMÈTES;**

**PAR A. M. LEGENDRE,**

Membre de l'Institut et de la Légion d'honneur, de la Société  
royale de Londres, &c.

---

**A PARIS,**

Chez **FIRMIN DIDOT**, Libraire pour les Mathématiques, la Marine,  
l'Architecture, et les Éditions stéréotypes, rue de Thionville, n° 116.

**AN XIII — 1805.**

La première édition de ce texte sur la méthode des moindres carrés apparaît dans l'annexe (pages 72-80) du livre intitulé *Nouvelles méthodes pour la détermination des orbites des comètes* (1805). Il est daté du 15 ventôse an XIII (6 mars 1805). Le texte fut reproduit par Puissant dans son *Traité de Géodésie*, édité par Courcier toujours en 1805 (pages 137-141) puis republié dans la deuxième édition de l'ouvrage de Legendre en 1806. L'originale des *Nouvelles méthodes pour la détermination des orbites des comètes* de 1805 a pour éditeur Didot, et l'édition de 1806 a été éditée par Courcier.

La méthode des moindres carrés est qualifiée par Stigler dans son *History of Statistics* (Belknap Press, 1986) comme le thème dominant – le leitmotiv – des mathématiques statistiques au XIX<sup>e</sup> siècle, il ajoute que la clarté de l'exposition de Legendre n'a pas été surpassée et que ce texte doit être compté dans l'histoire de la statistique comme l'une des introductions d'une nouvelle méthode les plus claires et les plus élégantes. Pour sa part, J. Bertrand, dans l'introduction de sa traduction de divers mémoires de Gauss intitulée *Méthode des moindres carrés*, (1855) affirme que Gauss considérait que la meilleure méthode pour combiner les observations était *un des problèmes les plus importants de la philosophie naturelle*.

On notera que la méthode utilisée par Legendre pour justifier la méthode des moindres carrés n'est pas probabiliste mais purement algébrique.

Pour Gauss, en 1809, La justification du principe des moindres carrés sera la maximalisation de la distribution *a posteriori* des erreurs.

Dans son papier de 1810, Laplace justifiera la méthode des moindres carrés par la minimisation de l'erreur estimée *a posteriori* sur un grand nombre d'échantillons.

Dans celui de 1811, Laplace indiquera que la méthode des moindres carrés est justifiée pour un grand nombre de mesures, parce qu'elle est équivalente à l'obtention d'une erreur moyenne minimale (ce qui équivaut à une variance minimale). Il utilisera alors une démonstration à l'aide des probabilités directes et non inverses comme dans sa démonstration de 1810.

En 1823, Gauss utilisera également la minimalisation des variances.

Nous reproduisons ci dessous la partie théorique de la publication de Legendre ainsi que le début de l'application numérique au calcul d'un arc de méridien de France effectué en 1795.

#### « Sur la Méthode des moindres carrés

Dans la plupart des questions où il s'agit de tirer des mesures données par l'observation les résultats les plus exacts qu'elles peuvent offrir, on est presque toujours conduit à un système d'équations de la forme

$$E = a + bx + cy + fz + \text{etc.}$$



dans lesquelles  $a, b, c$ , etc., sont des coefficients connus, qui varient d'une équation, à l'autre, et  $x, y, z$  etc., sont des inconnues qu'il faut déterminer par la condition que la valeur de  $E$  se réduise, pour chaque équation, à une quantité ou nulle ou très petite.

Si l'on a autant d'équations que d'inconnues  $x, y, z$  etc., il n'y a aucune difficulté pour la détermination de ces inconnues, et on peut rendre les erreurs  $E$  absolument nulles. Mais le plus souvent, le nombre des équations est supérieur à celui des inconnues, et il est impossible d'anéantir toutes les erreurs. Dans cette circonstance, qui est celle de la plupart des problèmes physiques et astronomiques, où l'on cherche à déterminer quelques éléments importants, il entre nécessairement de l'arbitraire dans la distribution des erreurs et on ne doit pas s'attendre que toutes les hypothèses conduiront exactement aux mêmes résultats ; mais il faut surtout faire en sorte que les erreurs extrêmes, sans avoir égard à leurs signes, soient renfermées dans les limites les plus étroites qu'il est possible.

De tous les principes qu'on peut proposer pour cet objet, je pense qu'il n'en est pas de plus général, de plus exact, ni d'une application plus facile que celui dont nous avons fait usage dans les recherches précédentes, et qui consiste à rendre *minimum* la somme des carrés des erreurs. Par ce moyen, il s'établit entre les erreurs une sorte d'équilibre qui, empêchant les extrêmes de prévaloir, est très propre à faire connaître l'état du système le plus proche de la vérité. La somme des carrés des erreurs  $E^2 + E'^2 + E''^2 + \text{etc.}$ , étant :

$$\begin{aligned} & (a + bx + cy + fz + \dots)^2 \\ & + (a' + b'x + c'y + f'z + \dots)^2 \\ & + (a'' + b''x + c''y + f''z + \dots)^2 + \text{etc.} \end{aligned}$$

si l'on cherche son *minimum* en faisant varier  $x$  seule, on aura l'équation :

$$0 = \int ab + x \int b^2 + y \int bc + z \int bf + \text{etc.}$$

dans laquelle par  $\int ab$  on entend la somme des produits semblables  $ab + a'b' + a''b'' + \text{etc.}$  par  $\int b^2$  la somme des carrés des coefficients de  $x$ , savoir  $b^2 + b'^2 + b''^2 + \text{etc.}$ , ainsi de suite.

Le *minimum*, par rapport à  $y$  donnera semblablement :

$$0 = \int ac + x \int bc + y \int c^2 + z \int cf + \text{etc.}$$

et le *minimum* par rapport à  $z$ ,

$$0 = \int af + x \int bf + y \int cf + z \int f^2 + \text{etc.}$$

où l'on voit que les mêmes coefficients  $\int bc, \int bf$ , etc., sont communs aux deux équations, ce qui contribue à faciliter le calcul. En général, pour former l'équation du minimum par rapport à l'une des inconnues, il faut multiplier tous les termes de chaque équation proposée par le coefficient de cette inconnue dans cette équation, pris avec son signe, et faire une somme de tous ces produits.

On obtiendra de cette manière autant d'équations du minimum qu'il y a d'inconnues, et il faudra résoudre ces équations par les méthodes ordinaires. Mais on aura soin d'abrégier tous les calculs, tant des multiplications que de la résolution, en n'admettant dans chaque opération que le nombre de chiffres entiers ou décimaux que peut exiger le degré d'approximation dont la question est susceptible.

Si par un hasard singulier, il était possible de satisfaire à toutes les équations en rendant toutes les erreurs nulles, on obtiendrait également ce résultat par les équations du minimum ; car si après avoir trouvé les valeurs de  $x, y, z$ , etc., qui rendent nulles  $E, E'$ , etc., on fait varier  $x, y, z$ , etc., de  $\delta x, \delta y, \delta z$ , etc., il est évident que  $E$  qui était

zéro deviendra par cette variation  $(a\delta x + b\delta y + c\delta z, \text{etc.})^2$ . Il en sera de même de  $E^2, E'^2, \text{etc.}$  D'où l'on voit que la somme des carrés des erreurs aura pour variation une quantité du second ordre par rapport à  $\delta x, \delta y, \delta z, \text{etc.}$  ; ce qui s'accorde avec la nature du *minimum*.

Si après avoir déterminé toutes les inconnues  $x, y, z, \text{etc.}$ , on substitue leurs valeurs dans les équations proposées, on connaîtra les diverses erreurs  $E, E', E'', \text{etc.}$ , auxquelles ce système donne lieu, et qui ne peuvent être réduites sans augmenter la somme de leurs carrés. Si parmi ces erreurs il s'en trouve que l'on juge trop grandes pour être admises, alors on rejettera les équations qui ont produit ces erreurs, comme venant d'expériences trop défectueuses, et on déterminera les inconnues par le moyen des équations restantes, qui alors donneront des erreurs beaucoup moindres. Et il est à observer qu'on ne sera pas obligé alors de recommencer tous les calculs ; car comme les équations du minimum se forment par l'addition des produits faits dans chacune des équations proposées, il suffira d'écarter de l'addition les produits donnés par les équations qui auront conduit à des erreurs trop considérables.

La règle par laquelle on prend le milieu entre les résultats de différentes observations, n'est qu'une conséquence très simple de notre méthode générale que nous appellerons *Méthode des moindres carrés*.

En effet, si l'expérience a donné diverses valeurs  $a', a'', a''', \text{etc.}$ , pour une certaine quantité  $x$ , la somme des carrés des erreurs sera  $(a' - x)^2 + (a'' - x)^2 + (a''' - x)^2 + \text{etc.}$ , et en égalant cette somme à un *minimum*, on a :

$$0 = (a' - x) + (a'' - x) + (a''' - x) + \text{etc.}$$

d'où résulte  $x = (a' + a'' + a''' + \text{etc.})/n$  ;  $n$  étant le nombre des observations.

Pareillement, si pour déterminer un point dans l'espace, on a trouvé, par une première expérience, les coordonnées  $a', b', c'$  ; par une seconde, les coordonnées  $a'', b'', c''$ , etc., ainsi de suite ; soient  $x, y, z$ , les véritables coordonnées de ce point : alors l'erreur de la première expérience sera la distance du point  $(a', b', c')$  au point  $(x, y, z)$  ; le carré de cette distance est :

$$(a' - x)^2 + (b' - y)^2 + (c' - z)^2$$

et la somme des carrés semblables étant égalée à un *minimum*, on en tire trois équations qui donnent  $x = (\int a)/n$  ;  $y = (\int b)/n$  ;  $z = (\int c)/n$ ,  $n$  étant le nombre de points donnés par l'expérience. Ces formules sont les mêmes par lesquelles on trouverait le centre de gravité commun de plusieurs masses égales, situées dans les points donnés ; d'où l'on voit que le centre de gravité d'un corps quelconque jouit de cette propriété générale.

Si on divise la masse d'un corps en molécules égales et assez petites pour être considérées comme des points, la somme des carrés des distances des molécules au centre de gravité sera un minimum.

On voit que la méthode des moindres carrés fait connaître, en quelque sorte, le centre autour duquel viennent se ranger tous les résultats fournis par l'expérience, de manière à s'en écarter le moins qu'il est possible. [...] »

On remarquera qu'il suffit de remplacer le signe somme  $\int$  par le signe  $\Sigma$  pour que le texte de Legendre soit conforme aux notations actuelles.

## 15.2. Application à la mesure des degrés du méridien

Legendre donne l'exemple d'application suivant :

« Supposant que le méridien est une ellipse dont les axes sont dans le rapport de 1 à  $1 + \alpha$ , si on désigne par D la longueur du 45° degré, et par S celle de l'arc compris entre les deux latitudes L et L', on aura par les formules connues, et en exprimant L' - L en degrés :

$$S = D(L' - L) - \frac{3}{2} \alpha D \cdot \frac{180}{\pi} \sin(L' - L) \cos(L' + L);$$

d'où résulte

$$L' - L = \frac{S}{D} + \frac{3}{2} \alpha \cdot \frac{180}{\pi} \sin(L' - L) \cos(L' + L).$$

Comme le 45° degré est d'environ 28 500 modules, égaux chacun à deux toises, on peut faire  $\frac{1}{D} = \frac{1 + \zeta}{28500}$ ,  $\zeta$  étant une fraction très petite, et on aura

$$L' - L = \frac{S}{28.500} + \frac{\zeta}{28.500} + \alpha \cdot \frac{270}{\pi} \sin(L' - L) \cos(L' + L), \text{ (a)}$$

équation qui pour chaque arc dont on connaît la longueur avec la latitude de ses extrémités, donnera une relation entre  $\alpha$  et  $\zeta$ .

« Voici maintenant les longueurs des différents arcs de la méridienne de France et les latitudes des parallèles qui les séparent, telles qu'elles résultent de l'opération exécutée par les célèbres astronomes Delambre et Méchain.

| Lieu<br>de l'observation. | Sa latitude.  | Arcs compris<br>exprimés en modules. | L' - L        | L' + L     |
|---------------------------|---------------|--------------------------------------|---------------|------------|
| Dunkerque .....           | 51° 2' 10" 50 | DP 62472.59                          | 2° 11' 20" 75 | 99° 53' 0" |
| Panthéon à Paris          | 48 50 49.75   | PE 76145.74                          | 2 40 7.25     | 95 1 32    |
| Evau.....                 | 45 10 42.50   | EC 84424.55                          | 2 57 48.10    | 89 23 37   |
| Carcassonne.....          | 43 12 54.40   | CM 52749.48                          | 1 51 9.60     | 84 34 39   |
| Montjoux .....            | 41 21 44 80   |                                      |               |            |

Nous avons donc quatre arcs dont les mesures étant substituées successivement dans l'équation (a), fourniront quatre équations entre  $\alpha$  et  $\zeta$ . Mais comme ces quatre équations ne peuvent pas être satisfaites toutes à la fois, nous supposerons qu'elles ont lieu, en attribuant une certaine erreur à la latitude de chaque lieu, et nous appellerons E', E'', etc. les corrections additives aux latitudes de Dunkerque, du Panthéon, etc. Ces erreurs n'entrent que dans le premier membre de chaque équation : elles sont trop petites pour affecter le terme multiplié par  $\alpha$  dans le second membre. Voici donc les équations qui résultent des quatre arcs mesurés dans l'opération de la méridienne,

$$E' - E'' = 0,002923 + \zeta(2,192) - \alpha(0,563)$$

$$E'' - E''' = 0,003100 + \zeta(2,672) - \alpha(0,351)$$

$$E''' - E^{IV} = 0,001096 + \zeta(2,962) + \alpha(0,047)$$

$$E^{IV} - E^V = 0,001808 + \zeta(1,851) + \alpha(0,263)$$

Comme il importe de considérer les erreurs séparément, on regardera comme une nouvelle inconnue l'erreur  $E'''$ , par exemple, et on aura les cinq équations :

$$E' = E''' + 0,006023 + \zeta(4,864) - \alpha(0,914)$$

$$E'' - E''' = 0,003100 + \zeta(2,672) - \alpha(0,351)$$

$$E''' = E'''$$

$$E^{IV} = E''' + 0,001096 - \zeta(2,962) - \alpha(0,047) \quad (b)$$

$$E^V = E''' + 0,002904 - \zeta(4,813) - \alpha(0,310)$$

Il faut maintenant faire en sorte que la somme des carrés de ces cinq erreurs soit un *minimum*, et d'abord cette condition exprimée par rapport à l'inconnue  $E'''$  dont tous les coefficients sont 1, donne par l'addition de toutes ces équations :

$$0 = 5E''' + 0,013123 - \zeta(0,239) - \alpha(1,622)$$

Substituant cette valeur dans les cinq équations (b), on aura

$$E' = 0,003398 + \zeta(4,912) - \alpha(0,590)$$

$$E'' = 0,000475 + \zeta(2,720) - \alpha(0,027)$$

$$E''' = -0,002625 + \zeta(0,048) + \alpha(0,324)$$

$$E^{IV} = -0,001529 + \zeta(2,914) + \alpha(0,277)$$

$$E^V = 0,000279 + \zeta(4,765) + \alpha(0,014)$$

Pour exprimer ensuite la condition du *minimum* par rapport à  $\zeta$ , il faut multiplier la première équation par 4,912 coefficient de  $\zeta$  ; la seconde par 2,720 ; la troisième par 0,048 ; la quatrième par - 2,914 ; la cinquième par - 4,765, et égaliser à zéro la somme de tous les produits. On opérera semblablement par rapport à  $\alpha$ , et on aura les deux équations :

$$0 = 0,020983 + \zeta(62,726) - \alpha(3,830)$$

$$0 = 0,003287 - \zeta(3,830) + \alpha(0,531)$$

d'où l'on tire  $\alpha = 0,00675$ , et  $\zeta = 0,0000778$ , donc

l'aplatissement =  $1/248$

et le 45° degré D =  $28\,500/(1 + \zeta) = 28\,497,78$  »

Legendre explique alors pourquoi ces valeurs diffèrent de celles connues par ailleurs.

### 15.3. Biographie de Legendre (1752-1833)

Adrien-Marie Legendre naît à Paris<sup>1</sup> le 18 septembre 1752. Sa famille, qui est très à l'aise, lui fait faire ses études au Collège Mazarin à Paris, où il est remarqué par l'abbé Marie, son professeur de mathématiques. Celui-ci insère, avec éloges, plusieurs articles de son élève dans son *Traité de mécanique* qu'il publie en 1774. Puis il le fait nommer, par l'entremise de d'Alembert, à l'École militaire de Paris où il va enseigner les mathématiques de 1775

1. Legendre est peut-être né à Toulouse et sa famille aurait rejoint ensuite Paris alors qu'Adrien-Marie était encore très jeune.

à 1780. Deux années après avoir quitté l'École militaire, Legendre obtient le prix mis au concours par l'Académie de Berlin sur la question de *déterminer la courbe décrite par les boulets et les bombes, en ayant égard à la résistance de l'air, et de donner les règles pour connaître les portées qui répondent à différentes vitesses initiales et à différents angles de projection*. Son mémoire intitulé *Recherches sur la trajectoire des projectiles dans les milieux résistants* est couronné en juin 1782. Newton a déjà traité le problème dans les *Principia*, mais Legendre améliore les calculs dans l'hypothèse d'une résistance de l'air proportionnelle au carré de la vitesse. En 1782, Lagrange dirige les mathématiques à l'Académie de Berlin, et le prix obtenu par Legendre attire son attention. En 1783, Legendre lit à l'Académie des sciences de Paris, un mémoire sur l'*Attraction des ellipsoïdes*. Pour déterminer l'attraction d'un tel ellipsoïde sur n'importe quel point extérieur il utilise une série de puissances. D'Alembert et Laplace ayant fait un excellent rapport sur ce travail et une place de membre adjoint mécanicien ayant été rendue vacante par la nomination de Laplace à la fonction d'associé, l'Académie des sciences désigne Legendre pour cette place de membre adjoint le 30 mars 1783.

En juillet 1784, Legendre lit à l'Académie, un *Mémoire sur la figure des planètes*<sup>1</sup> où il démontre que la figure elliptique peut seule convenir à l'équilibre d'une masse fluide, homogène, animée d'un mouvement de rotation et dont toutes les molécules s'attirent en raison inverse du carré de la distance. Il introduit dans cette étude ce que nous appelons les polynômes de Legendre<sup>2</sup>. Ses *Recherches d'analyse indéterminée*<sup>3</sup> de 1785 contiennent sa loi de réciprocité quadratique dont il donne une première démonstration partielle<sup>4</sup>. Le même traité contient aussi la théorie de la décomposition des nombres en trois carrés. En 1786, Legendre présente un *Mémoire sur la manière de distinguer les maxima des minima dans le calcul des variations*<sup>5</sup>. Les conditions dites de Legendre, établies dans ce papier, initient une abondante littérature. Les premiers rudiments de sa théorie des fonctions elliptiques sont présentés dans *Mémoires sur les intégrations par arcs d'ellipse et sur la comparaison de ces arcs*<sup>6</sup>. Par ailleurs, l'Académie l'a promu membre associé durant la même période (23 avril 1785).

En 1787, avec Cassini IV et Méchain, Legendre est chargé de certaines opérations géodésiques entreprises conjointement par l'Observatoire de Paris et celui de Greenwich. À cette occasion, il calcule non seulement tous les

1. Mémoires de l'Acad. des Sc., 1784, p. 370 ff.

2. Les polynômes de Legendre sont solutions d'une équation différentielle de la forme :  $(1 - x^2)y'' - 2xy' + n(n + 1)y = 0$

3. *Id.*, 1785, p. 465-560

4. La loi avait été énoncée sans preuve par Euler en 1783. Gauss la démontrera complètement en 1801.

5. *Id.*, 1786, p. 7-37 et 1787, p. 348-351.

6. *Id.*, 1786, p. 616, 644-673.

triangles situés en France, mais même ceux qui relient la côte d'Angleterre à Greenwich. Ce travail l'amène à Londres où il est nommé Fellow de la *Royal Society*. Il présente ses recherches à l'Académie des sciences en 1787 dans un *Mémoire sur les opérations trigonométriques dont les résultats dépendent de la figure de la terre*<sup>1</sup>. C'est dans ce mémoire qu'on voit apparaître, pour la première fois, la dénomination de *lignes géodésiques*, attribuée aux lignes de longueur minimum tracées sur une surface donnée. On y trouve aussi son théorème sur les triangles sphériques. En 1789 et 1790, dans son *Mémoire sur les intégrales doubles*<sup>2</sup>, Legendre complète son étude précédente de l'attraction des sphéroïdes en prenant en compte des sphéroïdes hétérogènes. En avril 1791 il est nommé, toujours avec Cassini et Méchain, responsable des opérations astronomiques nécessaires pour déterminer le mètre étalon, mais il demande à être relevé de cette mission un an plus tard.

En 1792, Legendre présente à l'Académie son *Mémoire sur les transcendentes elliptiques*. Mais les Académies sont supprimées en août 1793 et il doit faire imprimer lui-même son mémoire. Les temps deviennent durs pour lui car sa modeste fortune a fondu durant la Révolution. De plus, il épouse à cette époque une jeune fille de dix-neuf ans, Marguerite Couhin<sup>3</sup>. Durant 1794, il devient le Chef de la Commission Nationale exécutive de l'Instruction Publique et il publie ses *Éléments de géométrie* que Condorcet l'avait incité à écrire. Cet ouvrage aura un succès prodigieux bien que, par le plan et les démonstrations, il reproduise les défauts des traités d'Euclide et d'Archimède.

En 1795, l'Académie des sciences est réouverte sous le nom d'*Institut National des Sciences et des Arts*. Legendre obtient l'une des six places de la section de mathématiques (13 décembre 1795). Cet Institut est remanié par Napoléon en 1803, et Legendre fait alors partie de la section de Géométrie. Il devient vice-président de l'Académie des sciences en 1804-1805, puis président en 1805-1806. Legendre n'est pas nommé professeur à Polytechnique ni à l'École normale de l'an III, mais il succède à Laplace en 1799 comme examinateur de mathématiques pour l'entrée à l'école d'artillerie. Il gardera cette activité jusqu'en 1815, date à laquelle il démissionnera et se verra attribuer une pension de 3 000 francs qui lui sera retirée en 1824 à la suite de son refus de voter pour le candidat officiel lors d'une élection à un siège de l'Institut. Quand Lagrange meurt en 1813, Legendre lui succède au bureau des Longitudes, et il y restera sa vie durant.

1. *Id.*, 1787, p. 352 ff.

2. *Id.*, 1787, p. 454-486.

3. Plus tard il dira à Jacobi : « Je me suis marié après une révolution sanglante qui avait détruit ma petite fortune ; nous avons eu de grands problèmes et quelques moments très difficiles, mais mon épouse m'a aidé avec beaucoup de dévouement pour mettre peu à peu mes affaires en ordre et elle m'a donné la tranquillité nécessaire pour faire mon travail usuel et pour écrire de nouveaux travaux qui ont solidement augmenté ma réputation. »

Legendre publie en 1798 son *Essai sur la théorie des nombres* dans lequel il reprend les résultats des *Recherches d'analyse indéterminée* de 1785 et donne une nouvelle présentation de la loi de réciprocité quadratique.

En 1805 il publie *La nouvelle méthode pour la détermination des orbites des comètes*, ouvrage dans lequel on trouve en annexe la méthode des moindres carrés. Gauss, dans son traité *Theoria motus corporum coelestium* publié en 1809, indique qu'il utilisait la technique des moindres carrés depuis 1795. Mais la publication de Legendre étant la première chronologiquement, l'affirmation de Gauss irrita Legendre qui déclara, dit-on : « cette impudence excessive est incroyable chez un homme dont les mérites personnels sont suffisants pour qu'il n'ait pas besoin de s'attribuer les découvertes des autres. »

Legendre a produit un travail important sur les fonctions elliptiques dans les trois volumes de ses *Exercices du Calcul Intégral*<sup>1</sup>. En 1823, il publie ses *Recherches sur quelques objets d'analyse indéterminée et particulièrement sur le théorème de Fermat*<sup>2</sup> qui comportent la démonstration du grand théorème de Fermat, pour  $n = 5$ . Par ailleurs, dans son ouvrage intitulé *Traité des fonctions elliptiques et des intégrales eulériennes* (publié à Paris en 1825-1826), Legendre reprend ses travaux sur les fonctions elliptiques et traite les applications à la mécanique. Legendre tente de démontrer durant des années le cinquième postulat d'Euclide. Ses *Réflexions sur les différentes manières de démontrer la théorie des parallèles ou le théorème sur la somme des trois angles du triangle*<sup>3</sup> sont présentées à l'Académie en 1833.

Comme nous l'avons indiqué plus haut, en 1824, Legendre refuse de voter pour le candidat du gouvernement à l'élection de la présidence de l'Institut<sup>4</sup>, sa pension de 3 000 francs ne lui est plus versée et il vivra désormais dans la pauvreté. Legendre meurt à Paris le 9 janvier 1833, après une maladie douloureuse.

S.-D. Poisson, dans son *Discours prononcé aux funérailles de M. Legendre*<sup>5</sup>, a décrit le caractère de Legendre comme suit : « Notre Collègue a souvent exprimé le désir que, lorsque l'on parle de lui, ce soit uniquement à propos de ses travaux, qui furent, en fait, toute sa vie ». Il ajoute : « M. Legendre

- 
1. Dans le premier volume, publié en 1811, il présente les propriétés de base des intégrales elliptiques ainsi que celles des fonctions bêta et gamma. Dans le second volume, daté de 1817, Legendre donne d'autres résultats sur ces fonctions et applique ses résultats à la mécanique, à la rotation de la terre, à l'attraction des ellipsoïdes et à d'autres problèmes. Le dernier volume, publié en 1819, contient des tables de valeurs numériques d'intégrales elliptiques.
  2. *Mémoires de l'Académie des sciences* 6, 1823, p. 1-60.
  3. *Id.*, 12, 1833, p. 367-410.
  4. Ce qui fera dire à Abel, en octobre 1826, « Legendre est un homme extrêmement aimable, mais malheureusement aussi vieux que les pierres ».
  5. *Moniteur Universel* (20 Jan 1833), p. 162.

a eu cela de commun avec la plupart des géomètres qui l'ont précédé, que ses travaux n'ont fini qu'avec sa vie. Le dernier volume de nos mémoires renferme encore un mémoire de lui, sur une question difficile de la théorie des nombres ; et peu de temps avant la maladie qui la conduisit au tombeau, il se procura les observations les plus récentes des comètes à courtes périodes, dont il allait se servir pour appliquer et perfectionner ses méthodes. »

De son côté, dans son éloge de Legendre<sup>1</sup>, lu à l'Académie le 25 mars 1861, le président de cette institution, Léonce Élie de Beaumont, affirme : « Legendre, plus profond que populaire, a été notre Euler ; comme Euler et à son exemple, il a légué à l'avenir une foule de ces résultats analytiques que le génie seul sait obtenir et qui enrichissent pour toujours le domaine de l'esprit humain ».

## 15.4. Principales publications de Legendre

- *Essai sur la théorie des nombres*, Paris, Duprat, an VI, 1798.
- *Nouvelles méthodes pour la détermination des orbites des comètes*, Paris, Didot, 1805.
- *Traité des fonctions elliptiques et des intégrales eulériennes, avec des tables pour en faciliter le calcul numérique*, Paris, Huzard-Courcier, 1825-1828.

## 15.5. Références

- D. Poisson, *Discours prononcé aux funérailles de M. Legendre*, Moniteur universel, 20 Jan. 1833.
- L. Maurice, *Mémoire sur les travaux et écrits de M. Legendre*, Bibliothèque universelle des sciences, belles-lettres et arts, Science et arts 52 (1833), 45-82.
- E. de Beaumont, *Éloge historique d'Adrien-Marie Legendre*, Mémoires de l'Académie des sciences 32 (1864), 37-87.
- J. Itard, *Legendre*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 8, p. 135-143, 1970.

---

1. Mémoires de l'Académie des sciences, 32 (1864) xxxvii-xciv.





# Gauss

## 16.1. La loi normale et la méthode des moindres carrés de 1809

On sait que la théorie des moindres carrés relative aux erreurs de mesure a connu cinq étapes importantes : Le mémoire de Legendre de 1805, celui de 1809 de Gauss, ceux de 1810 et 1811 de Laplace et enfin la seconde théorie de Gauss de 1823.

Laplace dans son papier de 1786 et dans son premier mémoire de 1810 avait recherché une loi de distribution des erreurs menant à une estimation optimale selon la moyenne arithmétique des mesures, mais n'y parvint pas. Il n'énonça le théorème limite central que dans le supplément à son papier de 1810, après parution de la méthode de Gauss.

Gauss aborda le problème de la loi normale de la façon suivante : il considéra la méthode d'estimation et l'estimateur donnés et chercha s'il existait une loi de distribution des erreurs compatible avec ses hypothèses.

Gauss commence par établir le principe de probabilité *a posteriori* basé sur un modèle d'équiprobabilité, à l'aide d'un exemple faisant intervenir deux hypothèses s'excluant l'une l'autre et excluant également toutes les autres hypothèses différant des deux premières. C'est un principe identique à celui adopté par Laplace dans son *Mémoire de la Probabilité des causes par les événements*<sup>1</sup>. Laplace y énonce comme un principe le théorème de Bayes de la probabilité des causes prises *a priori* égales.

Pour le reste, Gauss, au lieu de raisonner sur une courbe d'erreurs et de là sur une méthode de combinaisons des observations, constate que la courbe doit avoir un maximum pour une erreur nulle, qu'elle doit être symétrique et être nulle pour des erreurs importantes. Puis il adopte un axiome et déclare : « On a coutume de regarder comme un axiome l'hypothèse que si une quantité a été obtenue par plusieurs observations immédiates, faites avec le même soin dans des circonstances semblables, la moyenne arithmétique des valeurs observées sera la valeur la plus probable de cette quantité, sinon

1. *Mémoires de l'Académie des sciences de Paris (Savants étrangers)*, tome VI, p. 621-656, 1774.

THEORIA  
MOTVS CORPORVM  
COELESTIVM

IN

SECTIONIBVS CONICIS SOLEM AMBIENTIVM

AUCTORE

CAROLO FRIDERICO GAUSS

---

HAMBVRGI SVMTIBVS FRID. PERTHES ET I. H. BESSER  
1809.

Venditur

PARISIIS ap. Trouttel & Würtz:  
STOCKHOLMIAE ap. A. Wiborg.  
MADRITI ap. Sancha.

LONDINI ap. R. H. Evans.  
PETROPOLI ap. Klostermann.  
FLORENTIAE ap. Molini, Landi & C.  
AMSTELODAMI in libreria: Kunst- und Industrie-Comptoir, dicta.

en toute rigueur, du moins avec une grande approximation, de telle sorte que le plus sûr soit toujours de s'y arrêter. » Avec cette hypothèse de valeur la plus probable égale à la moyenne arithmétique, il démontre que la seule courbe adéquate est la loi de distribution normale<sup>1</sup>.

On peut, en utilisant le concept de variable aléatoire, exposer comme suit le résultat de Gauss de 1809 :

On considère que les erreurs sur une quantité  $X$  inconnue, pour laquelle on dispose de plusieurs mesures  $x_1, x_2, \dots, x_n$ , sont :

$e_1 = x_1 - X, \dots, e_n = x_n - X$  et correspondent à des réalisations de  $n$  variables aléatoires indépendantes  $E_1, \dots, E_n$  de même loi continue de densité  $f$ , dépendant de la valeur inconnue de  $X$ .

Pour  $X$  donné, la probabilité d'effectuer des erreurs comprises entre  $e_1 + de_1, \dots, e_n + de_n$  est alors (compte tenu de l'indépendance)  $f(e_1) \cdot f(e_2) \dots f(e_n) \cdot de_1 \dots de_n$ .

On peut alors, selon la méthode de *Bayes*, les mesures  $x_1, \dots, x_n$  étant connues, chercher quelle est la valeur de  $X$  la plus vraisemblable.

Cette valeur de  $X$  rendra maximale la probabilité d'observation des mesures observées  $x_1, \dots, x_n$  et donc des erreurs  $e_1, \dots, e_n$ . Il s'agit de rechercher  $X$ , donc  $f$ , de sorte que  $f(x_1 - X) \dots f(x_n - X)$  soit maximum.

Le produit  $f(x_1 - X) \dots f(x_n - X)$  est maximum lorsque la somme  $\log f(x_1 - X) + \dots + \log f(x_n - X)$  est maximale.

En dérivant par rapport à  $X$  on obtient la condition

$$\frac{d \log f(x_1 - X)}{dX} + \dots + \frac{d \log f(x_n - X)}{dX} = 0$$

On sait que la moyenne arithmétique  $\bar{X} = \frac{x_1 + \dots + x_n}{n}$  correspond à la valeur recherchée de  $X$  :

(Comme déjà indiqué, Gauss écrit : « On a coutume de regarder comme un axiome l'hypothèse que si une quantité a été obtenue par plusieurs observations immédiates, faites avec le même soin dans des circonstances semblables, la moyenne arithmétique des valeurs observées sera la valeur la plus probable de cette quantité, sinon en toute rigueur, du moins avec une grande approximation, de telle sorte que le plus sûr soit toujours de s'y arrêter. »)

Cette moyenne vérifie l'équation en  $X$

$$\sum_{i=1}^n (x_i - X) = 0$$

1. Cf., par exemple, Hald A., *A history of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 351-356.

Gauss en déduit que, pour  $i$  allant de 1 à  $n$ , on a

$$\frac{d \log f(x_i - X)}{dX} = k(x_i - X).$$

On a, en intégrant,

$$\log f(x_i - X) = -k \frac{(x_i - X)^2}{2} + \text{cte soit :}$$

$$f(x_i - X) = C.e^{-k \frac{(x_i - X)^2}{2}},$$

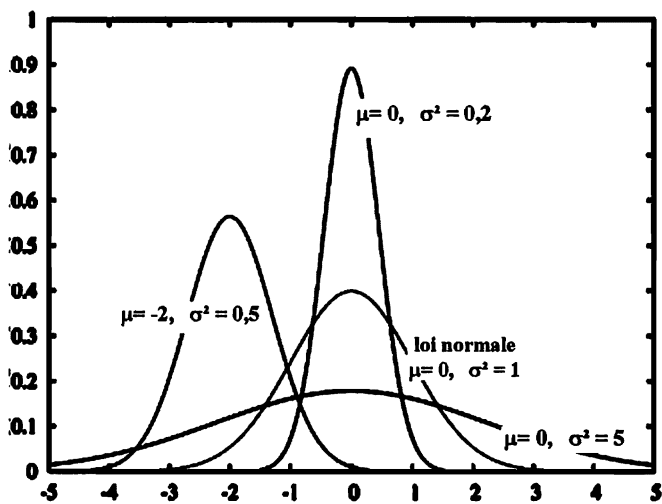
c'est-à-dire la densité de la loi normale.

On peut vérifier que si  $f(x_i - X) = C.e^{-k \frac{(x_i - X)^2}{2}}$ ,

$$\text{le produit } f(x_1 - X) \dots f(x_n - X) = C.e^{-k \frac{\sum_{i=1}^n (x_i - X)^2}{2}}$$

est maximum lorsque la somme des carrés des écarts des mesures  $\sum_{i=1}^n (x_i - X)^2$  est minimale.

En résumé, Gauss, dans son traitement de la théorie des erreurs de 1809, introduit en fait le principe du maximum de vraisemblance (puisqu'il cherche la valeur de la quantité inconnue  $X$  qui rendra maximale la probabilité d'observation des mesures observées  $x_1, \dots, x_n$  et donc des erreurs  $e_1, \dots, e_n$ ). Après ajout du postulat de la moyenne arithmétique, il obtient à la fois la loi de distribution normale et le principe des moindres carrés. La justification de ce principe des moindres carrés est donc la maximalisation de la distribution *a posteriori* des erreurs.



Fonction densité de probabilité  $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ ,  $x \in \mathbb{R}$ .

$$\frac{d\nu}{dr} \varphi' \nu + \frac{d\nu'}{dr} \varphi' \nu' + \frac{d\nu''}{dr} \varphi' \nu'' + \text{etc.} = 0$$

$$\frac{d\nu}{ds} \varphi' \nu + \frac{d\nu'}{ds} \varphi' \nu' + \frac{d\nu''}{ds} \varphi' \nu'' + \text{etc.} = 0$$

Hinc itaque per eliminationem problematis solutio plene determinata derivari poterit, quamprimum functionis  $\varphi'$  indoles innotuit. Quae quoniam a priori definiri nequit, rem ab altera parte aggredientes inquiremus, cuinam functioni, taceat quasi pro basi acceptae, proprie innixum sit principium trium, cuius praesentia generaliter agnoscitur. Axiomati scilicet loco haberi solet hypothesis, si quae quantitas per plures observationes immediatas, sub aequalibus circumstantiis aequalique cura institutas, determinata fuerit, medium arithmeticum inter omnes valores observatos exhibere valorem maxime probabilem, si non absolute rigore, tamen proxime saltem, ita ut semper tutissimum sit illi inhaerere. Statuendo itaque  $V = V' = V'' \text{ etc.} = p$ , generaliter esse debet  $\varphi'(M - p) + \varphi'(M' - p) + \varphi'(M'' - p) + \text{etc.} = 0$ , si pro  $p$  substituatur valor  $\frac{1}{\mu} (M + M' + M'' + \text{etc.})$ , quencunque integrum positivum exprimat  $\mu$ . Supponendo itaque  $M' = M'' = \text{etc.} = M - \mu N$ , erit generaliter, i. e. pro quovis valore integro positivo ipsius  $\mu$ ,  $\varphi'(\mu - 1)N = (1 - \mu)\varphi'(-N)$ , unde facile colligitur, generaliter esse debere  $\frac{\varphi' \Delta}{\Delta}$  quantitatem constantem, quam per  $k$  designabimus. Hinc fit  $\log \varphi \Delta = \frac{1}{2} k \Delta \Delta + \text{Const.}$ , siue designando basin logarithmorum hyperbolicorum per  $e$ , supponendoque  $\text{Const.} = \log \kappa$ ,

$$\Delta \varphi = \kappa e^{\frac{1}{2} k \Delta \Delta}$$

Porro facile perspicitur,  $k$  necessario negativam esse debere, quo  $\Delta$  reuera fieri possit maximum, quamobrem statuemus  $\frac{1}{2} k = -h$ ; et quum per theorema elegans primo ab ill. Laplace inuentum, integrale  $\int e^{-h \Delta \Delta} d\Delta$ , a  $\Delta = -\infty$  usque ad  $\Delta = +\infty$ , fiat  $= \frac{\sqrt{\pi}}{h}$ , (denotando per  $\pi$  semicircumferentiam circuli cuius radius 1), functio nostra fiet

$$\varphi \Delta = \frac{h}{\sqrt{\pi}} e^{-h \Delta \Delta}$$

178.

Functio modo eruta omni quidem rigore errorum probabilitates exprimere certo non potest: quum enim errores possibiles semper limitibus certis coërceantur,

Voici la traduction par J. Bertrand<sup>1</sup> du texte de Gauss paru dans l'ouvrage intitulé *Theoria motus corporum coelestium*<sup>2</sup> de 1809 :

« Soient  $V, V', V'',$  etc.  $\mu$  fonctions des  $\nu$  inconnues  $p, q, r, s,$  etc., et supposons que des observations directes aient donné, pour ces fonctions, les valeurs

$$V = M, V' = M', V'' = M'', \dots$$

En général, le calcul de ces inconnues constituera un problème indéterminé, déterminé ou plus que déterminé, suivant que l'on aura

$$\mu < \nu, \mu = \nu, \text{ ou } \mu > \nu$$

Nous ne nous occuperons ici que du dernier cas, dans lequel, évidemment, il ne serait possible d'obtenir une représentation exacte de toutes les observations, que si ces observations n'étaient affectées d'aucune erreur. Mais comme cela n'a jamais lieu dans la nature, on devra regarder comme possible tout système de valeurs des inconnues  $p, q, r, s,$  etc., desquelles résultent, pour les fonctions  $V = M, V' = M', V'' = M'',$  des valeurs qui ne surpassent pas les limites des erreurs que l'on peut commettre dans les observations, mais on ne doit pas regarder tous ces systèmes possibles comme jouissant du même degré de probabilité.

Supposons d'abord, dans toutes les observations, un état de choses tel qu'il n'y ait pas lieu de regarder l'une d'elle comme plus exacte qu'une autre, c'est-à-dire, que l'on doive regarder des erreurs égales dans chacune d'elles comme également probables. La probabilité qu'une erreur  $\Delta$  soit commise dans l'une des observations sera une fonction de  $\Delta$ , que nous nommerons  $\varphi(\Delta)$ . Quoique cette fonction ne puisse être assignée d'une manière précise, on peut du moins affirmer qu'elle doit devenir maximum pour  $\Delta = 0$ , avoir dans la plupart des cas la même valeur pour des valeurs de  $\Delta$  égales et de signes contraires, et, enfin, s'évanouir quand on donne à  $\Delta$  une valeur égale ou supérieure à l'erreur maximum ;  $\varphi(\Delta)$  doit donc, à proprement parler, être rapportée à la classe des fonctions discontinues, et, si nous nous permettons, pour la facilité du calcul, d'y substituer une fonction analytique, il faudra que cette dernière soit choisie de telle sorte qu'elle tende rapidement vers 0 à partir de deux valeurs de  $\Delta$ , l'une supérieure, l'autre inférieure à 0, et qu'en dehors de ces deux limites on puisse la considérer comme nulle. Or la probabilité que l'erreur soit comprise entre  $\Delta$  et une quantité  $\Delta + d\Delta$  qui en diffère infiniment peu, sera exprimée par  $\varphi(\Delta).d\Delta$ , et, par suite, la probabilité que l'erreur est comprise entre D et D', par  $\int_D^{D'} \varphi(\Delta)d\Delta$ .

Cette intégrale, prise depuis la plus grande valeur négative de  $\Delta$  jusqu'à sa plus grande valeur positive, ou plus généralement depuis  $\Delta = -\infty$  jusqu'à  $\Delta = \infty$ , devra nécessairement être égale à 1. On aura donc

$$\int_{-\infty}^{\infty} \varphi(\Delta)d\Delta = 1.$$

Supposons donc qu'on ait un système déterminé de valeurs des quantités  $p, q, r, s,$  etc. : la probabilité que l'observation donnera pour V la valeur M, sera exprimée par  $\varphi(M - V)$ , après qu'on aura substitué dans V les valeurs de  $p, q, r, s,$  etc., de même

1. *Méthode des moindres carrés* par C. F. Gauss, traduite par J. Bertrand, Paris, Mallet-Bachelier, 1855.
2. *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*, Auctore Carolo Friderico Gauss, Hamburgi sumptibus Perthes et Besser, 1809, page 210 et sqq.

$\varphi(M' - V')$ ,  $\varphi(M'' - V'')$ , etc., exprimeront les probabilités pour que les observations donnent aux fonctions  $V'$ ,  $V''$ , etc., les valeurs  $M'$ ,  $M''$ , etc. C'est pourquoi, tant qu'on pourra considérer toutes les observations comme des événements indépendants les uns des autres, le produit

$$\varphi(M - V) \cdot \varphi(M' - V') \cdot \varphi(M'' - V'') \dots = \Omega$$

exprimera la probabilité que toutes ces valeurs résulteront en même temps des observations.

De même qu'en se donnant des valeurs quelconques des inconnues, il en résulte, avant toute observation, une probabilité déterminée pour un système de valeurs des fonctions  $V$ ,  $V'$ ,  $V''$  etc., de même, après que l'observation aura donné pour ces fonctions des valeurs déterminées, il en résultera pour chaque système de valeurs des inconnues qui en découleront, une probabilité déterminée : car il est clair qu'on devra considérer comme les plus probables les systèmes qui donnent à l'événement observé la plus grande probabilité. L'appréciation de cette probabilité peut s'obtenir par le théorème suivant :

*Si, en adoptant une certaine hypothèse H, la probabilité d'un événement déterminé E est h, mais qu'en adoptant une autre hypothèse H', exclusive de la première et ayant a priori la même probabilité, la probabilité du même événement soit h' : je dis que lorsque l'événement E aura eu lieu, la probabilité que H soit la vraie hypothèse sera à la probabilité que H' soit la vraie hypothèse comme h est à h'.*

Pour le démontrer et afin de distinguer toutes les circonstances d'où peut dépendre, soit que l'hypothèse H ou H', ou toute autre ait lieu, l'arrivée d'un événement E ou d'un autre événement, formons un système des cas différents qui peuvent se présenter et que nous regarderons comme également probables *a priori* (c'est-à-dire tant qu'il y a doute si c'est l'événement E ou un autre qui aura lieu). Ces cas peuvent être ainsi distribués :

| nombre de cas | Hypothèse propre à ces cas | Événement qui doit en résulter |
|---------------|----------------------------|--------------------------------|
| $m$           | H                          | E                              |
| $n$           | H                          | Différent de E                 |
| $m'$          | H'                         | E                              |
| $n'$          | H'                         | Différent de E                 |
| $m''$         | Différente de H et de H'   | E                              |
| $n''$         | Différente de H et de H'   | Différent de E                 |

On aura d'après cela

$$h = \frac{m}{m+n}, \quad h' = \frac{m'}{m'+n'}$$

Or, avant l'arrivée de l'événement, la probabilité de l'hypothèse H était

$$\frac{m+n}{m+n+m'+n'+m''+n''}$$



Après l'événement qui exclut  $n + n' + n''$  cas, parmi ceux qui sont possibles, cette probabilité sera

$$\frac{m}{m + m' + m''} .$$

De même les probabilités de l'hypothèse H' avant et après l'événement sont respectivement

$$\frac{m' + n'}{m + n + m' + n' + m'' + n''} \text{ et } \frac{m'}{m + m' + m''} ;$$

mais, comme on a supposé que les hypothèses H et H' avaient avant l'événement la même probabilité, on aura

$$m + n = m' + n'$$

d'où résulte immédiatement la vérité du théorème.

Si maintenant on suppose qu'on n'a, pour déterminer les inconnues, que les observations

$$V = M, V' = M', V'' = M'', \dots$$

et que tous les systèmes de valeurs des inconnues étaient également probables avant ces observations, il est visible que la probabilité d'un certain système, après ces observations, sera proportionnelle à  $\Omega$ . C'est-à-dire que  $\lambda \Omega dp dq dr \dots$  exprimera la probabilité que les valeurs des inconnues soient respectivement comprises dans les limites infiniment voisines  $p$  et  $p + dp$ ,  $q$  et  $q + dq$ ,  $r$  et  $r + dr$ , ...,  $\lambda$  représentant une quantité indépendante de  $p, q, r, s$ , etc. ; et l'on aura évidemment

$$\frac{1}{\lambda} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \Omega dp dq dr \dots .$$

De là résulte naturellement que le système le plus probable des valeurs de  $p, q, r$ , etc., correspondra au maximum de  $\Omega$ , et tirera des  $\nu$  équations

$$\frac{d\Omega}{dp} = 0, \frac{d\Omega}{dq} = 0, \frac{d\Omega}{dr} = 0, \dots ;$$

si l'on pose

$$V - M = v, V' - M' = v', V'' - M'' = v'', \dots \text{ et } \frac{d\varphi(\Delta)}{\varphi(\Delta).d\Delta} = \varphi'(\Delta).$$

ces équations prendront la forme suivante

$$\frac{dv}{dp} \varphi'(v) + \frac{dv'}{dp} \varphi'(v') + \frac{dv''}{dp} \varphi'(v'') + \dots = 0$$

$$\frac{dv}{dq} \varphi'(v) + \frac{dv'}{dq} \varphi'(v') + \frac{dv''}{dq} \varphi'(v'') + \dots = 0$$

.....

De là résulte qu'on pourra obtenir par l'élimination une solution pleinement déterminée du problème, dès que la nature de la fonction  $\varphi'$  sera connue. Mais comme cette fonction ne peut être définie à priori, abordons la question à un autre point de vue et cherchons une fonction acceptée tacitement comme base, en vertu d'un principe simple et généralement admis. Or, on a coutume de regarder comme un axiome l'hypothèse que si une quantité a été obtenue par plusieurs observations

immédiates, faites avec le même soin dans des circonstances semblables, la moyenne arithmétique des valeurs observées sera la valeur la plus probable de cette quantité, sinon en toute rigueur, du moins avec une grande approximation, de telle sorte que le plus sûr soit toujours de s'y arrêter. Si l'on pose

$$V = V' = V'' \dots = p$$

et

$$p = \frac{M + M' + M'' + \dots}{\mu}$$

on devra avoir en général

$$\varphi'(M - p) + \varphi'(M' - p) + \varphi'(M'' - p) + \dots = 0$$

pour toute valeur entière et positive de  $\mu$ . Faisant ensuite

$$M' = M'' \dots = \text{etc.} = M - \mu N,$$

on aura généralement

$$\varphi'[(\mu - 1)N] = (1 - \mu) \varphi'(-N),$$

d'où l'on tire facilement que  $\frac{\varphi'(\Delta)}{\Delta}$  doit être en général une constante  $k$ . On aura donc

$$\log \varphi(\Delta) = \frac{1}{2} k \Delta^2 + \text{cte} = \frac{1}{2} k \Delta^2 + \log x,$$

d'où

$$\varphi(\Delta) = x e^{\frac{1}{2} k \Delta^2},$$

Or, on voit facilement que la constante  $k$  doit être négative, pour que  $\Omega$  puisse devenir maximum. Posons donc

$$\frac{1}{2} k = -b^2,$$

et comme, d'après un élégant théorème de Laplace, on a

$$\int_{-\infty}^{\infty} e^{-b^2 \Delta^2} d\Delta = \frac{\sqrt{\pi}}{b},$$

notre fonction deviendra

$$\varphi(\Delta) = \frac{h}{\sqrt{\pi}} e^{-b^2 \Delta^2}.$$

La fonction que nous venons de trouver ne peut pas exprimer, en toute rigueur, la probabilité des erreurs, puisque les erreurs possibles étant toujours renfermées entre certaines limites, la probabilité d'erreurs plus grandes devrait être toujours nulle, tandis que notre fonction a toujours une valeur finie. Cependant ce défaut, que présenterait également toute autre fonction analytique, n'a aucune importance dans les applications, parce que la valeur de notre fonction décroît si rapidement, pour peu que  $h\Delta$  ait une valeur considérable, qu'on peut, en toute sûreté, la regarder comme équivalente à zéro. D'ailleurs, la nature de la question ne permettra jamais d'assigner les limites des erreurs avec une rigueur absolue.

Au reste, la constante  $h$  peut être regardée comme servant de mesure à la précision des observations. Si en effet la probabilité de l'erreur  $\Delta$  dans un système d'observations est exprimée par

$$\frac{h}{\sqrt{\pi}} e^{-h^2 \Delta^2},$$

et dans un autre système d'observations plus ou moins exactes que les premières par

$$\frac{h'}{\sqrt{\pi}} e^{-h'^2 \Delta^2},$$

la probabilité que dans une observation du premier système l'erreur soit comprise entre les limites  $-\delta$  et  $+\delta$  sera exprimée par

$$\int_{-\delta}^{\delta} \frac{h}{\sqrt{\pi}} e^{-h^2 \Delta^2} d\Delta,$$

et de même la probabilité que l'erreur d'une observation du second système soit comprise entre les limites  $-\delta'$  et  $+\delta'$  sera exprimée par

$$\int_{-\delta'}^{\delta'} \frac{h'}{\sqrt{\pi}} e^{-h'^2 \Delta^2} d\Delta,$$

or ces intégrales sont manifestement égales lorsqu'on a

$$h\delta = h'\delta'.$$

Si par exemple on a  $h' = 2h$ , une erreur double dans le premier système sera commise aussi facilement qu'une erreur simple dans le second, de sorte que les dernières observations, pour nous servir d'une expression consacrée par l'usage, jouissent d'un degré de précision deux fois plus grand. »

Après avoir démontré la loi normale que nous appelons aujourd'hui de Laplace-Gauss, Gauss en déduit la méthode des moindres carrés. Il écrit :

« Voici maintenant quelques conséquences de cette loi. Il est clair qu'il faut, pour que le produit

$$\Omega = h^\mu \pi^{-\frac{\mu}{2}} e^{-h^2(\nu^2 + \nu'^2 + \nu''^2 + \dots)}$$

devienne maximum, que la somme

$$\nu^2 + \nu'^2 + \nu''^2 + \dots$$

devienne minimum. Donc le système de valeurs des inconnues  $p, q, r, s$ , etc., le plus probable correspond au cas où les carrés des différences entre les valeurs observées et les valeurs calculées des fonctions  $V, V', V''$ , etc., donnent la somme la plus petite possible, pourvu que toutes les observations soient également présumées précises.

Ce principe [...] doit être regardé comme un axiome, au même titre que le principe qui nous fait adopter la moyenne arithmétique des valeurs observées d'une même quantité comme la valeur la plus probable de cette quantité.

Le principe s'étend sans peine au cas d'observations d'une précision *inéga*le. Car si les précisions des observations par lesquelles on a trouvé

$$V = M, V' = M', V'' = M'', \dots,$$

# MÉTHODE

DES

# MOINDRES CARRÉS.

MÉMOIRES SUR LA COMBINAISON DES OBSERVATIONS,

PAR CH.-FR. GAUSS.

TRADUITS EN FRANÇAIS ET PUBLIÉS AVEC L'AUTORISATION DE L'ACADÉMIE,

PAR J. BERTRAND.



PARIS,

MALLET - BACHELIER, IMPRIMEUR - LIBRAIRE

DE L'ÉCOLE POLYTECHNIQUE, DU BUREAU DES LONGITUDES,

Quai des Augustins, n° 55.

1855

( L'Éditeur de cet ouvrage se réserve le droit de traduction. )

sont représentées respectivement par  $h, h', h'',$  etc., c'est-à-dire si l'on suppose que des erreurs réciproquement proportionnelles à ces quantités puissent être commises avec la même facilité, il est clair que cela revient au même que si, par des observations d'une égale précision (représentée par 1), les valeurs des fonctions  $hV, h'V', h''V''$  etc., avaient été trouvées égales à  $hM, h'M', h''M'',$  etc. ; c'est pourquoi le système le plus probable de valeurs des quantités  $p, q, r, s,$  etc., sera celui où la somme

$$h^2 v^2 + h'^2 v'^2 + h''^2 v''^2 + \dots,$$

c'est-à-dire la somme des carrés des différences entre les valeurs observées et calculées, multipliées respectivement par les carrés des nombres qui expriment le degré de précision, devient un minimum. »

Gauss poursuit en faisant remarquer que les calculs numériques sont simples quand les fonctions  $V, V', V'',$  etc., sont linéaires.

## 16.2. Les remarques de Joseph Bertrand

À ce stade, nous pensons utile de citer quelques remarques formulées par J. Bertrand<sup>1</sup> :

« L'histoire de la méthode des moindres carrés est singulière. Legendre, qui, le premier, l'a proposée, en 1805, aux calculateurs, ne la rattachait à aucune idée théorique. Gauss, qui depuis longtemps en faisait usage, a publié, en 1809, une démonstration acceptée sans contestation et reproduite quelques années après par Laplace. Gauss, cependant, avait fait de prudentes réserves, et ces réserves, dans son Mémoire de 1821 sur la combinaison des observations, se transforment en un abandon complet de la démonstration proposée en 1809. Les règles, qui subsistent, sont tout autrement démontrées. Le progrès paraît considérable. La démonstration de 1809 supposait, pour la probabilité des erreurs, une loi très particulière, dans laquelle ne figure qu'une seule constante ; celle de 1821 est affranchie de toute hypothèse.

Comment déduire rigoureusement d'une loi qui reste inconnue une règle certaine et précise ?

La simplicité du résultat doit redoubler l'étonnement.

Si des grandeurs  $l_1, l_2, \dots, l_n$ , ont été observées, la probabilité d'une erreur comprise entre  $z$  et  $z + dz$ , pour l'une d'elles, étant  $\varphi(z)dz$ , la probabilité du concours des erreurs  $z_1, z_2, \dots, z_n$  est proportionnelle à

$$\varphi(z_1)\varphi(z_2)\dots\varphi(z_n) \quad (1)$$

le produit doit être rendu maximum.

La fonction  $\varphi(z)$  restant inconnue, le problème paraît insoluble. L'hypothèse la plus plausible est, d'après le résultat du calcul, celle qui rend minima la somme

$$z_1^2 + z_2^2 + \dots z_n^2 \quad (2)$$

1. J. Bertrand, *Sur la méthode des moindres carrés*, CRAS, Paris, tome CVI, n° 16, 1888, p. 1115-1117.

Si le produit (1) est maximum quand la somme (2) est minima, il doit rester constant en même temps que celle-ci et la fonction  $\varphi(z)$  doit satisfaire à une équation de la forme

$$\varphi(x_1)\varphi(x_2)\dots\varphi(x_n) = F(x_1^2 + x_2^2 + \dots x_n^2)$$

On déduit de cette équation bien connue

$$\varphi(z) = G.e^{-k^2 z^2}.$$

C'est la loi proposée en 1809 ; elle seule peut justifier la règle qu'on en a déduite. [...]

L'explication est simple ; on a fait, sans rien dire, l'hypothèse qui précise la forme de la fonction ; elle consiste à supposer les erreurs infiniment petites. L'identité de tous les problèmes n'est pas, dès lors, plus étrange que celle de la loi des courbures, la même sur toutes les surfaces pour toutes les courbes issues d'un même point.

Toutes les équations, dans la théorie de Gauss, sont réduites au premier degré.

Si l'on nomme  $\varphi(z)dz$  la probabilité pour qu'une erreur d'observation soit comprise entre  $z$  et  $z + dz$ , on peut,  $z$  étant très petit, remplacer  $\varphi(z)$  par son développement

$$\varphi(0) + z\varphi'(0) + \frac{z^2}{2}\varphi''(0) + \frac{z^3}{6}\varphi'''(0) + \dots$$

Les erreurs constantes étant écartées, c'est une des hypothèses de Gauss, on a

$$\varphi(z) = \varphi(-z)$$

et, par conséquent,

$$\varphi'(0) = 0 \text{ et } \varphi'''(0) = 0.$$

On peut, puisqu'on est convenu de négliger  $z^2$  par rapport à  $z$ , négliger  $z^4$  par rapport à  $z^2$  et réduire  $\varphi(z)$  à la forme

$$\varphi(z) = \varphi(0) + \frac{1}{2}\varphi''(0)z^2 = a + bz^2$$

Mais on a, dans les mêmes conditions,

$$Ge^{-k^2 z^2} = G - Gk^2 z^2.$$

Si donc on pose

$$G = a, k^2 = -\frac{b}{a},$$

on aura

$$\varphi(z) = G.e^{-k^2 z^2}.$$

La fonction  $\varphi(z)$ , quelle que soit sa forme véritable, est réductible, dans les conditions où l'on veut résoudre le problème, à celle que Gauss avait proposée d'abord.

L'introduction d'une loi de probabilité indéterminée n'est pas le vrai progrès accompli dans le Mémoire de 1821. La condition à remplir y a été remplacée par une autre de beaucoup préférable. Au lieu de rendre maxima la probabilité d'une erreur nulle,

on cherche, dans la théorie définitive, à diminuer, pour chaque inconnue, la valeur probable du carré de l'erreur commise. »

### 16.3. La fonction d'erreur

En 1816, Gauss publia un *mémoire sur la détermination de la précision des observations*<sup>1</sup>.

La fonction erreur y est introduite. C'est une intégrale numérique. On la trouve

tabulée dans tous les manuels et elle est définie par  $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\varsigma^2} d\varsigma$ .

La probabilité pour qu'une variable normale centrée réduite<sup>2</sup> prenne une valeur dans l'intervalle  $[-z, z]$  est  $\operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)$ .

On appelle  $\Phi$  la fonction de répartition de la loi normale centrée réduite

$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$ . Elle est définie, pour tout réel  $x$ , par :

$$\Phi(x) = \int_{-\infty}^x \varphi(t) dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

On peut donner une approximation de la fonction de répartition en effectuant un développement en série de Taylor. Voici une approximation (à l'ordre 5) autour de 0, correcte pour  $|x| < 2$  :

$$\Phi(x) \approx 0,5 + 0,3989423 \left[ x - \frac{x^3}{6} + \frac{x^5}{40} \right].$$

### 16.4. La théorie de la combinaison des observations (*Theoria combinationis observationum erroribus minimis obnoxiae*)

En 1821 et 1823, Gauss partit d'une hypothèse différente de celle choisie en 1809. Il postula une variance minimale et en déduisit le même principe des moindres carrés dont il détailla, cette fois les techniques de mise en œuvre<sup>3</sup>.

1. *Zeitschrift für Astronomie und Verwandte Wissenschaften*, tome 1, page 185, 1816.

2. Une variable aléatoire centrée réduite a une espérance mathématique  $E(X)$  nulle et un écart type égal à 1. Par rapport à une variable aléatoire quelconque  $X$ , la variable centrée réduite est une variable translatée et homothétique  $\frac{X-E(X)}{\sigma(X)}$ .

3. *Theoria combinationis observationum erroribus minimis obnoxiae*. Comment. (Göttingen) V, 1819-1822, 33-62, 63-90 ; *Supplementum Theoriae combinationis observationum erroribus minime obnoxiae*, Comment. (Göttingen) VI, 1823-1827, 57-98 ; VII, 89-148.

Voici comment Gauss introduit la variance dans ce mémoire<sup>1</sup> :

« Désignons par la notation  $\varphi(x)$  la facilité relative d'une erreur  $x$  : on doit entendre par là, à cause de la continuité des erreurs, que  $\varphi(x) dx$  est la probabilité que l'erreur soit comprise entre les limites  $x$  et  $x + dx$ . Il n'est pas possible, en général, d'assigner la forme de la fonction  $\varphi(x)$ , et l'on peut même affirmer que cette fonction ne sera jamais connue dans la pratique. On peut néanmoins établir plusieurs caractères généraux qu'elle doit nécessairement présenter :  $\varphi(x)$  est évidemment une fonction discontinue ; elle s'annule pour toutes les valeurs de  $x$  non comprises entre les erreurs extrêmes. Pour toute valeur comprise entre ces limites, la fonction est positive [...] ; dans la plupart des cas, les erreurs égales et de signes contraires seront également probables, et l'on aura :

$$\varphi(x) = \varphi(-x)$$

Enfin, comme les petites erreurs sont plus facilement commises que les grandes,  $\varphi(x)$  sera en général maximum pour  $x = 0$  et diminuera sans cesse lorsque  $x$  croîtra.

L'intégrale

$$\int_a^b \varphi(x) dx$$

exprime la probabilité pour que l'erreur, encore inconnue, tombe entre les limites  $a$  et  $b$ . On en conclut que la valeur de cette intégrale prise entre les limites extrêmes des erreurs possibles sera toujours égale à l'unité. Et comme  $\varphi(x)$  est nulle pour les valeurs non comprises entre ces limites, on peut dire, dans tous les cas, que

$$\int_{-\infty}^{\infty} \varphi(x) dx = 1.$$

Considérons l'intégrale

$$\int_{-\infty}^{\infty} x\varphi(x) dx$$

et représentons sa valeur par  $k$ . Si les causes d'erreur sont telles, qu'il n'y ait aucune raison pour que deux erreurs égales et de signes contraires aient des facilités inégales, ou aura

$$\varphi(x) = \varphi(-x)$$

et, par suite  $k = 0$

Nous en concluons que, si  $k$  ne s'évanouit pas et a, par exemple, une valeur positive, il existe nécessairement une cause d'erreur qui produit uniquement des erreurs positives ou qui, tout au moins, les produit plus facilement que les erreurs négatives. Cette quantité  $k$ , qui est la moyenne de toutes les erreurs possibles, ou encore la valeur moyenne de  $x$  peut être désignée commodément sous le nom de *partie constante de l'erreur*. Du reste, on prouve facilement que la partie constante de l'erreur totale est la somme des parties des constantes des erreurs simples qui la composent.

Si la quantité  $k$  est supposée connue et qu'on la retranche du résultat de chaque observation, en désignant par  $x'$  l'erreur de l'observation ainsi corrigée, et la probabilité correspondante par  $\varphi'(x')$ , on aura

1. *Méthode des moindres carrés par C. F. Gauss*, traduite par J. Bertrand, Paris, Mallet-Bachelier, 1855



$$x' = x - k, \quad \varphi'(x') = \varphi(x) \text{ et, par suite,}$$

$$\int_{-\infty}^{\infty} \varphi'(x') dx' = \int_{-\infty}^{\infty} x \varphi(x) dx - k \int_{-\infty}^{\infty} \varphi(x) dx = k - k = 0$$

en sorte que les erreurs des observations corrigées n'ont pas de *partie constante*. Ce qui, du reste, semble évident *a priori*.

La valeur de l'intégrale  $\int_{-\infty}^{\infty} x \varphi(x) dx$ , c'est-à-dire la valeur moyenne de  $x$  fait connaître l'existence ou la non-existence d'une erreur constante, ainsi que la valeur de cette erreur ; de même l'intégrale

$$\int_{-\infty}^{\infty} x^2 \varphi(x) dx,$$

c'est-à-dire la valeur moyenne de  $x^2$ , paraît très propre à définir et à mesurer, d'une manière générale, l'incertitude d'un système d'observations ; de telle sorte qu'entre deux systèmes d'observations inégalement précises, on devra regarder comme préférable celui qui donne à l'intégrale :  $\int_{-\infty}^{\infty} x^2 \varphi(x) dx$ , une moindre valeur.

Mais quelle perte doit-on assimiler à une erreur déterminée ? C'est ce qui n'est pas clair en soi ;

cette détermination dépend en partie de notre volonté. Il est évident, d'abord, que la perte ne doit pas être regardée comme proportionnelle à l'erreur commise ; car, dans cette hypothèse, une erreur positive représentant une perte, l'erreur négative devrait être regardée comme un gain : la grandeur de la perte doit, au contraire, s'évaluer par une fonction de l'erreur dont la valeur soit toujours positive. Parmi le nombre infini de fonctions qui remplissent cette condition, il semble naturel de choisir la plus simple, qui est, sans contredit, le carré de l'erreur, et, de cette manière, nous sommes conduits au principe proposé plus haut.

Si l'on objecte que cette convention est arbitraire et ne semble pas nécessaire, nous en convenons volontiers. La question qui nous occupe a, dans sa nature même, quelque chose de vague et ne peut être bien précisée que par un principe jusqu'à un certain point arbitraire. La détermination d'une grandeur par l'observation peut se comparer, avec quelque justesse, à un jeu dans lequel il y aurait une perte à craindre et aucun gain à espérer ; chaque erreur commise étant assimilée à une perte que l'on fait, la crainte relative à un pareil jeu doit s'exprimer par la perte probable, c'est-à-dire par la somme des produits des diverses pertes possibles par leurs probabilités respectives.

Laplace a considéré la question d'une manière analogue, niais en adoptant, pour mesure de la perte, l'erreur elle-même prise positivement. Cette hypothèse, si nous ne nous faisons pas illusion, n'est pas moins arbitraire que la nôtre : faut-il, en effet, regarder une erreur double comme plus ou moins regrettable qu'une erreur simple répétée deux fois, et faut-il, par suite, lui assigner une importance double ou plus que double ? C'est une question qui n'est pas claire, et sur laquelle les arguments mathématiques n'ont aucune prise ; chacun doit la résoudre à son gré. On ne peut nier pourtant que l'hypothèse de Laplace ne s'écarte de la loi de continuité et ne soit, par conséquent, moins propre à une étude analytique ; la nôtre, au contraire, se recommande par la généralité et la simplicité de ses conséquences. »

Gauss, dans la suite de son mémoire, démontre qu'il est inutile de supposer que les lois des erreurs d'observation soient normales. Il écrit :

« Lorsqu'une observation dont la précision n'est pas absolue, fait connaître une certaine quantité liée analytiquement à une grandeur inconnue, le résultat de cette observation peut fournir pour l'inconnue une valeur erronée, mais dans la détermination de laquelle il n'y a rien d'arbitraire qui puisse donner lieu à un choix plus ou moins vraisemblable,

Mais si plusieurs fonctions de la même inconnue sont données par des observations imparfaites, chaque observation fournira une valeur de l'inconnue, et l'on pourra également obtenir des valeurs, par la combinaison de plusieurs observations. Il y a évidemment une infinité de manières d'y parvenir ; le résultat sera soumis, dans tous les cas, à une erreur possible. Selon la combinaison adoptée, l'erreur moyenne à craindre pourra être plus ou moins grande.

La même chose aura lieu si plusieurs quantités observées dépendent à la fois de plusieurs inconnues. Selon que le nombre des observations sera égal au nombre des inconnues, ou plus petit ou plus grand que ce nombre, le problème sera déterminé, ou indéterminé, ou plus que déterminé (du moins en général), et, dans ce troisième cas, les observations pourront être combinées d'une infinité de manières pour fournir les valeurs des inconnues. Parmi ces combinaisons, il faudra choisir les plus avantageuses, c'est-à-dire celles qui fournissent des valeurs pour lesquelles l'erreur moyenne à craindre est la moindre possible. Ce problème est certainement le plus important que représente l'application des mathématiques à la philosophie naturelle.

Nous avons montré, dans la *Théorie du Mouvement des Corps célestes*, comment on trouve les valeurs les plus probables des inconnues lorsque l'on connaît la loi de probabilité des erreurs des observations, et comme, dans presque tous les cas, cette loi par sa nature reste hypothétique, nous avons appliqué cette théorie à l'hypothèse très plausible, que la probabilité de l'erreur  $x$  soit proportionnelle à  $e^{-k^2 x^2}$  ; de là cette méthode que j'ai suivie, surtout dans les calculs astronomiques, et que maintenant la plupart des calculateurs emploient sous le nom de *Méthode des moindres carrés*.

Dans la suite, Laplace, considérant la question sous un autre point de vue, montra que ce principe est préférable à tous les autres, quelle que soit la loi de probabilité des erreurs, pourvu que le nombre des observations soit très grand. Mais lorsque ce nombre est restreint, la question est encore intacte ; de sorte que, si l'on rejette notre loi hypothétique, la méthode des moindres carrés serait préférable aux autres, par la seule raison qu'elle conduit à des calculs plus simples.

Nous espérons donc être agréables aux géomètres en démontrant, dans ce Mémoire, que la méthode des moindres carrés fournit les combinaisons les plus avantageuses des observations, non seulement approximativement, mais encore d'une manière absolue, et cela quelle que soit la loi de probabilité des erreurs et quel que soit le nombre des observations, pourvu que l'on adopte pour l'erreur moyenne, non pas la définition de Laplace, mais celle que nous avons donnée plus haut ».

## 16.5. Biographie de Gauss (1777-1855)

Karl Friedrich Gauss naît le 30 avril 1777 à Brunswick. Il est le fils unique de parents ouvriers. Son père exerce différents métiers en fonction des besoins : jardinier, gardien de canal ou maçon. Sa mère, fille d'un tailleur de pierre, est une femme intelligente et de caractère. De tous les mathématiciens, Gauss est sans doute le plus précoce. Il apprend sans aucune aide à calculer avant de parler. À trois ans, il corrige, dit-on, une erreur dans le calcul de la paye des ouvriers de son père. On dit également qu'à huit ans, son instituteur, Büttner, pour occuper ses élèves, demande de calculer la somme des cent premiers entiers. Gauss dépose immédiatement son ardoise, où est écrite la réponse, sur le bureau du maître, comme il était d'usage. Gauss passe alors l'heure suivante les bras croisés en attendant que ses camarades finissent le calcul. À la fin de l'heure, l'instituteur découvre les ardoises, la seule bonne réponse est celle de Gauss. Lorsqu'il a quatorze ans, Gauss est présenté au Duc de Brunswick, qui décide de financer ses études. Il entre alors au *Collegium Carolinum* de la ville où il lit les œuvres de Newton, Euler et Lagrange. À partir de 1795, il est étudiant à l'université de Göttingen.

Gauss n'a que dix-neuf ans lorsque, le 30 mars 1796, il décrit la construction du polygone de dix-sept côtés avec la règle et le compas<sup>1</sup>. Ce succès le détermine à devenir mathématicien. Il commence le jour même un journal qu'il tiendra pendant dix-huit ans, et dans lequel il inscrira cent quarante-six énoncés, extrêmement brefs, des résultats de ses travaux. Ce journal ne fait que dix-neuf pages et ne sera retrouvé qu'en 1898, puis publié par Félix Klein en 1901. Les travaux de Gauss en mathématiques sont nombreux, profonds et variés. Félix Klein présente Gauss comme le « sommet imposant qui domine tous les mathématiciens du XVIII<sup>e</sup> siècle ». Toujours à dix-neuf ans, Gauss découvre et démontre un remarquable théorème de la théorie des nombres : la loi de réciprocité quadratique.

Dès son adolescence, en 1792, Gauss réfléchit à une géométrie ne s'appuyant pas sur le postulat des parallèles d'Euclide. À partir de 1813, et jusqu'en 1831, il développe une géométrie non euclidienne, mais décide de ne pas diffuser ses résultats. Les notes et la correspondance de Gauss sur ce sujet ne seront effectivement publiées qu'en 1855 (l'année de sa mort). En 1832, Gauss apprend que le mathématicien hongrois Janos Bolyai vient de faire des travaux semblables. Le mathématicien russe Nikolai Ivanovitch Lobatchevski publie également, entre 1829 et 1837, des travaux sur une géométrie non euclidienne comparable. L'un des rares élèves de Gauss, Bernhard Riemann, publiera à son tour, en 1868, son célèbre mémoire concernant un second type de géométrie non euclidienne.

1. On trouve dans *Les Éléments* d'Euclide la construction, à la règle et au compas, des polygones réguliers à 3, 4, 5 et 15 côtés. Ceux à 7, 9, 11 et 13 côtés n'y figurent pas.

À l'âge de vingt ans, en 1797, Gauss indique qu'il est tellement submergé d'idées mathématiques, qu'il ne trouve le temps de n'écire qu'une petite partie d'entre elles. En 1801, Gauss publie ses fameuses *Disquisitiones arithmeticae*. Il y introduit la notion de congruence et trace les grandes lignes de la recherche en théorie des nombres pour au moins un siècle. Il n'est pas exagéré de dire que Gauss est pour la théorie des nombres, ce qu'Euclide a été pour la géométrie.

En 1799, Gauss passe son doctorat à l'université de Helmstedt et donne la première démonstration du théorème fondamental de l'algèbre. Seulement deux ans après sa thèse, Gauss change de direction : il s'intéresse désormais à l'astronomie.

En mathématique, Gauss a travaillé en solitaire et n'a eu ni collaborateurs, ni correspondants, et seulement peu d'étudiants. Parmi ses rares étudiants trois deviendront des mathématiciens célèbres : Peter Gustav Lejeune Dirichlet (1805-1859), qui lui succède à Göttingen en 1855, Richard Dedekind (1831-1916) et, bien sûr, Bernhard Riemann (qui à son tour succède à Dirichlet en 1859).

Début janvier 1801, Giuseppe Piazzi (1746-1826) découvre l'astéroïde Cérès, mais ne parvient pas à cerner son orbite complète : il n'a pu suivre l'astéroïde que sur neuf degrés de sa course. Gauss décide de déterminer sa trajectoire pour pouvoir ensuite le re-localiser. Pour cela il fait le calcul de l'orbite, utilise des méthodes numériques fondées sur les moindres carrés, mène ses calculs jusqu'à 21 décimales... et parvient à la détermination exacte de la position de l'astéroïde, que l'astronome Zach observe à l'endroit prévu par Gauss. Ses travaux assurent alors sa renommée.

En 1807, le baron von Humboldt, ami d'Arago, souhaite faire nommer Gauss au poste de directeur de l'observatoire de Göttingen. Il demande, dit-on, à Laplace quel est, selon lui, le plus grand mathématicien d'Allemagne. Laplace répond : « Pfaff<sup>1</sup> ». Humboldt s'étonne : « Que faites-vous donc de Gauss ? » Gauss, répond Laplace, est le plus grand mathématicien du monde ! C'est ainsi que Gauss est nommé professeur d'astronomie et directeur de l'observatoire de Göttingen en 1807 (il y restera toute sa vie). En parallèle à ses travaux en mathématiques et en astronomie, Gauss entreprend l'édification de l'observatoire de Göttingen, qui entre en opération en 1816, mais ne sera définitivement opérationnel qu'en 1821. À partir de 1817, Gauss consacre une bonne partie de son activité aux travaux de géodésie. En 1821, il est nommé conseiller scientifique pour l'établissement géodésique du cadastre dans le gouvernement du Danemark et du Hanovre, dont dépend Göttingen.

---

1. Johan Friedrich Pfaff a été le professeur de Gauss, et c'est sous sa direction que Gauss a obtenu son doctorat en 1799.

Il invente à cette occasion la géométrie différentielle qui jouera, un siècle plus tard, un si grand rôle en physique et, en particulier, en relativité générale.

Peu après son installation à Göttingen, le père de Gauss meurt. En 1809, Carl Friedrich perd sa femme Johanna<sup>1</sup> (qui n'a que vingt-neuf ans) en mettant au monde leur troisième enfant, Louis. Gauss est littéralement brisé par les événements. Il demande à son ami astronome Olbers de l'héberger pendant quelques semaines « pour retrouver de nouvelles forces dans les bras d'un ami, des forces pour une vie qui ne vaut que parce qu'elle appartient à mes trois jeunes enfants ». Mais le sort s'acharne et le dernier enfant, Louis, meurt l'année suivante. Gauss reste avec son premier fils, Joseph (1806-1873) et sa fille Wilhelmina (1808-1846). En 1810, il se remarie avec la meilleure amie de sa première épouse, Friederica Wilhelmina Waldeck, qu'il surnomme Minna. Elle lui donnera trois enfants : Eugen (1811-1896), Wilhelm (1813-1879) et Thérèse (1816-1864).

En 1820, Gauss se tourne vers la géodésie. Il s'y consacrera huit ans. Dès 1821, il devient conseiller des gouvernements du Hanovre et du Danemark et est chargé de réaliser des relevés géodésiques. Il invente, à cette occasion, l'héliotrope, appareil qui en réfléchissant les rayons solaires dans une direction bien déterminée, permet de faire des alignements précis. Après ses périodes mathématique (1796-1801), astronomique (1801-1820) et géodésique (1820-1828), Gauss se consacre à la physique. Dans ce domaine, ses travaux les plus connus se rapportent au magnétisme et à l'électricité (théorème de Gauss). Toutefois il se consacre également à l'optique (conditions de Gauss de l'optique géométrique), à la capillarité, à l'acoustique, à la cristallographie... Sur la recommandation de Gauss, Wilhelm Weber (1804-1891) devient professeur de physique à Göttingen. Peu après, en septembre 1831, quelques jours après la mort de la seconde femme de Gauss, une étroite collaboration débute entre les deux hommes. Une sincère amitié s'établit alors entre eux et permet à Gauss de surmonter sa peine. En 1833, Gauss publie un mémoire, *Intensitas vis magneticae terrestris*, dans lequel il utilise, pour la première fois, des unités absolues pour mesurer les quantités électriques et magnétiques. La collaboration avec Weber est une période fructueuse qui se concrétise par l'invention d'instruments novateurs tels que le télégraphe électrique (1833), le magnétomètre<sup>2</sup> ou l'électrodynamomètre. Mais, en 1837, des raisons politiques viennent altérer leur amitié. À partir de cette date ils poursuivront leurs travaux séparément. Dès lors, l'activité de Gauss, qui vient d'avoir soixante ans, se ralentit. En 1850, sa santé commence à se détériorer. Il assistera toutefois à la leçon inaugurale de son élève Riemann en juin 1854.

1. Carl Friedrich Gauss avait épousé Johanna Ostoff le 9 octobre 1805.

2. En reconnaissance de leurs travaux, l'unité de champ magnétique recevra le nom de gauss, et l'unité de flux magnétique celui de weber.

Gauss meurt durant son sommeil le 23 février 1855.

Gauss marque une transition dans l'histoire des mathématiques. Il est à la fois le premier mathématicien moderne, par ses méthodes, et le dernier des classiques, par le choix des sujets abordés. Comme le souligne E.T. Bell « il est encore présent partout dans l'univers des mathématiques ».

## 16.6. Principaux ouvrages de Gauss

Gauss a publié environ 150 mémoires scientifiques, dont :

- *Disquisitiones Arithmeticae*, Leipzig, 1801.
- *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*, Hamburg, 1809.
- *Theoria combinationis observationum erroribus minimis obnoxiae*, Göttingen, 1823-1828.
- *Theoria residuorum biquadraticorum*, Göttingen, 1828-1832.
- *Intensitas vis magneticae terrestres ad mensuram absolutam revocata*, Göttingen, 1832-1837.
- *Atlas der Erdmagnetismus* avec W. E. Weber, Leipzig, 1840.
- *Dioptrische Untersuchungen*, Göttingen, 1841.

## 16.7. Références

- R. Dedekind, *Gauss in seiner Vorlesung über die Methode der kleinsten Quadrate*, Festschr Kgl. Ges. Wiss., Göttingen, 45-59, 1901.
- K. O. May, *Gauss*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 5, p. 298-315, 1970.
- I. Todhunter, *On the method of least squares*, Trans. Cambridge Philos. Soc., 11, 219-238, 1869.
- W. K. Bühler, *Gauss: A Biographical Study*, Berlin, 1981.
- G. W. Dunnington, *Carl Friedrich Gauss: Titan of Science*, New York, 1955.
- T. Hall, *Carl Friedrich Gauss: A Biography*, 1970.
- R. L. Plackett, *The discovery of the method of least squares*, Biometrika 59, 239-251, 1972.

- O. B. Sheynin, *Gauss and the theory of errors*, AHES, vol. 20, p. 21-72, 1979.
- S.M. Stigler, *Gauss and the invention of least squares*, Ann. Stat., 9, 465-474, 1981.

# Poisson

*La probabilité d'un événement est la raison que nous avons de croire qu'il aura ou qu'il a eu lieu<sup>1</sup>.*

Le premier travail de Siméon-Denis Poisson relatif aux probabilités a été lu par son auteur le 13 mars 1820 à l'Académie des sciences et a pour titre *Mémoire sur l'avantage du banquier au jeu de trente et quarante*<sup>2</sup>. Par ailleurs, entre 1824 et 1829, Poisson a publié dans la revue *Connaissance des Temps*, un certain nombre de papiers sur la théorie des erreurs<sup>3</sup>. Mais ses travaux les plus importants datent de 1829 et 1837, ce sont ceux que nous commentons ci-dessous.

## 17.1. Le Mémoire sur la proportion des naissances des filles et des garçons

Poisson a lu à l'Académie des sciences de Paris, le 8 février 1829, un travail intitulé *Mémoire sur la proportion des naissances des filles et des garçons*.<sup>4</sup>

Le résumé qu'il en a donné dans les *Annales de Chimie et de Physique* comporte une introduction décrivant les observations faites sur plusieurs années en France sur le sujet, suivie du texte reproduit ci-dessous :

« Ce Mémoire est divisé en deux paragraphes. Dans le premier, on considère la répétition d'un événement simple, dont la probabilité est donnée *a priori*, et l'on parvient d'une manière nouvelle, à quelques égards, au fameux théorème de Jacques Bernoulli, d'après lequel, sur un très grand nombre d'épreuves, les événements arrivent proportionnellement à leurs probabilités respectives, avec une précision qui s'accroît sans limite à mesure que leur nombre augmente. Le second paragraphe est relatif aux probabilités des événements simples et des événements futurs, déterminées

1. *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités*, Paris, Bachelier, 1837, p. 30
2. Publié dans *Annales de Mathématiques pures et appliquées*, t. 16, p. 173-208, 1825.
3. Cf. B. Bru, « Poisson, le calcul des probabilités et l'instruction publique », in *Siméon-Denis Poisson et la science de son temps* (édité par M. Métivier, P. Costabel, P. Dugac), Palaiseau, École Polytechnique, 1981.
4. Publié dans *Mémoires de l'Académie des Sciences*, tome IX, p. 239-308, Didot, Paris, 1830. Un extrait avait été inclus dans les *Annales de Chimie et de Physique*, Tome XL, p. 39-46, 1829.



d'après les événements observés. Cette partie du calcul des hasards est l'inverse de celle qui fait l'objet du premier paragraphe ; les applications en sont les plus utiles et les plus fréquentes, car le plus souvent nous ignorons la probabilité des événements simples et nous devons la déduire des faits que nous observons, pour en conclure ensuite la probabilité des événements futurs. Mais il y a une différence essentielle entre la probabilité d'un événement simple qui nous est quelquefois donnée par hypothèse, et celle, que nous pouvons connaître par l'observation : relativement à celle-ci, si elle est susceptible, *a priori*, de tous les degrés possibles, depuis zéro jusqu'à la certitude, nous ne pouvons jamais déterminer que des limites entre lesquelles elle est comprise, qui se resserrent de plus en plus, ou deviennent de plus en plus probables, à mesure que nous employons un plus grand nombre d'observations ; mais il faudrait que ce nombre fût infini pour que la probabilité que nous en déduisons eût une valeur certaine. Il en résulte que la probabilité du nombre de fois qu'un événement arrivera sur un très grand nombre d'épreuves ne sera pas la même, selon que sa chance nous sera donnée par hypothèse, ou que nous en aurons déterminé, par l'expérience, des limites aussi resserrées qu'on voudra. On avait, en effet, déjà remarqué que les formules relatives au premier cas ne s'appliquent pas au second, si ce n'est lorsque le nombre des événements futurs est très petit par rapport à celui des événements passés. Pour compléter cette importante partie du calcul des hasards, il restait donc à trouver des formules qui convinssent au second cas, et qui fussent encore applicables quand le nombre des événements futurs est comparable à celui des épreuves qui ont fait connaître les limites de la probabilité simple. On verra dans ce Mémoire comment il a fallu, pour les obtenir, modifier les formules ordinaires qui se rapportent au cas où la valeur de cette probabilité est donnée avec certitude.

Les limites de probabilité des événements simples que l'on détermine par le calcul et l'observation, ne font pas connaître si la chance de l'un de ces événements surpasse celle d'une autre, d'une quantité égale ou supérieure à une fraction donnée, et quelle est la probabilité de cette différence. Cependant la comparaison des chances respectives de plusieurs événements, ou d'un même événement à des époques différentes, est une question intéressante et susceptible de nombreuses applications. On en trouvera la solution dans mon Mémoire. En appliquant cette solution aux naissances des deux sexes qui ont eu lieu en France pendant chacune des années 1817 et 1826, pour lesquelles, la proportion de ces naissances a présenté le plus de divergence, on en conclut, avec une grande probabilité, que la chance d'une naissance masculine a été plus considérable dans la première année que dans la seconde, et que néanmoins il n'y a pas plus d'un contre un à parier que la différence ait excédé deux millièmes. Ainsi, dans notre pays et à l'époque actuelle, la probabilité de la naissance d'un garçon, pour chaque année entière, est à peu près constante et comprise entre des limites qui s'écartent peu de 0,5159, ou d'environ 16/15. L'expérience seule peut nous apprendre si elle variera davantage par la suite, et si elle est la même dans les différentes saisons et chez les autres peuples. Nous savons déjà qu'elle n'est pas indépendante de l'état de la société, puisque le nombre des enfants qui naissent hors mariage influe sur la proportion des filles et des garçons.

Depuis que Pascal a donné les premiers exemples du calcul des hasards, ce calcul a été l'objet des travaux d'un grand nombre de géomètres, et le sujet des considérations les plus ingénieuses, soit pour former et résoudre les équations qui s'y rapportent,

soit pour convertir en nombres les formules auxquelles il conduit dans la plupart de ses applications. Si j'ai ajouté quelque chose à la partie de ce calcul qui lie, sans que leurs causes soient connues, les événements futurs aux événements passés, je le dois à la puissance de l'analyse que j'ai employée, et dont j'ai puisé le principe dans *la Théorie analytique des probabilités* ; ouvrage aussi éminemment remarquable par la variété des questions qui y sont traitées, que par la généralité de méthodes que Laplace a imaginées pour les résoudre. »

Dans son mémoire complet, Poisson donne une démonstration de la loi des grands nombres de Bernoulli, différente de celle qu'il a incluse plus tard dans son livre de 1837 *Recherches sur la probabilité des jugements*, et que nous avons donnée au chapitre consacré à Bernoulli. Ce travail, dans le mémoire de 1829-1830, est intitulé *Probabilité de la répétition d'un événement dont la chance est donnée*.

La démonstration s'étale sur une vingtaine de pages (p. 243-264).

Poisson appelle  $p$ , la probabilité d'un événement A, et  $q$  celle de l'événement contraire, de sorte qu'on ait  $p + q = 1$ , alors, écrit Poisson : « en appelant  $x'$  le nombre de fois que A arrivera sur le nombre  $n$  d'épreuves on peut dire que

la probabilité que la différence  $\frac{x'}{n} - p$  sera comprise entre les deux limites  $\pm u \sqrt{\frac{2pq}{n}}$  est

$$U = 1 - \frac{2}{\sqrt{\pi}} \int_u^\infty e^{-t^2} dt + \frac{e^{-u^2}}{\sqrt{2\pi npq}} \quad (1).$$

Et il ajoute :

« On pourra toujours prendre  $u$  assez grand pour que cette probabilité  $U$  diffère d'aussi peu qu'on voudra de la certitude[...]. La valeur de  $u$  ayant été convenablement choisie, et demeurant constante, les limites de  $\frac{x'}{n} - p$  se resserreront de plus en plus à mesure que le nombre  $n$  augmentera ; le rapport  $\frac{x'}{n}$  du nombre de fois que l'événement A arrivera au nombre total des épreuves, différera donc de moins en moins de la probabilité  $p$  de cet événement ; et l'on pourra toujours prendre  $n$  assez grand pour qu'il y ait la probabilité  $U$  que la différence  $\frac{x'}{n} - p$  sera aussi petite que l'on voudra ; ce qui est, comme on sait, le théorème de Jacques Bernoulli sur la répétition, dans un très grand nombre d'épreuves, d'un événement dont la chance est donnée *a priori*. »

Dans la suite du même texte<sup>1</sup>, Poisson démontre ce qui est appelé *la loi de Poisson*. Si  $X$  est la probabilité que sur un nombre d'épreuves  $n$ , un événement A, de probabilité  $p$  ne se produira pas plus de  $x$  fois, avec  $pn = \omega$ , on a

$$X = \sum_{k=0}^x e^{-\omega} \frac{\omega^k}{k!} = (1 + \omega + \frac{\omega^2}{2} + \dots + \frac{\omega^x}{1.2.3\dots x}) e^{-\omega}.$$

1. Page 261 du texte original de 1830.

Poisson ajoute :

« On voit que si  $x$  n'est pas un petit nombre, cette valeur de  $X$  différera très peu de l'unité. Si l'on a, par exemple,  $x = 10$  et  $\omega = 1$ , la différence  $1 - X$  sera à peu près de un cent millionième, c'est-à-dire qu'il est presque certain qu'un événement dont la chance est  $1/n$ , n'arrivera pas plus de dix fois, sur un très grand nombre d'épreuves.

Dans le cas où  $x = 0$ , on a  $X = e^{-\omega}$ , pour la probabilité que sur un très grand nombre  $n$  d'épreuves, un événement dont la chance est  $\omega/n$  n'arrivera pas une seule fois »

Il est à noter que l'économiste et statisticien russe Ladislav Bortkiewicz (1868-1931) publia en 1898 un livre sur la distribution de Poisson intitulé *Das Gesetz der kleinen Zahlen* (La loi des petits nombres) dans lequel il montra que des événements rares survenant dans une large population suivent la loi de Poisson, même lorsque la probabilité de ces événements est variable. Il se servit, pour vérifier ce fait, des statistiques de la cavalerie Prussienne, établies sur 20 ans, qui dénombraient le nombre de soldats tués par des ruades de cheval.

Bortkiewicz appliqua en 1913 ses méthodes statistiques à la décroissance radioactive : Un atome radioactif a une probabilité  $\rho$  de se désintégrer

pendant un laps de temps  $\Delta t$  qui est  $\rho = 1 - e^{-\frac{\Delta t}{\tau}}$ , où  $\tau$  est la constante radioactive de l'élément. Un radioélément possédant  $N_0$  atomes à l'instant

$t = 0$ , possédera donc  $N = N_0 e^{-\frac{t}{\tau}}$  atomes au temps  $t$ .

L'importance de la loi de Poisson ne fut reconnue que tardivement, sauf en Russie, grâce à l'école de Chebyshev.

La seconde partie du *Mémoire sur la proportion des naissances des filles et des garçons* est intitulée *Probabilités des événements simples et des événements futurs d'après les événements observés*.

Jusqu'ici, écrit Poisson<sup>1</sup>,

« nous avons supposé connue *a priori*, la chance de l'événement A, et nous en avons conclu la probabilité d'un événement futur, relatif à la répétition de A sur un très grand nombre d'épreuves ; mais dans les applications du calcul des hasards aux phénomènes naturels, et particulièrement dans la question indiquée par le titre de ce Mémoire, la valeur de  $p$  doit, au contraire, se déduire autant qu'il est possible, des événements observés en très grands nombres, pour servir ensuite à calculer la probabilité des événements futurs. »

Après calcul, il obtient le résultat suivant :

Si  $n$  est le nombre des événements futurs,  $m$  le nombre total des observations et  $s$  le nombre de fois que cet événement est arrivé,  $U$  exprime la probabilité

1. Page 265 du texte original de 1830.

que le nombre d'événements  $A$  sur un nombre d'épreuves  $n$ , sera renfermé entre des limites :

$$N \pm \frac{u}{m} \sqrt{2(n+1)(m-s)s}.$$

Si  $n$  est très petit par rapport à  $m$ ,

$$U = 1 - \frac{2}{\sqrt{\pi}} \int_u^{\infty} e^{-t^2} dt + \frac{me^{-u^2}}{\sqrt{2\pi n(m-s)s}}$$

Ce résultat coïncide avec l'équation (1) quand on fait dans cette équation,  $p = s/m$  et  $q = (m-s)/m$ .

Poisson ajoute ce commentaire<sup>1</sup> :

« Lors donc que le nombre  $n$  des événements futurs est très petit eu égard au nombre  $m$  des événements observés, les limites du nombre de fois que  $A$  arrivera et leur probabilité pourront se calculer en prenant pour la probabilité  $p$  de  $A$ , le rapport  $s/m$  du nombre de fois que cet événement est arrivé au nombre total des observations, comme si cette valeur de  $p$  était certaine et donnée *a priori*. Mais il n'en est pas ainsi quand les deux nombres  $n$  et  $m$  sont du même ordre de grandeur : quoique la valeur de  $p$  conclue de l'expérience, soit comprise, avec une probabilité qui diffère très peu de la certitude, entre des limites qui s'écartent aussi très peu de  $s/m$ , cependant, toutes choses d'ailleurs égales, le nombre de fois que  $A$  arrivera sera renfermé entre des limites moins étroites que celles qui auraient lieu s'il était certain qu'on eût  $p = s/m$ . On pourra alors supprimer le dernier terme de la valeur de  $U$ , qui sera de

$$\text{l'ordre de } \frac{1}{\sqrt{m}}, \text{ et prendre simplement } U = 1 - \frac{2}{\sqrt{\pi}} \int_u^{\infty} e^{-t^2} dt. ».$$

Le reste du mémoire est une application des résultats précédents à la proportion des naissances des filles et des garçons en France. La conclusion de Poisson, à ce propos, est la suivante<sup>2</sup> :

« Nous pouvons donc conclure qu'à l'époque actuelle et pour la France entière, la probabilité d'une naissance masculine<sup>3</sup> n'éprouve que de très petites variations d'une année sur l'autre, et prendre pour sa valeur, la moyenne des dix années que nous avons considérées, c'est-à-dire 0,5159. »

1. Pages 278-279 du texte original de 1830.

2. Page 307 du mémoire original de 1830.

3. Depuis qu'on la mesure, la proportion des naissances des garçons et des filles a semblé quasi invariable. Cette mesure est appelée *taux de masculinité*. Ce taux correspond au nombre de naissances de garçons pour cent naissances de filles. Il est, de nos jours, et a d'ailleurs toujours été, de 105 ou 106, c'est-à-dire en probabilité, comme l'indique Poisson, 0,512 à 0,514. Tout se passe comme si la nature compensait la surmortalité observée des garçons. Ce taux semble donc être une constante biologique de l'espèce humaine.

**RECHERCHES**  
SUR LA  
**PROBABILITÉ DES JUGEMENTS**  
EN MATIÈRE CRIMINELLE  
**ET EN MATIÈRE CIVILE,**

PASTÉURISÉES

**DES RÈGLES GÉNÉRALES DU CALCUL DES PROBABILITÉS;**

**PAR S.-D. POISSON,**

Membre de l'Institut et du Bureau des Longitudes de France; des Sociétés Royales de Londres et d'Édimbourg; des Académies de Berlin, de Stockholm, de Saint-Petersbourg, d'Upsal, de Boston, de Turin, de Naples, etc.; des Sociétés, Italienne, astronomique de Londres, Philomatique de Paris, etc.

---

**PARIS,**  
**BACHELIER, IMPRIMEUR-LIBRAIRE**  
POUR LES MATHÉMATIQUES, LA PHYSIQUE, ETC.

QUAI DES AUGUSTINS, N° 55.

---

**1837**

## 17.2. La publication de 1837

Le second ouvrage important de Poisson sur les Probabilités a pour titre *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités*<sup>1</sup> et a été publié en 1837.

Poisson définit, dans les premières pages de l'ouvrage, sa conception du but du calcul des probabilités<sup>2</sup> :

« Le calcul des probabilités a pour objet de déterminer dans chaque question d'éventualité ou de doute, le rapport du nombre des cas favorables à l'arrivée d'un événement, ou à la vérité d'une chose, au nombre de tous les cas possibles ; de sorte que nous puissions connaître, d'une manière précise, d'après la grandeur de cette fraction plus ou moins approchante de l'unité, la raison que nous avons de croire que cette chose soit vraie, ou que cet événement a eu ou aura lieu, et que nous puissions aussi, sans aucune illusion, comparer cette raison de croire, dans deux questions de nature toute différente. Il est fondé sur un petit nombre de règles que nous allons exposer, et qui se démontrent en toute rigueur [...]. Ses principes doivent être regardés comme un supplément nécessaire de la logique, puisqu'il y a un si grand nombre de questions où l'art de raisonner ne saurait nous conduire à une entière certitude. Aucune autre partie des mathématiques n'est susceptible d'applications plus nombreuses et plus immédiatement utiles. »

Le premier chapitre du livre donne quelques définitions relatives au calcul des probabilités dont celle de l'espérance mathématique :

« Le produit d'un gain et de la probabilité de l'obtenir est ce qu'on appelle *l'espérance mathématique* de chaque personne intéressée dans une spéculation quelconque. »

Poisson a été le premier à introduire la notion de variable aléatoire discrète. Il écrit<sup>3</sup> :

« Supposons [...] qu'il y ait un nombre donné  $\lambda$  d'événements, dont un seul devra arriver à chaque épreuve. Ce cas est celui où l'on considère une chose A d'une nature quelconque, susceptible d'un nombre  $\lambda$  de valeurs, connues ou inconnues, que je représenterai par  $a_1, a_2, \dots, a_\lambda$ , et parmi lesquelles une seule devra avoir lieu à chaque épreuve, de sorte que celle qui sera arrivée ou qui arrivera sera, dans cette question, l'événement observé ou l'événement futur. Soit aussi  $c_{i,i'}$ , la chance que la cause  $C_i$  si elle était certaine, donnerait à la valeur  $a_{i'}$  de A. Les valeurs de  $c_{i,i'}$  relatives aux divers indices  $i$  et  $i'$ , depuis  $i = 1$  jusqu'à  $i = i'$  et  $i'$  depuis  $i' = 1$  jusqu'à  $i' = \lambda$ , seront connues ou inconnues ; mais pour chaque indice  $i'$  on devra avoir

$$c_{i,1} + c_{i,2} + c_{i,3} + \dots c_{i,\lambda} = 1$$

1. Paris, Bachelier, 1837

2. *Id.*, p. 35-36.

3. Pages 140 et 141 de l'édition originale de 1837.

car si la cause C, était certaine, l'une des valeurs  $a_1, a_2, \dots, a_\lambda$   $a_{\lambda}$ , arriverait certainement en vertu de cette cause. »

En plus de la notion de variable aléatoire, Poisson introduit dans sa théorie la fonction de distribution de cette variable  $F(x) = P(A < x)$  et la dérivée de cette dernière la densité de probabilité. Il traite également le cas d'une variable aléatoire continue<sup>1</sup>.

Poisson indique, par ailleurs, que le sujet des probabilités des jugements a déjà été abordé par Condorcet et par Laplace. Il considère que, parmi les applications du calcul des probabilités, une des plus importantes est celle qui se rapporte à la probabilité des jugements, ou, en général, des décisions rendues à la pluralité des voix. Poisson poursuit<sup>2</sup> en écrivant :

« Condorcet est le premier qui ait essayé de la déterminer. [...] Son ouvrage renferme un discours préliminaire fort étendu, où l'auteur expose, sans le secours des formules analytiques, les résultats qu'il a obtenus, et où sont développées avec soin les considérations propres à montrer l'utilité de ce genre de recherches.

Dans son *Traité des probabilités*, Laplace s'est aussi occupé du calcul des chances d'erreur à craindre dans le jugement rendu contre un accusé, à une majorité connue, par un tribunal ou un jury composé d'un nombre de personnes également connu. La solution qu'il a donnée de ce problème, l'un des plus délicats de la théorie des probabilités, est fondée sur le principe qui sert à déterminer les probabilités des causes diverses auxquelles on peut attribuer les faits observés, principe que Bayes a présenté d'abord sous une forme un peu différente, et dont Laplace a fait ensuite le plus heureux usage, dans ses mémoires et dans son traité, pour calculer la probabilité des événements futurs d'après l'observation des événements passés ; mais, en ce qui concerne le problème de la probabilité des jugements, il est juste de dire que c'est à Condorcet qu'est due l'idée ingénieuse de faire dépendre la solution, du principe de Bayes en considérant successivement la culpabilité et l'innocence de l'accusé, comme une cause inconnue du jugement prononcé, qui est alors le fait observé, duquel il s'agit de déduire la probabilité de cette cause. L'exactitude de ce principe se démontre en toute rigueur ; son application à la question qui nous occupe ne peut non plus laisser aucun doute ; mais pour cette application, Laplace fait une hypothèse qui n'est point incontestable ; il suppose que la probabilité qu'un juré ne se trompera pas est susceptible de tous les degrés également possibles, depuis la certitude, représentée par l'unité, jusqu'à l'indifférence, qui répond dans le calcul à la fraction 1/2 et se rapporte à une égale chance d'erreur et de vérité. [...] Mais il existe une infinité de lois différentes de probabilité des erreurs qui satisfont à cette condition, sans qu'on soit obligé de supposer que la chance qu'un juré ne se trompera pas, ne puisse jamais descendre au-dessous de 1/2, et qu'au-dessus de cette limite, toutes ses valeurs soient également possibles. La supposition particulière de Laplace ne saurait donc être justifiée *a priori*. Soit à raison de cette hypothèse, soit à cause de leurs conséquences, qui m'ont paru inadmissibles, les solutions du problème

1. Page 254 de l'édition originale de 1837.

2. Pages 2 et 3 de l'édition originale de 1837.

de la probabilité des jugements que l'on trouve dans son *Traité des probabilités* et dans le *premier supplément* à ce grand ouvrage, et qui diffèrent l'une de l'autre, ont toujours laissé beaucoup de doutes dans mon esprit.

La formule de Laplace, pour exprimer la probabilité de l'erreur d'un jugement, ne dépend que de la majorité à laquelle il a été prononcé, et du nombre total des juges ; elle ne renferme rien qui soit relatif à leurs connaissances plus ou moins étendues dans la matière qui leur a été soumise. [...] »

Poisson entreprend alors de caractériser les recherches qu'il publie :

« Le caractère distinctif de cette nouvelle théorie de la probabilité des jugements criminels étant donc de déterminer d'abord, d'après les données de l'observation dans un très grand nombre d'affaires de même nature, la chance d'erreur du vote des juges, et celle de la culpabilité des accusés avant l'ouverture des débats, elle doit convenir à toutes les espèces nombreuses de jugements<sup>1</sup>. »

Les recherches de Poisson sur la probabilité des jugements sont basées sur la loi des grands nombres et sur la régularité statistiques de certains événements<sup>2</sup>. Il écrit<sup>3</sup> :

« La loi universelle des grands nombres est déjà pour nous un fait général et incontestable, résultant, d'expériences qui ne se démentent jamais. Cette loi étant d'ailleurs la base de toutes les applications du calcul des probabilités, on conçoit maintenant leur indépendance de la nature des questions et leur parfaite similitude, soit qu'il s'agisse de choses physiques ou de choses morales, pourvu que les données spéciales que le calcul exige, dans chaque problème, nous soient fournies par l'observation. Mais, vu son importance, il était nécessaire de la démontrer directement, c'est ce que j'ai tâché de faire ; et je crois y être enfin parvenu, comme on le verra dans la suite de cet ouvrage. »

On trouvera cette démonstration de Poisson dans notre chapitre consacré à Jacob Bernoulli (§ 5.1.4.).

Poisson utilise la loi des grands nombres pour déterminer deux paramètres destinés à calculer la probabilité d'un jugement erroné, à savoir : la probabilité que l'accusé soit coupable et la probabilité qu'un juré ne se trompe pas. Il se sert pour cela des statistiques disponibles et résume ses résultats en ces phrases :

« Dans les années qui ont précédé 1831, et pour la France entière, la probabilité de l'erreur d'une condamnation prononcée à la majorité *minima* de sept voix contre cinq était à très peu près 0,16 ou 0,04, selon qu'il s'agissait d'un crime contre les personnes ou d'un crime contre les propriétés ; sans distinction de l'espèce de crime, elle avait pour valeur 0,06. D'après la formule de Laplace, cette chance d'erreur serait la même dans tous les cas, et à peu près quintuple de 0,06. Mais il faut, en outre,

1. Page 25.

2. Cf. E. Barbin et Y. Marec, *Les recherches sur la probabilité des jugements de Siméon-Denis Poisson*, Histoire et Mesure, Année 1987, vol. 2, n°2, p. 39-58.

3. Page 12 de l'édition originale.



observer que l'intervention de la cour étant alors nécessaire, dans le cas de la plus petite majorité, cette chance d'erreur 0,06 se trouvait réduite à un peu moins d'un centième, si les juges confirmaient la décision des jurés ; en sorte que sur 1 597 condamnations qui ont eu lieu de cette manière, dans les cinq années écoulées depuis 1826 jusqu'à 1830, on peut croire qu'environ 15 ou 16 étaient erronées, en ce sens que les accusés n'étaient pas condamnables, mais non pas qu'ils fussent innocents<sup>1</sup>. »

Par ailleurs, dans le cours de l'ouvrage, Poisson reprend aussi une démonstration de la loi qui porte son nom dont il a été question dans le paragraphe relatif au mémoire de 1830. Nous en rappelons ci-dessous les propriétés et citons, simplement la phrase exacte écrite par Poisson à son propos :

Cette loi, appelée quelquefois loi des événements rares, est, comme nous l'avons vu plus haut, la loi de probabilité suivie par une variable aléatoire  $X$  qui indique le nombre de fois où un événement se produit sur un intervalle de temps  $T$ . Si sur cette période, l'événement se produit en moyenne un nombre de fois  $\lambda$  (appelé paramètre de la loi) la probabilité qu'il se produise  $k$  fois

c'est-à-dire que  $X$  prenne la valeur  $k$  est :  $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ ,

$\lambda$  étant un nombre strictement positif. L'espérance d'une telle loi, ainsi que sa variance, est  $\lambda$ , son écart type étant donc  $\sqrt{\lambda}$ . Pour  $T$  grand, on démontre que la loi binomiale de paramètres  $(T ; \lambda/T)$  converge vers la loi de Poisson.

Poisson, dans son livre de 1837, suppose que « les chances  $p$  et  $q$  de deux événements  $E$  et  $F$  sont constantes » et considère « la probabilité que dans un nombre d'épreuves  $\mu$  ou  $m + n$ ,  $E$  arrivera au moins  $m$  fois et  $F$  au plus  $n$  fois. Cette probabilité  $P$  sera la somme des  $m$  premiers termes du développement de  $(p + q)^\mu$ , ordonnée suivant les puissances de  $q$ . »

Après 17 pages de calculs, Poisson arrive à l'approximation suivante<sup>2</sup> :

« La probabilité qu'un événement dont la chance à chaque épreuve est la fraction très petite  $\frac{x}{\mu}$  n'arrivera pas plus de  $n$  fois dans un très grand nombre  $\mu$  d'épreuves est

$$P = (1 + x + \frac{x^2}{1.2} + \frac{x^3}{1.2.3} + \dots + \frac{x^n}{1.2.3\dots n}).e^{-x}.$$

Dans le cas  $n = 0$ , cette valeur de  $P$  se réduit à  $e^{-x}$  ; il y a donc cette probabilité  $e^{-x}$  que l'événement dont il s'agit n'arrivera pas une seule fois dans le nombre  $\mu$  d'épreuves, et conséquemment, la probabilité  $1 - e^{-x}$  qu'il arrivera au moins une fois.

Dès que  $n$  ne sera plus un très petit nombre, la valeur de  $P$  différera très peu de l'unité [...].

1. Page 25.

2. Pages 205-206, n° 81.

Si l'on a, par exemple  $x = 1$ , et qu'on suppose  $n = 10$ , la différence  $1 - P$  sera à peu près un cent millionième, de sorte qu'il est presque certain dont la chance très faible est  $1/\mu$  à chaque épreuve, n'arrivera pas plus de dix fois, dans le nombre  $\mu$  d'épreuves. »

### 17.3. Biographie de Poisson (1781-1840)

Le père de Siméon-Denis Poisson, Siméon Poisson (1744-1808), est un soldat ayant participé à la guerre de Sept Ans (1756-1763) dans la région de Hanovre. Toutefois, le mauvais traitement dont il a été l'objet de la part d'officiers nobles lui a fait quitter l'armée. Il occupe alors un poste administratif modeste. Siméon-Denis naît le 22 juin 1781 à Pithiviers. Plusieurs de ses frères et sœurs aînés sont morts en bas âge et lui-même a été pris de convulsions alors qu'il était encore bébé et ses parents ont cru de nouveau perdre un de leurs enfants. Sa mère fait alors appel à une nourrice et Siméon-Denis réussit à se sortir de cette période critique. Son père a eu une grande influence sur son jeune fils et lui a consacré beaucoup de son temps, lui apprenant en particulier à lire et écrire. Notre futur mathématicien a huit ans lorsque la révolution française éclate. Son père, séduit par les idées de la Révolution, s'enthousiasme pour cette cause et est nommé bailli<sup>1</sup> de la commune de Pithiviers.

Siméon Poisson souhaitait que son fils devienne médecin. Comme l'un des oncles de notre futur mathématicien était chirurgien à Fontainebleau, on l'envoya dans cette ville pour y bénéficier des premiers enseignements de son parent. Toutefois la peur du sang et la crainte de voir des blessures engagent mal Siméon-Denis vers des études médicales. En 1796 il entre à l'École Centrale<sup>2</sup> de la ville où l'on s'aperçoit vite de ses talents en mathématiques. Son professeur, M. Billy, l'encourage dans cette voie et, citant La Fontaine, dit à son propos : « Petit Poisson deviendra grand, Pourvu que Dieu lui prête vie. » On le prépare au concours de la jeune, mais déjà réputée École Polytechnique<sup>3</sup>. Il y est reçu major à dix-sept ans, en 1798. Poisson s'adapte rapidement à son nouvel environnement, travaille avec enthousiasme, tout en

1. Agent chargé de fonctions administratives et judiciaires.

2. Ancêtre des lycées.

3. La fondation de l'École Polytechnique résulte des travaux de la Commission des Travaux Publics du 7 vendémiaire an III (28 septembre 1794). Le but de l'École est de « Former des ingénieurs en tous genres et de rétablir l'enseignement des sciences exactes qui a été suspendu pendant les crises de la Révolution, de donner une haute formation scientifique à des jeunes gens, soit pour être employés par le Gouvernement aux travaux de la République, soit pour reporter dans leurs foyers l'instruction qu'ils auront reçue et y prodiguer les connaissances vraiment utiles. » Le premier cours a lieu le 21 décembre 1794 (1<sup>er</sup> nivôse an III). L'École, d'abord nommée École Centrale des Travaux Publics, prend à partir du 1<sup>er</sup> septembre 1795 le nom d'École Polytechnique. La sélection est remarquable, on compte dans la première promotion (X 1794) : Biot, Malus et Poinot !

se donnant le loisir de se rendre au théâtre et de participer à diverses activités sociales. Poisson, comme ses camarades, bénéficient de cours donnés par des savants prestigieux : l'analyse est enseignée par Lagrange, Monge enseigne la géométrie descriptive avec un don pédagogique hors du commun, la chimie est enseignée par Berthollet qui apporte à l'expérimentation dans ce domaine la rigueur de l'algèbre, etc. Poisson est vite remarqué par Lagrange et Laplace<sup>1</sup>. Joseph-Louis Lagrange, qui enseignait la théorie des fonctions, reconnaît vite les talents de son jeune élève et deviendra son ami. Quant à Pierre Simon Laplace il considère Siméon-Denis presque comme son fils<sup>2</sup>.

En 1800, Poisson publie deux mémoires, l'un sur la méthode d'élimination de Bezout, l'autre sur le nombre des intégrales d'une équation de différences finies. Lacroix et Legendre, qui examinent ces mémoires, recommandent qu'ils soient publiés dans le *Recueil des savants étrangers*, un honneur exceptionnel pour un jeune homme de dix-huit ans. Poisson ne quitte pas l'École Polytechnique dès l'obtention de son diplôme, il y est nommé répétiteur adjoint d'analyse. En 1802, lorsque Joseph Fourier est nommé par Bonaparte préfet de l'Isère, Poisson le remplace en tant que professeur suppléant et il deviendra professeur titulaire en 1806. Durant cette période, Poisson étudie les équations différentielles aux dérivées partielles et traite d'un certain nombre de problèmes physiques tels que le mouvement du pendule dans un milieu résistant ou la théorie du son. Dès 1806, Laplace, Lagrange, Lacroix, Legendre et Biot souhaitent que Poisson entre à l'Académie des sciences dans la section mathématiques, mais aucune place n'est disponible<sup>3</sup>. L'importance de ses travaux lui vaut toutefois la nomination à des postes prestigieux : il est nommé astronome au bureau des Longitudes (1808) et professeur de mécanique rationnelle à la Faculté des sciences de Paris dès sa création (1809). Entre 1808 et 1811, Poisson publie trois articles aux Comptes Rendus de l'Académie des Sciences (CRAS)<sup>4</sup> ainsi que son cours,

1. Laplace a été, entre 1796-1799, examinateur à Polytechnique, il deviendra en 1816 président du conseil d'administration de l'École, qui garde encore l'empreinte de ses réformes.
2. Poisson fera, à l'Institut, l'éloge funèbre de Lagrange en 1813 et celle de Laplace en 1827.
3. Si ce n'est, potentiellement, celle de Bossut, alors âgé de 76 ans. Mais, de fait, ce mathématicien a vécu jusqu'à 83 ans.
4. Ses publications dans les CRAS portent sur l'astronomie et la mécanique : *Sur les inégalités des mouvements moyens des planètes*, (liées aux perturbations des autres planètes), en 1808 ; *Sur le mouvement de rotation de la Terre* (1809) et *Sur la variation des constantes arbitraires dans les questions de mécanique* (1809). C'est dans ce mémoire lu à l'Institut le 16 octobre 1809, et inséré au XV<sup>e</sup> cahier du Journal de l'École Polytechnique, que Poisson a introduit ces célèbres « crochets de Poisson ». Plus d'un siècle plus tard (en août 1925, précisément) Dirac en fera une des pierres angulaires du formalisme quantique. La façon dont Dirac a procédé mérite d'être conté : Un dimanche, au cours de sa promenade hebdomadaire, Dirac poursuit sa réflexion sur la façon de formaliser la mécanique quantique et fait le rapprochement entre le commutateur introduit par Heisenberg et les crochets de Poisson. Il écrit : « C'était un dimanche, je ne pouvais pas aller en bibliothèque, tout ce qu'il me

particulièrement clair, fait à l'École Polytechnique<sup>1</sup>. Étienne Malus, qui avait participé à la campagne d'Égypte jusqu'en 1801, avait contracté la peste. Élu en août 1810 à l'Académie des sciences dans la section physique générale, sa santé décline rapidement et une issue fatale est malheureusement prévisible. Les savants cités plus haut voient là l'occasion de faire entrer Poisson à l'Institut et, pour optimiser ses chances, l'incitent à concourir au Grand Prix d'Électricité de l'Académie dont le thème était : « Déterminer par le calcul et confirmer par l'expérience la façon dont l'électricité est distribuée sur la surface des corps électriques, soit isolés soit en présence l'un de l'autre, par exemple sur la surface de deux sphères électrisées en présence l'une de l'autre. Afin de simplifier le problème, il est seulement demandé l'examen des cas où l'électricité répartie sur chaque surface demeure toujours de même nature. » C'est ainsi que Poisson soumet la première partie de sa solution à l'Académie le 9 mars 1812 dans un papier intitulé *Sur la distribution de l'électricité à la surface des corps conducteurs*. Sa contribution à la physique lui permet d'être élu à l'Académie des sciences le 23 mars 1812, dans la section de physique générale, au fauteuil de Malus, décédé un mois plus tôt, le 24 février. Poisson était également un membre assidu de la Société d'Arcueil, qui était en quelque sorte un second Institut (voir encadré).

#### La Société d'Arcueil

Laplace fonde, avec Claude Berthollet, la Société d'Arcueil dont le but est de : « ramener la vision newtonienne du monde aux dimensions de la physique moléculaire, en appliquant, par les mathématiques, les lois de Newton aux phénomènes tels que la lumière, la chaleur, le son, l'électricité et le magnétisme ». La Société est constituée en société semi-officielle en 1805, sous la protection de l'Empereur, avec des statuts, des séances régulières deux fois par mois, le jeudi après-midi ou le dimanche, présidées alternativement par Berthollet et par Laplace. Les sociétaires réunis autour de Laplace et Berthollet sont Biot, Gay-Lussac, l'explorateur Humboldt, le chimiste Thénard, le botaniste de Candolle, le minéralogiste Collet-Descostils et le fils de Berthollet. Malus, Arago, Poisson et Dulong rejoindront plus tard la Société. La plupart des membres ont alors moins de trente ans ! « On arrivait, raconte un sociétaire, vers 1 heure, on travaillait jusqu'à 4 h 1/2, on jouait aux barres dans le jardin, on dînait et, à 9 heures, tout le monde était parti ».

En 1814 Poisson est fait chevalier de la Légion d'honneur, en 1815 il est nommé examinateur à École Militaire ainsi qu'examinateur de sortie à l'École Polytechnique. En 1817, à trente-six ans, notre mathématicien rencontre une orpheline anglaise d'origine italienne, Nancy de Bardi, qu'il épouse en 1817. Le mariage est heureux et son épouse l'encouragera et l'assistera toute

---

restait à faire était d'attendre impatiemment, pendant la nuit et le début de la matinée, que les bibliothèques ouvrent ». Il vérifie le bien-fondé de ses connaissances sur les crochets de Poisson et indique « tout le travail découla assez directement de cette première idée ». Il racontera plus tard que l'instant de sa révélation a été le plus émouvant de sa vie.

1. Son cours de l'École Polytechnique est publié dans son *Traité de mécanique* (en 2 volumes).

sa vie<sup>1</sup>. Ils auront trois fils, dont l'un fera une belle carrière militaire. Poisson partage son existence entre sa famille et ses recherches mathématiques : il embrasse pratiquement tous les domaines de la physique mathématique<sup>2</sup>. Citons quelques-uns de ses travaux :

- Étude du potentiel interne aux masses de gravitation : il étend les résultats de Laplace (« équation de Laplace ») et introduit (en 1813) dans son mémoire sur l'électrostatique, la fameuse « équation de Poisson ».
- Études sur les intégrales définies et sur les séries de Fourier (1813-1823) qui préparèrent le terrain des recherches de Dirichlet et Riemann.
- Études des vibrations des plaques élastiques. Il introduit le coefficient (dit de Poisson) caractérisant le rapport de l'allongement sur la variation d'épaisseur. Il travaille sur ce sujet avec Sophie Germain, Gabriel Lamé et Claude Navier.
- Travaux sur la conduction de la chaleur (1815) qui feront dire à Fourier : « Poisson à trop de talent pour l'appliquer au travail des autres. Utiliser ce talent pour découvrir ce qui est déjà connu est du gâchis ».
- En optique (1817-1819) une controverse célèbre s'établit entre Poisson (partisan de la théorie corpusculaire de la lumière) et Fresnel (qui introduit la théorie ondulatoire). Poisson, en s'appuyant sur la démarche de Fresnel croit pouvoir démontrer un résultat absurde : selon ce calcul il déduit que « le centre de l'ombre d'un écran circulaire opaque... [doit en]... incidence peu oblique... être aussi éclairé que si l'écran n'existait pas ». Poisson présente cette conséquence comme une condamnation des idées de Fresnel, mais ce dernier, prévenu par Arago fait l'expérience, qui confirme la prévision que Poisson pensait absurde ! La théorie ondulatoire triomphe.
- Études sur le calcul des variations (1833) qui seront, beaucoup plus tard à l'origine des méthodes des éléments finis.

La renommée de Poisson ne cesse de croître : il est fait baron en 1821 (mais il n'utilisera jamais ce titre) ; il est nommé en 1824 officier dans l'ordre de la Légion d'honneur, il devient président de l'Académie des sciences en 1826 puis en 1840 ; il succède à Laplace au bureau des longitudes en 1827 ; il obtient la prestigieuse médaille Copley de la *Royal Society* en 1832 et est fait, la même année Pair de France.

En 1837, Poisson publie son livre intitulé *Recherches sur la probabilité des jugements en matière criminelle et matière civile*, dans lequel apparaît pour la

1. Nancy de Bardi meurt le 26 octobre 1873 à près de 82 ans à Saint-Lô.

2. Il écrira plus de 300 mémoires.

première fois la « distribution de Poisson ». Il introduit également l'expression « loi des grands nombres ».

En 1839, la santé de Poisson décline rapidement. Il retourne à Pithiviers en pensant que l'air de la campagne lui permettra de se rétablir. Il poursuit néanmoins d'anciens travaux, en particulier sur la balistique en prenant en compte les corrections établies par Gaspard Coriolis. Toutefois son état ne s'améliore pas, on le ramène à Sceaux, ville où il meurt le 25 avril 1840.

Arago rapporte que Poisson avait l'habitude de dire « La vie n'est bonne qu'à deux choses : apprendre les mathématiques et les professer ». En 1851, on érige à Pithiviers sa statue qui est inaugurée par son collègue académicien Jean-Baptiste Biot<sup>1</sup>.

## 17.4. Ouvrages de Poisson

Poisson a publié environ 180 mémoires scientifiques dont :

- *Traité de mécanique*, Paris, 1811.
- *Sur la distribution de l'électricité à la surface des corps conducteurs*, Mémoires de la classe des sciences mathématiques de l'Institut, Paris, 1812 et 1813.
- *Théorie mathématique de la chaleur*, Paris, 1835.
- *Sur la probabilité des résultats moyens des observations*, Connaiss. Temps pour 1827, p. 273-302.
- *Suite du Mémoire sur la probabilité du résultat moyen des observations, inséré dans Connaissance des Temps de l'année 1827*, Connaiss. Temps pour 1832, p. 3-22.
- *Mémoire sur la proportion des naissances des filles et des garçons*, CRAS, 9, 239-308, 1830.
- *Recherches sur la probabilité des jugements, principalement en matière criminelle*, CRAS, 11, 473-494, 1835.
- *Note sur la loi des grands nombres*, CRAS, 12, 377-382, 1836.
- *Note sur le calcul des probabilités*, CRAS, 12, 395-400, 1836.
- *Formules relatives aux probabilités qui dépendent de très grands nombres*, CRAS, 12, 603-613, 1836.

---

1. Malheureusement cette statue sera déboulonnée puis fondue par les occupants pendant la seconde guerre mondiale.

- *Recherches sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile, précédées des Règles Générales du Calcul des Probabilités*, Bachelier, 1836.

## 17.5. Références

- P. Costabel, *Poisson*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 15, p. 480, 1970-1980.
- O. B. Sheynin, *S.-D. Poisson's work in probability*, Arch. hist. exact sci., 18, 1978, n° 3, p. 245-300.
- *Siméon-Denis Poisson et la science de son temps* (édité par M. Métivier, P. Costabel, P. Dugac), Palaiseau, École Polytechnique.
- A. Dale, *A History of inverse probability...*, chapitre 7, Springer, 1991.

# Les statistiques d'Ulpien à Quetelet

La naissance de l'arithmétique politique et de la démographie est liée à quelques noms célèbres : Ulpien, John Graunt, Edmund Halley, Leibniz, De Moivre, Simpson, Johann Süssmilch, Willem Kersseboom, Buffon, Nicholas I Bernoulli, Antoine Deparcieux, Leonhard Euler et Adolphe Quetelet.

La démographie est généralement définie comme l'étude quantitative et qualitative des populations humaines et de leurs évolutions à partir de la connaissance de la natalité, la fécondité, la mortalité, la nuptialité et les migrations de ces populations.

On trouve dans divers chapitres du présent ouvrage des commentaires concernant les travaux statistiques de Halley, De Moivre, Simpson, Nicholas Bernoulli et Buffon. Nous dirons donc ici quelques mots d'autres mémoires ou livres relatifs aux mêmes sujets.

## 18.1. Ulpien

Concernant les statistiques avant Graunt, le juriste Romain Ulpien est souvent cité :

Originaire de Tyr en Syrie, Ulpien (Domitius Ulpianus), fut, sous Alexandre Sévère, préfet de l'annone puis, en 222, préfet du prétoire. La fonction judiciaire à Rome vit, au III<sup>e</sup> siècle, le Sénat cesser d'être cour de justice souveraine. La juridiction criminelle sénatoriale de première instance disparut sous Septime Sévère au profit du préfet du prétoire et du préfet de la Ville. Le Sénat ne garda plus que le droit de juridiction sur ses propres membres. Il y eut 7 préfectures, qui constituaient le sommet de la hiérarchie équestre : les deux préfectures des flottes de Misène et de Ravenne, les deux préfectures provinciales et de Mésopotamie, les trois préfectures urbaines des vigiles, de l'annone et du prétoire. L'échelon le plus élevé de la carrière équestre était la préfecture du prétoire. Tout de suite après venait la préfecture d'Égypte, puis la préfecture de l'annone.

Au cours d'un soulèvement de la garde prétorienne, Ulpien fut tué en 228.



La table d'Ulpien a été rédigée vers 220 après J.-C.

Cette table<sup>1</sup>, destinée à la détermination des annuités des rentes, est tirée du *Corpus Iuris Civilis, Digesta*, XXXV, 2, 68. Les *Digesta* ou *Pandectae* étaient incluses dans un ensemble de lois, éditées en 533 après J.-C., consistant en une compilation des écrits de grands juristes Romains, tels Ulpien. Voici cette table :

|      |       |       |       |       |    |     |    |       |       |     |
|------|-------|-------|-------|-------|----|-----|----|-------|-------|-----|
| 0-19 | 20-24 | 25-29 | 30-34 | 35-39 | 40 | ... | 49 | 50-54 | 55-59 | 60- |
| 30   | 28    | 25    | 22    | 20    | 19 | ... | 10 | 9     | 7     | 5   |

La première ligne correspond à l'âge de la personne désignée pour recevoir la rente et la seconde, selon Nicholas I Bernoulli, l'espérance de vie de ladite personne.

## 18.2. Graunt (1620-1674)

Le livre de Graunt *Natural and Political Observations made upon the bills of Mortality*, paru à Londres en 1662, est le premier livre de démographie mathématique, c'est-à-dire appliquant des calculs aux données sur la population. Graunt a en effet entrepris dans son ouvrage une étude statistique des données de mortalité de la ville de Londres qui donnent les décès par cause et non par âge. Graunt a cherché à assigner une durée moyenne à la vie humaine et à déterminer la vitesse d'accroissement de la population en cherchant des régularités éventuelles dans les naissances et les décès, certains savants de l'époque pensant d'ailleurs que cette régularité impliquait une intention divine. Graunt, en plus de ses *Observations Naturelles et Politiques* fait aussi des observations d'ordre épidémiologique et économique. Au total son livre contient 106 propositions, observations ou questions qui découlent de son étude.

Graunt a effectué une estimation raisonnée de la population et a été le premier à mettre en évidence la constance du rapport des naissances des garçons et des filles. Arbuthnot, Nicholas I Bernoulli, Daniel Bernoulli, Laplace, Poisson et bien d'autres reprendront l'étude probabiliste de ce phénomène démographique.

Le *Journal des sçavans* du lundi 2 Août 1666 a publié une recension du livre de Graunt, rédigée par le Sr G.P. (c'est-à-dire l'Abbé Jean Gallois, qui fut éditeur et rédacteur du journal de 1666 à 1674) dont nous donnons, ci-dessous, quelques extraits :

« C'est une chose particulière aux Anglais de faire des *Billets de la mortalité*, c'est-à-dire des listes qui contiennent combien il naît de personnes chaque semaine, combien il en meurt, et quelle est la cause de leur mort. Les maladies contagieuses qui ont

1. Greenwood, M. *Jal stat. Soc.*, 103, 1940, p. 246-248.

fait de grands ravages en Angleterre ont donné lieu à ces billets. Car la curiosité qu'on a eue de savoir si la peste diminuait ou augmentait, a fait qu'on a d'abord tenu registre de tous ceux qui mouraient de cette maladie : Ensuite après que la contagion a cessé, on a continué à marquer le nombre de ceux qui naissent, et à spécifier leur maladie. Enfin pour voir la quantité de ceux qui naissent, aussi bien que de ceux qui meurent, on a ajouté le nombre des baptêmes. [...]

L'auteur de ce livre ayant examiné tous ces Billets de mortalité, y a fait plusieurs réflexions curieuses, dont voici les principales :

1. Il compare les maladies entre elles et il remarque que sur deux cent vingt-neuf mille personnes qui sont mortes à Londres pendant vingt ans, les convulsions, les vers, les dents et d'autres semblables maladies en ont emporté 71 124. Il en est mort 12 210 de la petite vérole et de la rougeole. De fièvre, il n'est est pas mort plus d'un entre quarante. Les maladies chroniques, comme l'hydropisie, phtisie ou maladie de poumon etc. en ont fait mourir environ 70. Les maladies externes, par exemple, les blessures, les Cancres etc. emportent beaucoup moins de monde, de 60 personnes il n'y en a qu'une qui soit morte de ces maladies. La peste fait d'étranges désordres de temps en temps, car en 1603 et en 1625, elle emporta presque la cinquième partie des Habitants de Londres. Cet auteur ajoute qu'il se fait fort peu de meurtres dans cette ville, et que pour deux cents personnes qui meurent de maladie, à peine y en a-t-il une qui ait été tuée. Qu'il ne meurt presque personne de faim, et que de quatre cents il n'y en a pas un qui meurt de disette. Que de cent femmes, il n'en meurt pas plus d'une en couche, et qu'entre 200, il n'y en a presque qu'une qui meurt en travail. Il examine ensuite quelles sont les maladies qui augmentent ou qui diminuent, et il remarque que les Rikets, qui sont une nouvelle maladie particulière aux Anglais, augmentent tellement, que pour 14 personnes qui en moururent en 1634, il en est mort jusqu'à 560, en 1660. Au contraire, il semble que la Pierre diminue. Pour la Goutte elle demeure toujours en même état, et de mille, il n'en meurt pas plus d'un de cette maladie, quoiqu'il se trouve beaucoup de personnes qui en soient tourmentées.

2. Il compare les divers âges de la vie entr'eux, et il observe qu'il y a vingt fois plus d'enfants qui viennent à terme, que de ceux qui viennent morts ou qui meurent en naissant. Que de cent enfants qui naissent en même temps, il n'en reste, six ans après que soixante-quatre ; au bout de 16 ans il n'y en a plus que quarante ; au bout de 26 ans il ne s'en trouve que vingt-cinq ; au bout de 36 ans il n'en demeure que seize ; au bout de 46 ans il n'y en a plus que dix ; au bout de 56 ans il n'en reste que six ; au bout de 66 ans il n'en demeure que trois, et au bout de 76 ans, ces trois sont réduits à un. Enfin au bout de 80 ans il ne reste plus personne. »

Le tableau qui suit résume ces résultats :

| <i>AGE (ans)</i> | <i>DECES</i> | <i>SURVIVANTS</i> |
|------------------|--------------|-------------------|
| 0                | 36           | 100               |
| 6                | 24           | 64                |
| 16               | 15           | 40                |
| 26               | 9            | 25                |
| 36               | 6            | 16                |
| 46               | 4            | 10                |
| 56               | 3            | 6                 |
| 66               | 2            | 3                 |
| 76               | 1            | 1                 |
| 80               |              | 0                 |

L'article du *Journal des Sçavans* poursuit :

« 3. Il compare les Hommes avec les Femmes, et il considère en quelle proportion ils naissent et ils meurent. Il dit que dans Londres il naît quatorze Garçons pour treize Filles et dans la campagne quinze Garçons pour quatorze Filles [...].

4. Il compare les Habitants de la ville avec ceux de la campagne. Il remarque que bien qu'à Londres on meurt moins par secousses qu'à la campagne, parce que la fumée et les vapeurs de cette ville la garantissent en quelque manière des mauvaises impressions de l'air ; néanmoins il meurt plus de monde à proportion dans Londres ; car de 32 personnes il y en meurt pour le moins une, et dans la campagne il n'en meurt pas plus d'une de 50.

5. Il compare le nombre de ceux qui naissent et de ceux qui meurent. Il avoue qu'il est difficile d'en rien déterminer de certain, parce qu'on ne peut connaître le nombre de ceux qui sont nés que par les registres des baptêmes ; et cependant de ceux qui sont nés à Londres depuis l'année 1650 jusqu'en 1660, il n'y en a pas, à ce qu'il dit, la moitié qui ait été baptisés, plusieurs personnes ayant été pour lors dans cette erreur que le baptême n'était pas absolument nécessaire aux enfants. Néanmoins, il dit qu'autant qu'il peut conjecturer, il meurt à Londres environ 12 personnes pour 11 qui naissent, mais qu'à la campagne il y a ordinairement 63 personnes qui naissent pour 52 qui meurent. [...]

6. Il suppose par diverses manières combien il y a de monde dans Londres et dans toute l'Angleterre. Il dit qu'il se fait à Londres par an 15 000 enterrements, entre lesquels il s'en trouve 5 000 d'enfants morts avant d'être nés ou aussitôt qu'ils sont nés, ou morts des dents. Qu'il s'y fait 12 000 baptêmes, et partant qu'il y a dans Londres environ 384 000 personnes tant Hommes que Femmes qui font en tout 48 000 familles. De plus il dit que dans le reste de l'Angleterre il y a quatorze fois plus de monde que dans Londres. Que dans l'Angleterre et dans le pays de Galles

il y a environ 25 millions d'arpents de terre à 16 pieds la perche. Qu'il s'y trouve à peu près 10 000 paroisses, dans chacune desquelles, l'une portant l'autre, il y a 600 personnes, d'où il s'ensuit qu'il y a dans le Royaume d'Angleterre environ six millions de personnes. Il ajoute que bien qu'il meure dans Londres plus de monde qu'il n'en naît, néanmoins cette ville augmente au lieu de diminuer, parce qu'il y vient tous les ans environ 6 000 personnes de la campagne. Par ce calcul qu'il fait, il trouve que le peuple de Londres en 20 ans est augmenté de 7 à 12 et en 40 ans de 23 à 52.

7. Il fait plusieurs autres remarques curieuses : Par exemple que de 7 malades il n'en meurt ordinairement qu'un. Que de chaque mariage il naît environ quatre enfants. Qu'il y a dans chaque famille de Londres environ 8 personnes etc. »

Le premier problème qu'ont cherché à résoudre les arithméticiens politiques était donc de mesurer une population et sa croissance en l'absence de recensements. Dans cette démarche, la ville de Londres constituait un excellent terrain, grâce à sa collection de Bulletins de mortalité.

L'ouvrage posthume de William Petty<sup>1</sup> (1623-1687) *Political Arithmetic* comporte non seulement l'évaluation de la population et du rapport des nombres d'hommes et de femmes, mais également celle des biens et marchandises. Cet ouvrage laisse penser que Graunt et Petty se sont mutuellement influencés et de longs débats ont pris naissance pour savoir lequel de ces deux auteurs devait être crédité de la création des études statistiques. Divers travaux, dont ceux de Kesserboom et de Süssmilch comparèrent aussi les populations de Londres et de Paris.

### 18.3. La correspondance des frères Huygens

Cette correspondance, qui n'a été publiée qu'en 1920, est relative aux rentes viagères, à la vie moyenne et à la vie probable des humains a été initiée entre les frères Huygens par une lettre<sup>2</sup> de Lodewijk à Christiaan en date du 22 août 1669 dont voici un extrait :

« À propos d'âge, j'ai fait une table ces jours passés du temps qu'il reste à vivre, à des personnes de toute sorte d'âge. C'est une conséquence que j'ai tirée de cette table du livre anglais of the bills of mortality, de laquelle je vous envoie ici une copie, afin que vous preniez la peine de faire un peu les mêmes supputations, et que nous puissions voir comment nos calculs s'accorderont. J'avoue que j'ai eu assez de peine

1. *Political Arithmetick or a Discourse Concerning The Extent and Value of Lands, People, Buildings ; Husbandry, Manufacture, Commerce, Fishery, Artizans, Seamen, Soldiers ; Publick Revenues, Interest, Taxes, Superlucration, Registries, Banks Valuation of Men, Increasing of Seamen, of Militia's, Harbours, Situation, Shipping, Power at Sea, &c. As the same relates to every Country in general, but more particularly to the Territories of His Majesty of Great Britain, and his Neighbours of Holland, Zealand, and France. By Sir William Petty, Late Fellow of the Royal Society. London, Printed for Robert Clavel at the Peacock, and Hen. Mortlock at the Phoenix in St. Paul's Church-yard. 1690.*
2. CH. Huygens, *Œuvres complètes*, tome VI, page 482, La Haye, Martinus Nijhoff, 1895.

d'en venir à bout, mais à vous il n'en sera pas de même, et les conséquences qui en résultent sont fort plaisantes et peuvent même être utiles pour les constitutions des rentes à vie. La question est jusqu'à quel âge doit vivre naturellement un enfant aussitôt qu'il est conçu. Puis un enfant de 6 ans, puis un de 16 ans, de 26 etc. Si vous y trouvez de la difficulté ou trop d'embarras, je m'offre à vous faire part de ma méthode, qui est assurée, par la première occasion.

Adieu.

Selon mon calcul vous vivrez environ jusqu'à l'âge de 56 ans et demi. Et moi jusqu'à 55. »

La réponse détaillée de Christiaan est contenue dans ses lettres à Lodewijk du 21 et 28 novembre 1669. Dans cette dernière, il écrit :

« Ce sont donc deux choses différentes que l'espérance ou la valeur de l'âge futur d'une personne et l'âge auquel il y a égale apparence qu'il parviendra ou ne parviendra pas. Le premier est pour régler les rentes à vie, et l'autre pour les gageures. Je verrai si vous avez fait la même distinction. »<sup>1</sup>

J.-M. Rohrbasser et J. Véron<sup>2</sup> ont publié un mémoire dans lequel ils soulignent que les deux frères Huygens divergent sur l'approche adéquate du problème et que Lodewijk a en tête l'espérance de vie ou la vie moyenne, tandis que Christiaan cherche les probabilités d'atteindre différents âges. R. Pressat<sup>3</sup> a fait remarquer, par ailleurs que Christian « en calculant la survie du dernier décédé d'un groupe de deux personnes a fait apparaître les notions de probabilité conditionnelle et d'espérance de vie conditionnelle. »

## **18.4. Le calcul politique de Leibniz (1646-1716)**

Le premier ouvrage scientifique publié par Leibniz a été la *Dissertatio de Arte Combinatoria*..., un texte de quatre-vingts pages qui constituait son mémoire de thèse de philosophie et qui parut à Leipzig en 1666. Leibniz avait suivi à Iéna, en 1664, les cours du mathématicien Ehrard Weigel, et il avait alors conçu l'idée d'une méthode philosophique de combinaisons analogue à la méthode mathématique, idée qu'il exposa deux ans plus tard dans sa thèse. Leibniz désirait élaborer un alphabet des pensées humaines. Dans sa théorie, tous les concepts ne sont que des combinaisons d'un certain nombre d'autres concepts plus simples, tout comme les mots sont la combinaison de lettres. Toutes les combinaisons de concepts peuvent être, à leur tour, décomposées en idées simples, de telle sorte que ce nouvel alphabet des pensées pourrait servir de logique des inventions, en complément de la logique des démons-

- 
1. Ch. Huygens, *Œuvres complètes*, tome VI, p. 537-538, La Haye, Martinus Nijhoff, 1895.
  2. J.-M. Rohrbasser et J. Véron, *Lodewijk et Christiaan Huygens : la distinction entre vie moyenne et vie probable*, Math et Sci. hum, 38<sup>e</sup> année, n° 149, 2000, p. 7-21.
  3. R. Pressat, *Christian Huygens et la table de mortalité de Graunt*, Math et Sci. hum, 39<sup>e</sup> année, n° 153, 2001, p. 29-36.

DISSERTATIO  
De  
ARTE COMBINATORIA,

In qua

Ex Arithmeticae fundamentis Complicationum ac Transpositionum  
Doctrina novis præceptis extruitur; & usus ambarum per uni-  
versum scientiarum orbem ostenditur; nova etiam  
Artis Meditandi,

Sed

Logicae Inventionis semina  
sparguntur.

*Præfixa est Syncreticus Tractatus, & additamenti loco  
Demonstratio*

EXISTENTIAE DEI,  
ad Mathematicar. certitudi-  
nem exacta.

AUCTORE

GOTTFREDO GUILLIELMO  
LEIBNÜZIO Lipsensi,  
Phil. Magist. & J. U. Baccal.

---

L I P S I Æ,

APUD JOH. SIMON. FICKHUM ET JOH.  
POLYCARP. SEIBOLDUM

*in Platea Nigra.*

Literis SPÖRELIANIS.

A. M. DC. LXXI.

trations déjà connue. Leibniz utilise dans son traité les permutations qu'il appelle *variationes ordinis*, et les combinaisons. Les exemples qu'il donne sont relatifs, entre autres, au Droit et à la Musique. Mais les applications philosophiques lui semblent plus importantes, si bien qu'il termine son traité par ce qu'il croit être une démonstration mathématique de l'existence de Dieu.

Ce que Leibniz appelait, par ailleurs, *calcul politique* était équivalent à l'arithmétique politique de Petty. Il concevait l'ordre de la mortalité comme un processus aléatoire. Dans un mémoire écrit vers 1682, intitulé *Essay de quelques raisonnements nouveaux sur la vie humaine et sur le nombre des hommes*<sup>1</sup>, Leibniz, après avoir rappelé son travail antérieur sur les rentes à vie, fait les suppositions suivantes : 81 enfants nouvellement nés, mourront uniformément, soit un par an dans les 81 années suivantes ; tous les hommes ont la même vitalité ; la limite de la vie humaine est de 81 ans (ou 70 ans dans d'autres écrits). Pour Leibniz, ces faits sont dus à la Providence.

Leibniz basait donc ses considérations, à cette époque, non sur l'expérience et sur des tables de mortalité, mais une modélisation mathématique simpliste et gratuite, ignorant le calcul des probabilités *a posteriori*. Il a abordé cependant sur ces bases cinquante-six diverses questions<sup>2</sup> relatives à l'âge auquel on est le plus atteint par les maladies, le nombre d'enfants qui atteignent l'âge adulte, la durée moyenne de la vie humaine, l'accroissement et le décroissement du genre humain ou la valeur de rentes viagères. Le résultat est, par exemple, une longueur moyenne de la vie humaine de quarante ans et une population stationnaire, le nombre de naissances étant, selon Leibniz, égal au nombre de décès.

## 18.5. Arbuthnot (1667-1735)

John Arbuthnot a été à l'origine du premier livre consacré aux probabilités publié en Anglais en 1692. Ce livre intitulé *On the laws of Chance, A method of calculation of the hazards of Game Plainly demonstrated, And applied to Games as present most in Use*<sup>3</sup> comportait une traduction du *De ratiociniis in ludo aleae* de Huygens (occupant les pages 5 à 48) et des applications à divers autres jeux.

Arbuthnot est considéré, par ailleurs, comme l'auteur d'un des premiers mémoires à l'origine des statistiques mathématiques.

1. Reproduit dans J.-M. Rohrbasser et J. Véron, *Leibniz et les raisonnements sur la vie humaine*, Paris, Ined, 2001.
2. *Hauptschriften zur Versicherungs- und Finanzmathematik*, éd. par E. Knobloch et J.-M. Graf von der Schulenburg, Berlin, Akademie Verlag, 2000.
3. London: Printed by Benj. Motte, and sold by Randall Taylor, 1692.

OF THE  
Laws of Chance,  
OR, A  
METHOD  
OF  
*Calculation of the Hazards*  
OF  
GAME,  
Plainly demonstrated,  
And applied to GAMES at  
present most in Use,  
Which may be easily extended to the most  
intricate Cases of Chance imaginable.

---

L O N D O N:

Printed by *Benj. Motte*, and sold by  
*Randall Taylor* near *Stationers-Hall*, 1692.



Nous traduisons comme suit, le début du célèbre papier<sup>1</sup>, dont le titre est *An Argument for Divine Providence, taken from the constant regularity observ'd in the birth of both sexes* [Un argument de preuve de la divine providence résultat de la régularité constante observée dans la naissance des deux sexes] :

« Parmi les très nombreuses empreintes de la Divine Providence que l'on peut trouver dans les Œuvres de la Nature, il y en a une, remarquable, consistant en l'équilibre exact maintenu entre les hommes et les femmes ; car, par ce moyen, il est fait en sorte que l'espèce ne disparaîtra pas, puisque tout mâle peut avoir une femelle, et d'un âge approprié. Cette égalité de mâles et de femelles n'est pas un effet du hasard mais résulte de la Providence Divine œuvrant pour le Bien, ce que je vais démontrer :

Soit une pièce à deux faces, croix et pile que nous appellerons *Met F*. Pour trouver les chances de trouver n'importe quel nombre déterminé de telles pièces, prenons le binôme *M + Félevé* à la puissance représentant le nombre de pièces donné ; les coefficients des termes du binôme seront les chances cherchées. Par exemple, pour deux pièces à deux faces *M* et *F*, les chances sont  $M^2 + 2MF + F^2$ , c'est-à-dire une chance pour un double *M*, une chance pour un double *F*, et deux chances pour un *M* et un *F*. Pour quatre pièces, les chances sont  $M^4 + 4M^3F + 6M^2F^2 + 4MF^3 + F^4$  ; c'est-à-dire une chance pour *M* quadruple, une pour *F* quadruple, quatre pour *M* triple et *F* unique, quatre pour *F* triple et *M* unique, six pour *F* double et *M* double ; enfin, universellement, si le nombre de pièces est *n*, toutes les chances seront exprimées dans la série

$$M^n + \frac{n}{1} M^{n-1} F + \frac{n(n-1)}{1.2} M^{n-2} F^2 + \frac{n(n-1)(n-2)}{1.2.3} M^{n-3} F^3 + \dots etc.$$

Il apparaît que lorsque le nombre de pièces est pair, il y a autant de *M* que de *F*, dans le terme médian de la série, tandis que les autres termes ont plus de *M* ou plus de *F*. [...]

Arbuthnot observe ensuite que sur un très grand nombre d'observations (années de statistiques dans le cas des naissances), c'est-à-dire un très grand nombre de pièces *n*, il est improbable d'avoir le même nombre de *M* et de *F* et tout aussi improbable que *M* soit nettement plus grand que *F*. En effet, pour un nombre *n* pair, la probabilité d'obtenir le terme médian est 1/2 pour *n* = 2 ; elle est de 3/8 pour *n* = 3 et de 259/1 024 pour *n* = 10, etc. Arbuthnot ayant à sa disposition les statistiques des naissances d'enfants baptisés à Londres durant 82 années, et observant que le rapport du nombre de naissances de garçons par rapport aux filles est légèrement supérieur à 1 chaque année et quasi constant d'année en année, en déduit que la probabilité de telles coïncidences étant de  $2^{-82}$ , le fait ne peut être dû au hasard mais bien, selon lui, à la Providence Divine qui compense la plus faible espérance de vie des garçons par rapport aux filles, par un nombre de naissances plus élevé.

1. Phil. Trans., London, 27 (1710-1712) p. 186-190.

Die göttliche  
**Ordnung**  
in den Veränderungen  
des menschlichen  
Geschlechts,

aus  
der Geburt, Tod, und  
Fortpflanzung  
desselben

verfaßt von  
**Johann Peter Süßmilch,**  
Prediger beym hochlöblichen Altschleinschen  
Regiment.

---

Nebst einer Vorrede  
Herrn Christian Wolffens

---

Berlin,  
Zu finden bey J. C. Spener.  
1741.

Le raisonnement d'Arbuthnot est, en fait, le premier exemple d'inférence statistique. Poisson établira cent vingt ans plus tard la loi relative aux événements rares (cf. notre chapitre 17).

## 18.6. Süssmilch, le père de la démographie allemande (1707-1767)

Johann Peter Süssmilch, pasteur luthérien de Berlin a publié en 1741 un véritable ouvrage de démographie intitulé *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts aus der Geburt, Tod und Fortpflanzung des selben erwiesen...*<sup>1</sup>.

Cependant, ce livre constitue pour l'essentiel un traité de théologie. L'auteur veut par ces calculs déduits de données démographiques établir qu'un ordre et des régularités existent dans les faits humains et que cet ordre résulte du rôle et de l'omniprésence de la Providence divine.

## 18.7. Euler (1707-1783)

Euler a publié en 1767 dans les Mémoires de l'Académie des sciences de Berlin<sup>2</sup>, un mémoire intitulé *Recherches générales sur la mortalité et la multiplication du genre humain*<sup>3</sup>. Dans ce travail, Euler souligne que toutes les questions relatives à la mortalité et à la multiplication du genre humain sont abordées à partir de tables de mortalité dont les données dépendent de divers paramètres spécifiques. Il se propose donc « de traiter ici en général la plupart de ces questions sans se borner aux résultats que les registres d'un certain endroit fournissent et ensuite il sera aisé de faire l'application à chaque endroit qu'on voudra. »

« J'observe d'abord », écrit Euler, « que toutes ces questions prises en général dépendent de deux hypothèses [...]. Je nommerai la première l'hypothèse de la mortalité par laquelle on détermine, combien d'un certain nombre d'hommes, qui sont nés à la fois, seront encore en vie après chaque nombre d'années écoulées. Ici la considération de la multiplication n'entre point du tout en compte, et partant il faut constituer la seconde hypothèse, que je nommerai celle de la multiplication, et par laquelle je marque de combien le nombre de tous les hommes est augmenté ou diminué pendant le cours d'un an. Cette hypothèse dépend donc de la quantité des mariages et de la fécondité, pendant que la première est fondée sur la vitalité ou le pouvoir de vivre, qui est propre aux hommes. »

Euler aborde six problèmes distincts :

1. Cf. Süssmilch (Johann Peter), *L'ordre divin dans les changements de l'espèce humaine, démontré par la naissance, la mort et la propagation de celle-ci* (texte intégral de l'édition de 1741, traduit et annoté par Rohrbasser Jean-Marc), 1998, Paris, INED.
2. Tome XVI, 1767, p. 144-164.
3. Eneström 334.

## RECHERCHES GÉNÉRALES

SUR

### LA MORTALITÉ ET LA MULTIPLICATION DU GENRE HUMAIN.

PAR M. EULER.

I.

**L**es registres des naissances & des morts à chaque âge, qu'on publie en plusieurs endroits tous les ans, fournissent tant de questions différentes sur la mortalité & la multiplication du genre humain, qu'il seroit trop long de les rapporter toutes. Or les unes dépendent pour la plupart en sorte des autres, qu'en ayant développé une ou deux, toutes les autres se trouvent pareillement déterminées. Comme les solutions doivent être tirées des registres mentionnés, il est à remarquer, que ces registres diffèrent beaucoup selon la diversité des villes, villages & provinces, où ils ont été dressés: & par la même raison les solutions de toutes ces questions se trouvent fort différentes selon les registres sur lesquels elles sont fondées. C'est pourquoi je me propose de traiter ici en général la plupart de ces questions sans me borner aux résultats que les registres d'un certain endroit fournissent: & ensuite il sera aisé de faire l'application à chaque endroit qu'on voudra.

2. Or j'observe d'abord, que toutes ces questions prises en général dépendent de deux hypothèses; lesquelles étant bien fixées il est aisé d'en tirer la solution de toutes. Je nommerai la première l'hypothèse de la mortalité par laquelle on détermine, combien d'un certain nombre d'hommes, qui sont nés à la fois, seront encore en vie après chaque nombre d'années écoulées. Ici la considération de la multiplication n'entre point du tout en compte, & partant il faut conti-

1. Un certain nombre d'hommes dont tous soient du même âge étant donné, trouver combien ils seront probablement après  $n$  ans.
2. Trouver la probabilité qu'un homme d'un certain âge soit encore en vie après un certain nombre d'années.
3. On demande la probabilité qu'un homme d'un certain âge mourra dans le cours d'une année donnée.
4. Trouver le terme, auquel un homme d'un âge donné peut espérer de parvenir, de sorte qu'il est également probable qu'il meure avant ce terme qu'après.
5. Déterminer les rentes viagères, qu'il est juste de payer à des hommes d'un âge quelconque tous les ans, jusqu'à leur mort, pour une somme qu'ils auront avancée d'abord.
6. Quand les intéressés sont des enfants nouvellement nés, et que le paiement des rentes viagères ne doit commencer que lorsqu'ils auront atteint un certain âge, déterminer la quantité de ces rentes.

Il en donne les solutions sous forme d'équations établies à partir d'une table de mortalité virtuelle dans laquelle il définit pour une population  $N$  et pour chaque âge compris entre  $n$  et  $(n + 1)$  années, le nombre probable de survivants  $(n)N$ , le nombre de décès annuel pour cette tranche d'âge étant donc  $(n)N - (n + 1)N$ .

Euler introduit ensuite les hypothèses de fécondité en plus de celles de mortalité et résout les problèmes suivants :

1. Les hypothèses de mortalité et de fécondité étant données, si on connaît le nombre de tous les vivants, trouver combien il y en aura de chaque âge.
2. Les mêmes choses étant données, trouver le nombre des hommes qui mourront dans un an.
3. Connaissant tant le nombre des naissances que des enterrements qui arrivent dans le cours d'une année, trouver le nombre de tous les vivants et leur augmentation annuelle pour une hypothèse de mortalité donnée.
4. Le nombre des naissances et des enterrements d'une année étant donnés, trouver combien de chaque âge il y aura parmi les morts.
5. Connaissant le nombre de tous les vivants, de même que le nombre des naissances avec les nombres des morts de chaque âge durant le cours d'une année, trouver la loi de mortalité.

Il est à noter que pour clarifier par des exemples le calcul des rentes, Euler indique que ce calcul exige de nombreuses tables résultant d'observations et il cite et utilise des tables publiées par Kersseboom<sup>1</sup>.

## 18.8. Deparcieux (1703-1768)

Antoine Deparcieux, de la Société Royale des Sciences de Montpellier, a entrepris en 1746 de rectifier les tables de mortalité existantes, dont celle établie par *Halley* dans le but de déterminer les rentes viagères. Il publie aussi, de fait, la première table de mortalité française. Il indique dans l'avertissement de son ouvrage *Essai sur les probabilités de la durée de la vie humaine*<sup>2</sup> :

« La détermination des Rentes viagères, qui sont le sujet de la troisième partie de cet ouvrage, dépend de deux principes ; 1°, de l'intérêt que l'argent doit rapporter ; 2° du plus ou moins de probabilités qu'il y a que la personne qui constitue la rente vivra jusqu'à tel ou tel autre âge. Cette seconde partie ne peut être établie que par des observations faites sur l'ordre de mortalité du genre humain. »

*Deparcieux* se base, pour établir ses propres tables de mortalité, sur l'examen des registres d'une tontine<sup>3</sup> et des registres de décès des couvents parisiens sur une longue période (1607-1745). Sa table fut longtemps en usage auprès des assureurs français, pour le calcul des rentes viagères.

Le problème général des annuités à divers âges et celui des intérêts composés sont abordés dans les 34 premières pages du livre.

La seconde partie commence par décrire les caractéristiques des diverses tables de mortalité déjà disponibles :

Après avoir indiqué que William Petty avait cherché à établir une table de mortalité à partir des registres de Londres, Deparcieux estime que les flux permanents de l'immigration et de l'émigration dans cette grande ville rendaient les données peu précises, ce qu'avait déjà signalé Halley. Concernant les tables de ce dernier, Deparcieux les critique : Halley, selon lui, « n'a rien dit des tontines, ni des rentes qui sont en partie tontines, et en partie viagères simples, ni de quelques autres manières de faire des rentes à vie. D'ailleurs son mémoire est écrit en Anglais, et il n'est connu en France que

1. Kersseboom, Willem, *Essais d'arithmétique politique contenant trois traités sur la population de la province de Hollande et Frise occidentale, la durée de survie des veuves, la durée des mariages, la relation entre la population et le nombre de naissances, le nombre de couples...* (1742), réédition Paris, Institut national d'études démographiques, 1970.

2. Paris, Guérin, 1746.

3. Le mot *tontine* vient du nom d'un banquier italien, Lorenzo Tonti, qui proposa ce système à Mazarin : Les souscripteurs versent une somme dans un fonds et reçoivent les dividendes de la somme investie. Quand un souscripteur décède, sa part est répartie entre les survivants. Le dernier survivant récupère le capital.

de quelques savants ; et il est écrit d'une manière si concise, que quand on le traduirait en Français, peu de gens pourraient l'entendre. »

Puis Deparcieux commente la table, établie par Smart, et utilisée par Simpson dans son ouvrage *Annuities and Reversions* de 1742. Simpson aurait corrigé cette table dressée d'après les registres de Londres, pour tenir compte, écrit Deparcieux, « de la quantité prodigieuse d'étrangers qui vont s'y établir et mourir. » Mais Simpson n'aurait pas indiqué sa méthode de correction, rendant ainsi douteux l'usage de sa table.

Deparcieux ayant compilé les tables de De Moivre, Simpson et Kersseboom, a réuni dans une table finale les « Comparaisons des différentes tables qui ont été faites pour montrer l'ordre de mortalité du genre humain, ou les probabilités que les personnes de chaque âge ont de vivre jusqu'à un autre âge. » On y trouve, outre les résultats de De Moivre (Halley) pour les âges compris entre 1 et 84 ans, ceux de Simpson pour les âges compris entre 6 et 75 ans et ceux de Kersseboom de la naissance à 70 ans par pas de cinq ans, les siens propres relatifs aux tontines de 1689 et 1696 et à divers ordres religieux français.

L'ouvrage de Deparcieux a été complété en 1760 par un Supplément qui comporte essentiellement des tables de prix d'annuités différées pour trois taux d'intérêt différents.

## 18.9. Lambert (1728-1777)

Dans le troisième volume de ses *Beyträge zum Gebrauche der Mathematik und deren Anwendung*<sup>1</sup> (Contributions à l'étude des mathématiques et de leurs applications) publié en 1772, Lambert a inclus un chapitre sur la démographie.

Lambert cherche une mathématisation de la loi de mortalité, en niant l'hypothèse de Leibniz selon laquelle il y a équiprobabilité de décéder à tout âge.

L'équation représentant la loi de mortalité correspondant aux tables de Londres de 1728-1757 est selon Lambert :

$$y = 10.000 \left( \frac{96 - x}{96} \right)^2 - 6.176 \left( e^{-\frac{x}{13,682}} - e^{-\frac{x}{2,43114}} \right),$$

$y$  étant le nombre de survivants, pour une population initiale de 10 000 personnes et  $x$  l'âge en années à partir de la naissance. Cette équation a été citée par Lacroix dès 1822.<sup>2</sup>

1. *Beyträge zum Gebrauche der Mathematik und deren Anwendung*, Berlin, 3 volumes, 1765, 1770, 1772.
2. S. F. Lacroix, *Traité élémentaire du calcul des probabilités*, 2<sup>e</sup> édition, Paris, Bachelier, 1822 page 196.

Lambert indique que la courbe est composée de l'addition d'une parabole et de deux droites, la parabole pouvant indiquer que l'espèce humaine s'éteint à la manière d'un récipient cylindrique qui se vide de son eau, les deux droites étant des « lignes logistiques »<sup>1</sup>.

## 18.10. Quetelet (1796-1874)

Adolphe Quetelet a joué, au milieu du XIX<sup>e</sup> siècle, un rôle majeur dans l'organisation de la collecte des données démographiques et dans les méthodes d'analyse de ces données à l'aide d'outils mathématiques.

Le but de Quetelet était l'étude des lois qui régissent l'homme du point de vue physique, intellectuel et moral. Dans son travail intitulé *Sur l'homme et le développement de ses facultés ; Essai d'une physique sociale*, publié en 1835, il affirme :

« L'homme naît, se développe, et meurt d'après certaines lois. »

Puis Quetelet décrit ainsi son but :

« L'objet de cet Ouvrage est d'étudier, dans leurs effets, les causes, soit naturelles, soit perturbatrices qui agissent sur le développement de l'homme ; de chercher à mesurer l'influence de ces causes, et le mode d'après lequel elles se modifient mutuellement.

Je n'ai point en vue de faire une théorie de l'homme, mais seulement de constater les faits et les phénomènes qui le concernent, et d'essayer de saisir, par l'observation, les lois qui lient ces phénomènes ensemble.

L'homme que je considère ici est, dans la société, l'analogue du centre de gravité dans les corps ; il est la moyenne autour de laquelle oscillent les élémens sociaux : ce sera, si l'on veut, un être fictif pour qui toutes les choses se passeront conformément aux résultats moyens obtenus pour la société. Si l'on cherche à établir, en quelque sorte, les bases d'une *physique sociale*, c'est lui qu'on doit considérer, sans s'arrêter aux cas particuliers ni aux anomalies, et sans rechercher si tel individu peut prendre un développement plus ou moins grand dans l'une de ses facultés. »

L'homme moyen d'une population est donc, selon Quetelet, un individu dont les caractéristiques physiologiques sont chacune égale à la moyenne des caractéristiques physiologiques des autres individus de cette population.

C'est dans les *Lettres à S.A.R. le duc régnant de Saxe-Cobourg et Gotha sur la théorie des probabilités appliquée aux sciences morales et politiques*<sup>2</sup> qui font suite à une série de leçons données en 1836, à la demande du Roi Léopold 1<sup>er</sup>, à deux de ses neveux, Ernest et Albert de Saxe-Cobourg et Gotha, que Quetelet expose l'usage de la loi de Gauss (qu'il appelle loi des possibilités) pour ses études

1. Cf. M. Barbut, J. M. Rohrbasser, J. Véron, *Math et Sci. hum.*, 43<sup>e</sup> année, n° 171, 2005, p. 51-85.

2. Bruxelles, Hayez, 1846.



# **SUR L'HOMME**

**ET LE**

**DÉVELOPPEMENT DE SES FACULTÉS,**

**OU**

**ESSAI DE PHYSIQUE SOCIALE;**

**PAR A. QUETELET,**

Secrétaire perpétuel de l'Académie royale de Bruxelles, Correspondant de  
l'Institut de France, de la Société royale astronomique de Londres, des  
Académies royales de Berlin, de Turin, etc.

---

Appliquons aux sciences politiques et morales la méthode fondée sur l'observation et sur le calcul, méthode qui nous a si bien servi dans les sciences naturelles.

*Laplace, Essai ph. sur les probabilités.*

**TOME PREMIER.**

---

**PARIS,**  
**BACHELIER, IMPRIMEUR-LIBRAIRE,**  
**QUAI DES AUGUSTINS, n° 55.**

**1835**

statistiques. L'ouvrage comporte par ailleurs, quarante-six lettres groupées en quatre thèmes : *Sur les probabilités* ; *Des moyennes et des limites* ; *De l'étude des causes* ; *De la statistique*. En fait, Quetelet a importé de la théorie probabiliste des erreurs la distribution gaussienne et le choix de la valeur moyenne arithmétique comme valeur la plus probable mesurée mais n'a pas appliqué la méthode des moindres carrés pour l'optimisation des calculs.

Dans la lettre XXX<sup>1</sup> intitulée *Sur les causes variables qu'on peut regarder comme constantes – Loi des grands nombres*, Quetelet expose qu'il considère qu'un ensemble de mesures sur un grand nombre d'hommes est équivalent à un grand nombre de mesures sur un homme moyen, les dispersions étant équivalentes à des erreurs. Il interprète dans ce but le théorème de convergence de Jacob Bernoulli et écrit :

« Cet habile géomètre avait employé une partie de sa vie à démontrer ce résultat qui nous paraît si simple aujourd'hui, à savoir : Que la moyenne donnée par une série d'expériences tombe près du nombre cherché, dans des limites d'autant plus étroites que les épreuves ont été plus multipliées [...] Mais Bernoulli a établi ses calculs dans l'hypothèse que le nombre cherché était fixe et constant. Il peut arriver que cette quantité éprouve de petites variations [...]. Or le principe de Bernoulli est encore applicable dans ce cas, et c'est ce qui a été démontré par M. Poisson au moyen d'une savante analyse [...]. Dans ce cas, les expériences doivent en général être très multipliées : c'est par ce motif que M. Poisson a nommé la loi des grands nombres l'extension du principe de Bernoulli.[...] Nous avons déjà eu l'occasion de considérer [...] des mesures prises sur un grand nombre d'hommes du même âge. Nous avons vu alors que toutes les tailles réunies s'accordaient à donner une taille moyenne dont elles s'écartaient individuellement d'après une loi bien marquée, et absolument comme si un individu avait été mesuré un grand nombre de fois par des moyens plus ou moins défectueux. »

Pour Quetelet, l'itération des épreuves dont on observe les résultats fonctionne comme un filtre qui « détruit tous les effets des causes accidentelles » et met à jour les effets des seules « causes constantes » des phénomènes.

En fait, toute la *physique sociale* de Quetelet est basée sur ce principe de loi de distribution et de valeur moyenne. Suite à sa correspondance avec les mathématiciens Laplace, Fourier et Poisson, Quetelet devient ainsi le premier à utiliser la courbe normale autrement que pour la répartition d'erreurs. Il considère d'ailleurs les probabilités et les statistiques comme deux disciplines très proches l'une de l'autre et il écrit :

« La théorie des probabilités prit naissance presque en même temps que la statistique, sa sœur puînée, dont elle devait devenir la compagne la plus sûre et la plus indispensable. Cette concordance n'est point accidentelle, mais l'une de ces sciences

---

1. Pages 213-216.

interroge en quelque sorte par ses calculs et coordonne ce que l'autre obtient par ses observations et ses expériences. »<sup>1</sup>

Les mesures d'une caractéristique humaine sont, pour Quetelet, groupées suivant une loi normale. C'est en étudiant les mensurations de conscrits français et celles d'environ 5 000 soldats écossais (publiées en 1817 dans la revue *Edinburgh Medical Journal*) que Quetelet a remarqué les régularités temporelles des distributions des données biométriques de l'homme, comme le poids, la taille, le périmètre thoracique, qui avaient, de plus, une forme gaussienne. Il distingue, par ailleurs, trois types de moyennes : la moyenne objective, qui correspond à un objet réel, sur lequel on réalise un certain nombre de mesures ; la moyenne subjective, qui résulte du calcul d'une tendance centrale ; le troisième cas est une simple moyenne arithmétique sur des objets qui n'obéissent pas à une loi normale.

Par sa méthode qui applique les lois des probabilités aux phénomènes humains, Quetelet a mis en évidence la constance des taux de criminalité et de suicides par année et d'autres paramètres démographiques.

Dans les mêmes lettres de 1846<sup>2</sup>, Quetelet donne sa définition de la statistique :

« La statistique a donc pour objet de nous présenter l'exposé fidèle d'un état, à une époque déterminée [...]. Je pense que la définition que je propose, et qui du reste s'écarte peu de celle donnée par plusieurs savants modernes, circonscrit suffisamment les attributions de la statistique, pour qu'on ne puisse pas la confondre avec les sciences historiques ou les autres sciences politiques et morales qui s'en rapprochent le plus. Elle ne s'occupe d'un état que pour une époque déterminée ; elle ne réunit que les éléments qui se rattachent à la vie de cet état, s'applique à les rendre comparables et les combine de la manière la plus avantageuse pour reconnaître tous les faits qu'ils peuvent nous révéler. »

Dans les « Notes » du même ouvrage<sup>3</sup>, on peut lire :

« [Certains statisticiens] voudraient tout réduire à des nombres et faire consister la science dans un vaste recueil de tableaux ; d'autres, au contraire, semblent craindre les nombres et ne les regardent que comme donnant des idées incomplètes et superficielles des choses. Ces deux excès seraient également nuisibles ; mais il est vrai de dire que, quand des résultats peuvent être appréciés par des chiffres, c'est toujours aux chiffres qu'il faudra recourir pour avoir des appréciations exactes et des résultats comparables. »

Quetelet tenait compte dans ses études de tous les paramètres pouvant influencer les données sociales collectées et insistait sur la différence entre

1. *Physique sociale*, Bruxelles, Muquard, 1869, tome 1 p. 134.

2. Lettre XXXV, p. 264-269.

3. Page 432.

les phénomènes périodiques naturels et sociaux. Il écrivait en 1842 dans une note à l'Académie royale des sciences et belles-lettres de Bruxelles :

« Les phénomènes périodiques peuvent être partagés en deux grandes classes : les uns appartiennent aux sciences physiques et naturelles ; les autres sont plutôt du domaine de la statistique et concernent l'homme, vivant au milieu de l'état social ; car la société elle-même, avec ses tendances à se soustraire le plus possible aux lois naturelles, n'a pu échapper à cette périodicité qui nous occupe.

Les phénomènes périodiques *naturels* sont généralement indépendants des phénomènes périodiques *sociaux* ; mais il n'en est pas de même de ceux-ci à l'égard des premiers. Ce serait donc avoir fait un premier pas sur ce terrain si peu défriché, et qui semble promettre tant aux travaux de ceux qui sauront l'exploiter, que d'avoir commencé l'étude simultanée de tous les phénomènes périodiques qui se rattachent aux sciences physiques et naturelles. »

Les phrases les plus célèbres de Quetelet sont celles par lesquelles il affirme que l'organisation d'une société est responsable de certaines lois statistiques, certes non invariables, mais prévisibles dans un contexte donné :

« Il est un budget qu'on paie avec une régularité effrayante, c'est celui des prisons, des bagnes et des échafauds ; c'est celui-là qu'il faudrait s'attacher à réduire ; et, chaque année, les nombres sont venus confirmer mes prévisions, à tel point, que j'aurais pu dire, peut-être avec plus d'exactitude : Il est un tribut que l'homme acquitte avec plus de régularité que celui qu'il doit à la nature ou au trésor de l'État, c'est celui qu'il paie au crime ! Triste condition de l'espèce humaine ! Nous pouvons énumérer d'avance combien d'individus souilleront leurs mains du sang de leurs semblables, combien seront faussaires, combien empoisonneurs, à peu près comme on peut énumérer d'avance les naissances et les décès qui doivent avoir lieu.

La société renferme en elle les germes de tous les crimes qui vont se commettre, en même temps que les facilités nécessaires à leur développement. C'est elle, en quelque sorte, qui prépare ces crimes, et le coupable n'est que l'instrument qui les exécute. Tout état social suppose donc un certain nombre et un certain ordre de délits qui résultent comme conséquence nécessaire de son organisation. »<sup>1</sup>

Ces études de Quetelet sur la constance de la fréquence des crimes et délits suscitèrent une large discussion des concepts de liberté et de déterminisme social. Une controverse prit naissance au sein des sociologues du XIX<sup>e</sup> siècle autour de la notion « d'homme moyen » qui, d'abord appliquée aux qualités physiques, eut tendance à être appliquée aux qualités intellectuelles et morales jusqu'à être éventuellement présentée comme un type idéal. Ainsi cette théorie fut critiquée par Émile Durkheim et Maurice Halbwachs qui refusèrent une telle généralisation.

Au sujet de l'*homme moyen*, J. Bertrand a émis un jugement féroce<sup>2</sup> :

- 
1. *Sur l'homme et le développement de ses facultés. Essai de physique sociale*, Paris, 1835, tome I, p. 9-10.
  2. J. Bertrand, *Calcul des Probabilités*, Paris, Gauthier-Villars, 1888, p. XLI-XLIII.

« M. Quetelet, dans un livre riche de faits judicieusement recueillis, a voulu définir et préciser le mot *homme* indépendamment des hommes particuliers considérés comme accidents.[...] Dans le corps de l'homme moyen, l'auteur belge place une âme moyenne. L'homme type sera donc sans passions et sans vices, ni fou ni sage, ni ignorant ni savant, souvent assoupi : c'est la moyenne entre la veille et le sommeil, ne répondant ni oui ni non ; médiocre en tout. Après avoir mangé pendant trente-huit ans la ration moyenne d'un soldat bien portant, il mourrait, non de vieillesse, mais d'une maladie moyenne que la Statistique révélerait pour lui. »

## **18.11. Biographie de Graunt (1620-1674)**

John Graunt est né à Londres le 24 avril 1620. Mercier de profession, il a été l'un des premiers statisticiens. Le seul ouvrage de Graunt, qui l'a cependant rendu célèbre, *Natural and Political... Observations upon the Bills of Mortality* date de 1662 et est considéré comme la fondation des disciplines des statistiques et de la démographie. On ignore exactement quelle assistance éventuelle il a pu recevoir de Sir William Petty pour la rédaction de cet ouvrage.

Graunt, dans son livre a listé toutes les connaissances que l'on pouvait obtenir à partir de tables de mortalité et de toutes les statistiques relatives à la vie des citoyens, puis il a analysé tous les paramètres susceptibles d'être à l'origine de fausses conclusions et indiqué toutes les découvertes faites à partir des tables dont il disposait.

Grâce à son livre, Graunt, a été élu membre de la Royal Society en 1662.

John Graunt est mort le 18 avril 1674 à Londres.

### **Références relatives à Graunt**

- K. Pearson, *The History of Statistics in the 17th and 18th centuries*, London, Griffin, 1978.
- F. N. Edgerton III, *John Graunt*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 5, p. 506-508, 1970.

## **18.12. Biographie d'Arbuthnot (1667-1735)**

John Arbuthnot est né en avril 1667<sup>1</sup> à Arbuthnot, Kinkardineshire, sur la côte nord est de l'Écosse. Son père, Alexander Arbuthnot, était pasteur et sa mère, née Margaret Lammie, fille de pasteur. John est l'aîné des sept enfants du couple. Il reçoit de son père les bases du latin et du grec. John entre en 1681 au Marischal College<sup>2</sup> où il étudie, en particulier, les mathématiques

1. On ne connaît pas sa date de naissance exacte, on sait simplement qu'il a été baptisé le 29 avril 1667.
2. Son père y avait également fait ses études. James Clerk Maxwell y enseignera, près de deux siècles plus tard.

et la physique et obtient son diplôme en 1685. On sait très peu de chose sur la vie de John jusqu'à la mort de son père en 1691<sup>1</sup> (période sombre durant laquelle son père était impliqué dans des luttes cléricales). John s'établit alors à Londres où il donne des leçons de mathématiques. Il traduit le fameux traité de Huygens *De ratiociniis in ludo aleae* (publié en 1657) auquel il ajoute des considérations concernant divers jeux. Sa publication, anonyme, s'intitule *Of the Laws of Chance* et est datée de 1692. C'est le premier ouvrage sur les probabilités publié en anglais.

Arbuthnot devient précepteur de Jeffrey Jeffreys, le fils d'un membre du parlement. Ils vont ensemble à Oxford où Arbuthnot peut échanger des idées avec son ami David Gregory<sup>2</sup>. Par ailleurs, il étudie la médecine de 1694 à 1696 et soutient son doctorat le 11 septembre 1696. Lorsque la Reine Anne accède au trône en 1702, Arbuthnot devient médecin à la cour. Il se marie avec Margaret Wemyss et leur premier enfant, George, naît en 1703. L'année suivante, Arbuthnot est élu membre de la *Royal Society* puis devient, le 30 octobre 1705, médecin de la Reine Anne.

Il poursuit par ailleurs ses travaux scientifiques et soumet, en 1710, son célèbre article relatif au léger excédent de naissances masculines par rapport aux naissances féminines durant les années allant de 1629 à 1710. Cet article, publié dans les *Philosophical Transactions*, est considéré comme la première application des probabilités aux statistiques sociales.

Après la mort de la Reine Anne, le 1<sup>er</sup> août 1714, il séjourne en France. Il revient ensuite en Angleterre comme médecin de hauts personnages.

Arbuthnot, qui possède un important embonpoint, souffre d'asthme et de calculs rénaux. Sa santé se dégrade en 1734 et il meurt dans sa maison de Cork Street à Londres le 27 février 1735.

John Arbuthnot était également écrivain (souvent satirique). Il appartenait au Scriblerus Club et c'est dans ce cadre qu'il inspira le livre III des *Voyages de Gulliver* de Jonathan Swift et de nombreux ouvrages d'Alexander Pope. On lui doit le personnage de John Bull. Le portrait suivant a été fait par l'un de ses amis, le docteur Johnson : « [Arbuthnot] était un homme de grande compréhension, habile et compétent dans sa profession, versé en sciences, très avisé en littérature antique, et capable de mettre en œuvre ses connaissances grâce à une imagination brillante et active ; un érudit avec beaucoup d'esprit ; une intelligence, qui, durant sa vie mouvementée, a maintenu et a découvert une noble ardeur religieuse ».

- 
1. John Arbuthnot était réticent à laisser des traces, il disait : « la biographie est l'une des nouvelles terreurs de la mort ».
  2. Lors de son séjour à Oxford, Arbuthnot a rencontré de nombreux savants, parmi lesquels Isaac Newton, John Radcliffe et Samuel Pepys.

### **18.13. Principaux ouvrages d'Arbuthnot et références**

- *An Essay of the Usefulness of Scientific Nature* (1701).
- *Tables of Ancient Coins, Weights and Measures* (1727).
- *The Miscellaneous Works of the later Dr Arbuthnot* (1751).

#### **Références**

- H. Freudenthal, *John Arbuthnot*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 1, p. 208-209, 1970.
- G. A. Aitken, *The Life and Works of John Arbuthnot*, Clarendon Press, 1892.
- L. M. Beattie, *John Arbuthnot : Mathematician and Satirist*, Harvard University Press, 1935.

### **18.14. Biographie de Deparcieux (1703-1768)**

Antoine Deparcieux est né le 28 octobre 1703 au mas du Clotet du hameau de Cessous, paroisse de Peyremale, qui appartient aujourd'hui à la commune de Portes près d'Alès, dans le Gard. Il est d'origine très modeste : son père, Jean-Antoine Deparcieux, est fermier. Ce dernier épouse en 1674 Jeanne Donzel de Peyremale. Ils ont sept enfants nés entre 1675 et 1703. Antoine est le plus jeune, il apprend à lire et à écrire chez le prieur de Portes, puis est envoyé à l'école de Saint Florent sur Auzonnet. Il fréquente cette école pendant dix ans tout en participant aux travaux de la ferme parentale. Orphelin de père à douze ans et ayant perdu sa mère, c'est son grand frère, Pierre Deparcieux, de vingt-sept ans plus âgé que lui, qui prend soin d'Antoine. En 1718, ce dernier l'envoie au collège des jésuites d'Alès, probablement grâce à un support financier externe. Antoine se montre très à l'aise dans les sujets scientifiques.

Vers l'âge de vingt ans il quitte définitivement sa famille. On le retrouve à Lyon où il offre ses services au collège des Jésuites en échange de la permission de suivre les cours. Il apprend les mathématiques grâce à l'un des pères.

En 1730 il part pour Paris où ses débuts sont difficiles par manque d'argent. Il a la chance de rencontrer le géomètre De Montcarville, membre de l'Académie des sciences et professeur au Collège de France. Ce dernier le protège, lui donne des compléments d'enseignements en mathématique et mécanique et en fait son assistant au Collège de France. Pour améliorer sa

condition financière, Deparcieux réalise des cadrans solaires qui sont très appréciés : il compte même le Duc de Nevers parmi ses clients. C'est à ce moment qu'il commence à se consacrer à l'étude des applications sociales des mathématiques. Il devient l'ami intime de Fontenelle, de Lavoisier, de Cassini, de Vaucanson et de Réaumur. Il publie en 1733 un premier traité de trigonométrie, qu'il étend en 1741 sous le titre de *Nouveaux traités de trigonométrie rectiligne et sphérique avec un traité de Gnomonique*<sup>1</sup>. Ses traités « sont accompagnés de tables des sinus, tangentes... des logarithmes des nombres de 1 à 20 000, et des logarithmes des sinus et tangentes »<sup>2</sup>. En 1746, Deparcieux compose, à la demande de M. de Boullongne, intendant des Finances, son ouvrage le plus célèbre intitulé : *Essais sur les probabilités de la vie humaine, d'où l'on déduit la manière de déterminer les rentes viagères tant simples que tontines, précédé d'une courte explication sur les rentes à terme, ou annuités, et accompagné d'un grand nombre de tables*<sup>3</sup>.

Ses travaux le font nommer à l'Académie des sciences, d'abord en tant qu'ad-joint géomètre le 11 février 1746, puis associé géomètre le 16 mai 1756<sup>4</sup>, et enfin pensionnaire géomètre le 9 juin 1768<sup>5</sup>.

Antoine Deparcieux se passionne pour la technique et réalise bientôt également d'autres équipements que des cadrans solaires. En 1751 il réalise, à la demande de Madame de Pompadour, une machine pour élever les eaux du château de Crécy-Couvé, près de Dreux. Sa machine hydraulique permet d'élever de 163 pieds (environ 50 mètres) les eaux de la rivière de Blaise pour les besoins du château, des jardins et des maisons du lieu. À partir de 1762, il étudie, à la demande de la ville de Paris, la possibilité d'amener l'eau de l'Yvette à Paris. Il prévoit de capter l'eau en amont de Gif-sur-Yvette et de l'amener à Paris, au pied de la tour de l'Observatoire, à l'aide d'un aqueduc de 30 kilomètres. Il constitue un dossier très complet, mais, faute de financement, le projet ne sera jamais concrétisé. Il invente également une presse destinée à la fabrication du tabac.

En 1765, Louis XV nomme Antoine Deparcieux « Censeur Royal des livres »<sup>6</sup> et lui accorde un logement de fonction dans les galeries du Louvre.

1. La gnomonique est l'art de concevoir, calculer et tracer des cadrans solaires.
2. Deparcieux donne les tables des sinus, tangentes, sécantes ( $1/\cosinus$ ), etc. avec 7 décimales et les tables de logarithmes de sinus et de tangentes avec 8 décimales. On y trouve par ailleurs d'intéressantes formules trigonométriques.
3. Cet ouvrage contient les célèbres *Tables de Mortalité* qui ont été utilisées par les compagnies d'Assurances vie et les banques pendant tout le XIX<sup>e</sup> et le début du XX<sup>e</sup> siècle.
4. Il remplace d'Alembert qui devient pensionnaire.
5. Il sera également nommé membre des Académies des sciences de Berlin et de Stockholm.
6. Le censeur royal est un érudit qui autorise, par délégation du roi, la publication d'un manuscrit sous forme imprimée. Une quarantaine de censeurs royaux sont nommés, chacun dans un domaine de spécialité. Ce concept a été initié par Richelieu.



Malheureusement, dès cette époque, Deparcieux commence à souffrir de paralysie et de rhumatisme. Il meurt le 2 septembre 1768 à Paris.

Durant toute sa vie Deparcieux a été modeste, jamais trop ambitieux, gardant toujours à l'esprit ses origines. Il ne s'est jamais marié et n'a pas eu d'enfant, expliquant « qu'il n'avait pas de temps pour cela ». Son petit-neveu, portant le même nom que lui, Antoine Deparcieux (1753-1799), est également connu. Il a remplacé Pilâtre de Rozier comme professeur d'astronomie, puis Condorcet comme professeur de mathématiques. Il publiera le *Traité des annuités ou des rentes à vie* de son grand-oncle en 1781.

### **18.15. Principaux ouvrages de Deparcieux et références**

- *Traité de trigonométrie rectiligne et sphérique*, 1738.
- *Nouveau traité de trigonométrie*, avec table des sinus et logarithmes, 1740.
- *Traité complet de gnomonique*, 1741.
- *Essai sur les probabilités de la durée de la vie humaine*, 1746.

#### **Références**

- H. L. L. Busard, *Deparcieux*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 4, p. 38-39, 1970.

### **18.16. Biographie de Quetelet (1796-1874)**

Lambert Adolphe Jacques Quetelet est né à Gand, maintenant en Belgique Flamande<sup>1</sup>, le 22 février 1796. Son père François-Augustin-Jacques-Henri Quetelet est d'origine picarde. Il a vécu un temps en Grande Bretagne où il avait acquis la nationalité anglaise. Devenu secrétaire d'un noble écossais il a voyagé en Europe et, en 1787, vers l'âge de 31 ans, il s'est établi à Gand. C'est là qu'il rencontre sa femme, Anne-Françoise Vandervelde. C'est dans un doux environnement familial que notre futur mathématicien passe sa petite enfance, mais alors qu'il n'a que sept ans un malheur survient : son père meurt.

Adolphe fait ses études au lycée, où il se montre très talentueux en mathématiques ainsi qu'en littérature, mais à la fin de sa scolarité il décide de travailler pour aider financièrement sa famille. Il obtient en 1813 un poste d'enseignant en mathématique dans une école située dans le village d'Aude-

---

1. À cette époque la Belgique dépend des Pays-Bas, elle ne prendra son indépendance qu'en 1830.

naerde. Deux ans plus tard il trouve un autre poste dans un collège, cette fois à Gand. Au lycée, Adolphe Quetelet s'était fait un ami, Germinal Dandelin, avec qui il partageait la passion des mathématiques, de la littérature et de la musique. Ils ont composé ensemble un livret pour un opéra ainsi que quelques drames. Quetelet était également en très bonne relation avec un professeur d'astronomie, Genter. C'est sans doute ce dernier qui orienta définitivement Quetelet vers les mathématiques.

Notre futur statisticien obtient en 1819 son doctorat de la toute jeune université de Gand<sup>1</sup>. Sa thèse de mathématiques porte sur une classe de courbes utilisées en optique. Il obtient alors un poste de professeur de mathématiques à l'Athenaeum de Bruxelles. En février 1820, il n'a alors que vingt-quatre ans, il est élu à l'Académie Royale des sciences de Belgique<sup>2</sup>.

Très intéressé par l'astronomie il parvient à convaincre l'administration du Royaume des Pays-Bas de la nécessité de construire un observatoire astronomique à Bruxelles, sur le modèle de ceux qui existent déjà en France et en Angleterre. Pour préparer ce projet, il se rend à Paris en décembre 1823 et y reste quelques mois. Il y rencontre les meilleurs astronomes et mathématiciens français : Alexis Bouvard, François Arago, Augustin Fresnel, Pierre Simon Laplace, Alexandre von Humboldt (dont on sait l'attachement à la France et l'amitié avec Arago), Joseph Fourier, Siméon-Denis Poisson. Ce séjour aura une importance décisive pour la suite de son histoire. En effet, il y découvre l'usage que ces savants font du calcul des probabilités pour contrôler les erreurs de mesure en astronomie. De retour à Bruxelles il obtient un poste plus élevé à l'Athenaeum et donne son premier cours sur le calcul des probabilités durant l'année académique 1824-1825. Il commence également à donner des conférences publiques au Musée de Bruxelles sur des matières telles que la géométrie, les probabilités, la physique, et l'astronomie, ainsi que des cours sur l'histoire des sciences.

En 1825, Quetelet se marie avec la fille d'un médecin français, M. Curtet. Le couple aura un fils, Ernest<sup>3</sup>, et une fille. Par ailleurs son projet d'observatoire se poursuit. En 1827 on lui confie la tâche d'en choisir les instruments. Il se rend en Grande Bretagne avec son ami Dandelin, et ils visitent et recueillent des conseils auprès des universités, observatoires et sociétés savantes tant à Londres, que dans le reste de l'Angleterre ainsi qu'en Écosse et en Irlande.

1. L'université de Gand a été fondée en octobre 1817 par William I<sup>er</sup>, roi des Pays-Bas et du Grand-Duché de Luxembourg.
2. L'Académie royale de Belgique a été fondée à Bruxelles en 1769 en tant qu'Académie purement littéraire. En 1816 elle est réorganisée en Académie royale des sciences et des belles-lettres et placée sous l'autorité de William I<sup>er</sup>. Quetelet y joue un rôle très actif et lui confère une organisation solide. Il en sera le directeur en 1832 et 1833, puis le Secrétaire, poste qu'il occupera jusqu'à sa mort.
3. Ernest deviendra astronome et succédera à son père comme directeur de l'observatoire de Bruxelles.

En 1829, cette fois accompagné de son épouse, il rend visite à des astronomes hollandais et allemands et devient familier des observatoires de ces pays. En 1830 il part pour l'Italie et la Sicile, toujours dans le but de mieux connaître le fonctionnement des observatoires. Finalement celui de Bruxelles ouvre en 1832 et, tout naturellement, il en devient le premier directeur. Il emménage à l'observatoire et c'est là qu'il fera ses travaux sur les statistiques, la géophysique, la météorologie et l'astronomie (étude des essaims de météorites).

Pendant la période allant de 1825 à 1835 il rédige plusieurs articles sur les statistiques sociales, en particulier *Recherches sur le penchant au crime aux différents âges* et surtout *Sur l'homme et le développement de ses facultés, essai d'une physique sociale*. Ce livre, publié en 1835, lui conférera une renommée internationale. Quetelet est le premier à utiliser la distribution de Gauss pour une autre application que celle concernant les erreurs de mesure. Dans son article de 1835, il introduit la notion d'homme moyen, défini comme la valeur centrale autour de laquelle les mesures d'un trait humain sont réparties selon une gaussienne. À la demande du gouvernement, Quetelet recueille et analyse les statistiques sur le crime et la mortalité et met au point des améliorations pour le recueil des données.

Quetelet organise, en 1853, le premier « congrès international de statistique ». L'objectif est de pouvoir faire se rencontrer les statisticiens des « bureaux de statistique » créés, souvent sous son influence, dans plusieurs pays au cours des deux décennies précédentes. Il s'agit pour Quetelet de créer un réseau coordonné de recueil d'informations sur la population, l'économie et les statistiques « morales » ayant des nomenclatures et des procédures d'enregistrement les plus standardisées possibles. Il fait en sorte qu'une autre conférence internationale, sur la météorologie cette fois, se déroule de façon concomitante. Ainsi la première conférence « traite de la statistique générale des différents pays et des moyens de mettre l'unité entre les documents officiels destinés à l'administration et à la science », tandis que l'autre « concerne la marine et l'accord qu'il s'agit d'établir entre les travaux des divers peuples pour arriver à connaître les lois qui règlent les mouvements des mers et de l'atmosphère, la profondeur et la température des eaux, et en général tout ce qui peut intéresser le navigateur ». Quetelet cherche à promouvoir un système de statistique internationale reposant sur quatre principes : 1) mettre en place un réseau de mesures ; 2) standardiser les procédures de mesures ; 3) impliquer les États dans cette mise en place ; 4) s'assurer du caractère utile (bonne administration de la société, navigation maritime) et non uniquement scientifique de ce système.

Durant l'été 1855, Quetelet est frappé par la maladie. Cette dernière n'est pas très grave et il ne tarde pas à bénéficier d'un bon rétablissement physique mais mentalement il ne sera plus aussi vivace qu'auparavant et sa mémoire lui fait progressivement défaut.

Un de ses biographes en fait le portrait suivant : « Modeste et généreux, convaincu mais respectueux des avis d'autrui, toujours calme et prévenant, un homme de grande culture et de conversation agréable, il s'est fait des amis partout où il est allé et les a gardés. Un homme d'un tact exquis, et d'un formidable enthousiasme... Possédant de nombreux centres d'intérêts intellectuels, en même temps qu'une prodigieuse capacité de travail... toujours plein d'entrain et jovial, il avait de l'esprit et le sens de l'humour. »

Adolphe Quetelet s'est éteint le 17 février 1874 à Bruxelles.

## **18.17. Principaux ouvrages de Quetelet et références**

- *Recherches sur la population, les naissances, les décès, les prisons, les dépôts de mendicité, etc. dans le royaume des Pays-Bas*, Nouv. Mém. Acad. R., Bruxelles, 4, 117-192, 1827.
- *Sur l'homme et le développement de ses facultés, essai d'une physique sociale*. Paris, 1835.
- *Sur l'appréciation des documents statistiques, et en particulier sur l'appréciation des moyennes*, Bull. Comm. Cent. Star., Belg., 2, 205-286, 1845.
- *Lettres à S. A. R. le duc régnant de Saxe-Cobourg et de Gotha sur les Probabilités*, Brussels, 1846.
- *Du système social et des lois qui le régissent*, Paris, 1848.

### **Références**

- H. Freudenthal, *Quetelet*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 11, p. 236-238, 1970.
- O. B. Sheynin, *Quetelet as a statistician*, AHES, vol. 36, p. 281-325. 1986.



# Cournot

## 19.1. Les écrits de Cournot concernant les fondements du calcul des probabilités

Les travaux de Cournot relatifs aux probabilités et aux statistiques ne comportent aucune innovation mathématique, mais ils sont importants dans l'histoire du calcul des probabilités parce qu'ils constituent une tentative d'interprétation des fondements de ce calcul. Pour Cournot, le hasard est un concept objectif et tout événement est produit par une cause.

Dans son *Essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique*<sup>1</sup> Cournot affirme :

« De même que toute chose doit avoir sa raison, tout ce que nous appelons événement doit avoir une cause. Souvent la cause d'un événement nous échappe, ou nous prenons pour cause ce qui ne l'est pas ; mais, ni l'impuissance où nous nous trouvons d'appliquer le principe de causalité, ni les méprises où il nous arrive de tomber en voulant l'appliquer inconsidérément, n'ont pour résultat de nous ébranler dans notre adhésion à ce principe, conçu comme une règle absolue et nécessaire. »

Et, plus loin<sup>2</sup> :

« À la vérité, les géomètres ont appliqué leur théorie des chances et des probabilités à deux ordres de questions bien distinctes, et qu'ils ont parfois mal à propos confondues : à des questions de *possibilité*, qui ont une valeur tout objective, ainsi qu'on vient de l'expliquer, et à des questions de *probabilité*, dans le sens vulgaire du mot, qui sont en effet relatives, en partie à nos connaissances, en partie à notre ignorance. »

En 1843, Cournot définissait le hasard, dans une proposition devenue célèbre, comme la "rencontre de deux séries causales indépendantes" :

« Les événements amenés par la combinaison ou la rencontre de phénomènes qui appartiennent à des séries indépendantes, dans l'ordre de la causalité, sont ce qu'on nomme des événements fortuits ou des résultats du hasard. »<sup>3</sup>

1. Paris, Hachette, 1851, article 29.

2. Article 37, 1851.

3. A. Cournot, *Exposition de la théorie des chances et des probabilités*, page 73, Hachette, 1843.

**EXPOSITION**  
**DE LA THÉORIE**  
**DES CHANCES**  
**ET**  
**DES PROBABILITÉS**

**PAR H. A. A. COURNOT**

**INSPECTEUR GÉNÉRAL DES ÉTUDES**

Quantumvis ratiois terminis  
comprehendi posse videtur quod fortalis  
sunt etque incerta, tanta mirabilis ars  
censetur, cui ista quoque subiacent

**HOVENUS**



**PARIS**  
**LIBRAIRIE DE L. NACHETTE**  
**RUE PIERRE-GARRAUX, N° 12**

Dans ses *Considérations sur la marche des idées et des événements dans les temps modernes*<sup>1</sup>, on peut lire :

« Nous croyons avoir éclairci dans d'autres ouvrages et défini, plus exactement que ne l'avaient fait nos devanciers, l'idée du *hasard* en montrant que ce n'est point, comme on l'a tant répété, un fantôme créé pour nous déguiser à nous-mêmes notre ignorance, ni une idée relative à l'état variable et toujours imparfait de nos connaissances, mais bien au contraire la notion d'un fait vrai en lui-même, et dont la vérité peut être dans certains cas établie par le raisonnement, ou plus ordinairement constatée par l'observation, comme celle de tout autre fait naturel. Le fait naturel ainsi établi ou constaté consiste dans l'*indépendance* mutuelle de plusieurs séries de causes et d'effets qui concourent accidentellement à produire tel phénomène, à amener telle rencontre, à déterminer tel événement, lequel pour cette raison est qualifié de *fortuit* et cette indépendance entre des chaînons particuliers n'exclut nullement l'idée d'une suspension commune de tous les chaînons à un même anneau primordial, par-delà les limites, ou même en deçà des limites où nos raisonnements et nos observations peuvent atteindre. De ce que la Nature agit sans cesse et partout le cornet du hasard, et de ce que le croisement continu des chaînes de conditions et de causes secondes, indépendantes les unes des autres, donne perpétuellement lieu à ce que nous nommons des *chances* ou des combinaisons fortuites, il ne s'ensuit pas que Dieu ne tienne point dans sa main les unes et les autres, et qu'il n'ait pu les faire sortir toutes d'un même décret initial. On ne manque pas plus de respect à Dieu en étudiant les lois du hasard (car le hasard même a ses lois que met en évidence la multiplicité des épreuves), qu'en étudiant les lois de l'astronomie ou de la physique. La raison même nous impose l'idée du hasard ; et le tort imputable à notre ignorance consiste, non à nous forger cette idée, mais à la mal appliquer, ce dont il n'y a que trop d'exemples, même chez les plus habiles. Elle est le principe de toute espèce de critique, soit qu'il s'agisse des plus hautes spéculations de la philosophie, ou des recherches de l'érudition, ou de la pratique la plus ordinaire de la vie. Elle est la clef de la *statistique*, et donne un sens incontestable à ce que l'on a appelé la *philosophie de l'histoire*, à ce que nous aimerions mieux appeler l'*étiologie historique*, en entendant par là l'analyse et la discussion des causes ou des enchaînements de causes qui ont concouru à amener les événements dont l'histoire offre le tableau ; causes qu'il s'agit surtout d'étudier au point de vue de leur indépendance ou de leur solidarité. Pour donner quelque valeur aux nombreux systèmes dont la philosophie de l'histoire a été l'objet, il faut toujours revenir aux principes de la critique ou de l'étiologie historique. »

## 19.2. Le principe de Cournot

L'existence des séries causales indépendantes, dont la rencontre correspond au hasard, dont il vient d'être question, amène Cournot à affirmer que le monde réel permet d'appréhender des probabilités objectives concernant la mesure de possibilités effectives, par opposition aux événements physiquement impossibles.

1. Livre 1, chapitre 1, Hachette, 1872.



On a donné le nom de principe de Cournot (appellation proposée par Maurice Fréchet) à un principe des probabilités négligeables. Le but de ce principe était, en fait, pour Cournot, d'assurer la valeur objective du calcul des probabilités. En effet, un événement physiquement impossible est, selon Cournot, un événement dont la probabilité est infiniment petite. Cette seule remarque donne substance, c'est-à-dire une valeur objective et phénoménologique, à la théorie mathématique des probabilités. En 1851, Cournot écrivait<sup>1</sup> :

« À la notion du hasard s'en attache une autre qui est de grande conséquence en théorie comme en pratique : nous voulons parler de la notion de l'*impossibilité physique* [...] l'événement physiquement impossible (celui qui de fait n'arrive pas, et sur l'apparition duquel il serait déraisonnable de compter tant qu'on n'embrasse qu'un nombre fini d'épreuves ou d'essais, c'est-à-dire tant qu'on reste dans les conditions de la pratique et de l'expérience possible) est l'événement qu'on peut assimiler à l'extraction d'une boule blanche par un agent aveugle, quand l'urne renferme une seule boule blanche pour une infinité de boules noires ; en d'autres termes, c'est l'événement qui n'a qu'une *chance* favorable pour une infinité de chances contraires. Mais on a donné le nom de *probabilité mathématique* à la fraction qui exprime le rapport entre le nombre des chances favorables à un événement et le nombre total des chances : en conséquence, on peut dire plus brièvement, dans le langage reçu des géomètres, que l'événement physiquement impossible est celui dont la probabilité mathématique est infiniment petite, ou tombe au-dessous de toute fraction, si petite qu'on la suppose. On peut dire aussi que l'événement physiquement certain est l'événement dont le contraire est physiquement impossible, ou l'événement dont la probabilité mathématique ne diffère de l'unité par aucune fraction assignable, si petite qu'on la suppose : événement qu'il ne faut pourtant pas confondre avec celui qui réunit absolument toutes les combinaisons ou toutes les chances en sa faveur, et qui est certain, d'une certitude mathématique. D'un autre côté, il résulte de la théorie mathématique des combinaisons que, quelle que soit la probabilité mathématique d'un événement dans une épreuve aléatoire, si l'on répète un très grand nombre de fois la même épreuve, le rapport entre le nombre des épreuves qui amènent l'événement A et le nombre total des épreuves doit différer très peu de la probabilité de l'événement A : de sorte que, par exemple, si l'événement A a pour lui les deux tiers des chances, et qu'on embrasse dix mille épreuves, le nombre des épreuves qui amènent l'événement A sera, à peu de chose près, les deux tiers de dix mille. Si l'on peut accroître indéfiniment le nombre des épreuves, on fera décroître indéfiniment, et l'on rendra aussi petite qu'on le voudra, la probabilité que la différence des deux rapports dépasse une fraction donnée, si petite qu'elle soit, et l'on se rapprochera ainsi de plus en plus des cas d'impossibilité physique cités tout à l'heure. Dans le langage rigoureux qui convient aux vérités abstraites et absolues des mathématiques et de la métaphysique, une chose est possible ou elle ne l'est pas : il n'y a pas de degrés de possibilité ou d'impossibilité. Mais, dans l'ordre des faits physiques et des réalités qui tombent sous les sens, lorsque des événements contraires peuvent arriver et arrivent effectivement, selon les combinaisons fortuites

1. *Essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique*, Paris, Hachette, 1851, § 33.

de certaines causes variables et indépendantes d'une épreuve à l'autre, avec d'autres causes ou conditions constantes qui régissent solidairement l'ensemble des épreuves, il est naturel de regarder chaque événement comme ayant une disposition d'autant plus grande à se produire, ou comme étant d'autant plus possible, de fait ou physiquement, qu'il se reproduit plus souvent dans un grand nombre d'épreuves. La probabilité mathématique devient alors la mesure de la *possibilité physique*, et l'une de ces expressions peut être prise pour l'autre. »

Il est à noter qu'Émile Borel<sup>1</sup>, à partir de considérations spatio-temporelles concernant notre Univers a affirmé, lui aussi, un principe de probabilités négligeables :

La conclusion qui s'impose, écrit Borel,

« est que les probabilités qui s'expriment par un nombre inférieur à  $10^{-1000}$  sont non seulement négligeables dans la pratique courante de la vie, mais universellement négligeables, c'est-à-dire doivent être traitées comme *rigoureusement égales à zéro* dans toutes les questions qui concernent notre Univers. Le fait qu'elles ne sont pas effectivement nulles peut éventuellement intéresser le métaphysicien ; pour le savant, elles *sont* nulles et les phénomènes auxquels elles se rapportent sont absolument *impossibles*. »

Thierry Martin a publié en 1994, une étude détaillée portant le titre *La valeur objective des probabilités selon Cournot*<sup>2</sup>. On peut y lire :

« Nombre d'historiens ont vu dans l'œuvre de Cournot le triomphe d'une interprétation fréquentiste de la probabilité. [...] On ne peut... sans mutiler la pensée de Cournot la réduire à une interprétation objective de type fréquentiste. [...] Les stratifications du concept de probabilité permettent, lorsqu'elles sont distinguées et non plus confondues, d'éviter les *équivoques* que Cournot aperçoit chez ses prédécesseurs mêlant les significations objectives et subjectives des probabilités... »

### 19.3. Biographie de Cournot (1801-1877)

Antoine Augustin Cournot naît le 28 août 1801 à Gray en Haute-Saône. Son père est un commerçant issu d'une famille de fermiers. Entre 1809 et 1816, Antoine fréquente le Collège Jésuite de Gray. Durant les quatre années suivantes, il travaille en amateur comme clerc dans une étude de juristes. Il est enfin admis dans la classe de mathématiques spéciales du Collège Royal de Besançon puis intègre l'École Normale Supérieure en 1821. Mais l'École est fermée par le grand maître de l'Université de France, l'abbé Frayssinous, le 6 septembre 1822. Il poursuit alors son cursus universitaire pour obtenir une licence à la Sorbonne. Lacroix et Hachette sont parmi ses professeurs. Dans ses *Souvenirs*<sup>3</sup>, Cournot raconte sa fréquentation du salon de Hachette, où il

1. E. Borel, *Sur les probabilités universellement négligeables*, CRAS, t. 190, 1930, p. 540.

2. *Math. et Sc. humaines*, tome 127, p. 5-17, 1994.

3. Éditées à titre posthume, en 1913.

rencontra notamment le physicien danois Ørsted en 1823 dans une soirée, honorée aussi par Ampère. Dans ces mêmes années, Cournot commença à fréquenter Poisson, dont il sera très proche dans les années 1830.

Cournot a aussi pour condisciple et ami Dirichlet. Il devient, à la même période précepteur du fils du maréchal Gouvion Saint-Cyr, aide ce dernier à rédiger ses Mémoires, et reste dans la famille jusqu'au décès du maréchal en 1833.

Il obtient sa licence en Sciences en 1823 et quatre ans plus tard une licence en droit.

Au plan scientifique, il publie huit articles dans le *Bulletin des sciences* du baron de Férussac entre 1821 et 1831. En février 1829, il soutient une thèse de mécanique, et il traduit en 1834 le traité d'astronomie d'Herschell auquel il ajoute un supplément relatif à *la distribution des orbites cométaires* qui est son premier travail de probabilités et de statistiques où, selon ses propres termes, il applique la « théorie des chances » à la « philosophie naturelle ». Un an plus tard, il donne une traduction des *Éléments de Mécanique*, de Dionysius Lardner avec un chapitre additif sur « la mesure des forces et le travail des machines ».

À partir de 1833, il commence une carrière d'universitaire et de haut fonctionnaire : sur la recommandation de Poisson il est nommé professeur d'analyse à l'université de Lyon en 1834, puis, l'année suivante, recteur de l'académie de Grenoble et professeur à la faculté des sciences de cette ville. En 1836, il devient inspecteur général des études, préside le jury de l'agrégation de mathématiques, et, de 1856 à 1861, il est recteur de l'académie de Dijon.

En 1838, il publie son ouvrage *Recherches sur les principes mathématiques de la théorie des richesses* qui fait de lui le fondateur de l'économie mathématique. Il est ainsi l'un des premiers à avoir formulé un modèle de l'offre et de la demande.

La même année il s'intéresse, comme l'avaient fait Poisson et Condorcet, à l'interprétation des statistiques légales et publie une étude intitulée *Mémoire sur les applications du calcul des chances à la statistique judiciaire*<sup>1</sup> qu'il reprendra en 1843 dans son ouvrage principal sur les probabilités *Exposition de la théorie des chances et des probabilités*. Dans ce traité, Cournot donne une interprétation originale du hasard qui serait la "rencontre de deux séries causales indépendantes", les événements en eux-mêmes étant parfaitement déterminés quant à leur cause et à leur effet.

À partir de 1851, Cournot publie une série de travaux philosophiques : il propose une théorie des « révolutions scientifiques » étudie la connexion

1. *Journal de Liouville*, 4, 1838, p. 257-334.

entre la science et l'histoire et définit une classification des sciences dans son *Essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique* de 1851.

Cournot a aussi examiné les implications philosophiques des théories de l'évolution et il déclare que l'homme, à l'âge industriel, est un agent de la nature, qui laisse sur elle des traces irréversibles.

Antoine Augustin Cournot est mort, à Paris, le 30 mars 1877.

## 19.4. Principales publications de Cournot et références

- *Recherches sur les principes mathématiques de la théorie des richesses*, 1838.
- *Mémoire sur les applications du calcul des chances à la statistique judiciaire*, Journal des mathématiques pures et appliquées, 4, p. 257-334, 1838.
- *Traité élémentaire de la théorie des fonctions et du calcul infinitésimal*, 1841.
- *Exposition de la théorie des chances et des probabilités*, 1843.
- *De l'origine et des limites de la correspondance entre l'algèbre et la géométrie*, 1847.
- *Essai sur les fondements de nos connaissances et sur les caractères de la critique philosophique*, 1851.
- *Traité de l'enchaînement des idées fondamentales dans les sciences et dans l'histoire*, 1861.
- *Principes de la théorie des richesses*, 1863.
- *Considérations sur la marche des idées et des événements dans les temps modernes*, 1872.
- *Matérialisme, vitalisme, rationalisme : études des données de la science en philosophie*, 1875.
- *Souvenirs*, 1913 (publication posthume avec une introduction de E.P. Bottinelli).

## Références

- G. Granger, *Cournot*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, 1970, vol. 3, p. 450-454.
- Th. Martin, *Probabilités et philosophie des mathématiques chez Cournot*, Revue d'histoire des mathématiques, n° 1, Paris, 1995, p. 111-138.

- Th. Martin, *Probabilités et critique philosophique selon Cournot*, Paris, Vrin, coll. « Mathesis », 1996.
- Th. Martin, « Cournot et les mathématiques », in *Les philosophes et les mathématiques*, sous la direction d'Évelyne Barbin et Maurice Caveing, Paris, Ellipses, p. 193-211.

# Bienaymé

*Le calcul des probabilités donne lieu à de singulières illusions, auxquelles les meilleurs esprits n'ont pu toujours se soustraire<sup>1</sup>.*

## 20.1. Mémoire sur la probabilité des erreurs d'après la méthode des moindres carrés

Dans cet important papier<sup>2</sup> présenté le 27 octobre 1851 à l'Institut et publié en 1852, Bienaymé expose une généralisation de la démonstration, donnée par Laplace en 1811, relative à la normalité asymptotique de fonctions linéaires des erreurs sur plusieurs variables aléatoires ; une approche de la distribution du  $\chi^2$  pour l'exposant de la loi normale multivariée et une méthode des moindres carrés envisagée comme conduisant aux ellipsoïdes de confiance les plus petits.

Avant de citer des extraits du texte de Bienaymé, nous croyons utile d'exposer quelques rappels :

Une variable aléatoire  $X$  est dite avoir une distribution normale si sa densité de probabilité est égale à :

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

pour la paire de valeurs des paramètres  $\mu$  et  $\sigma$ .

La distribution normale de paramètres  $\mu$  et  $\sigma$  est notée  $N(\mu, \sigma^2)$ .

Loi normale de distribution bivariee :

Si on considère deux variables aléatoires  $X_1, X_2$ , de valeurs moyennes  $\mu_1, \mu_2$ , et d'écarts-type  $\sigma_1, \sigma_2$ , on appelle covariance de  $X_1, X_2$  l'espérance du produit  $(X_1 - \mu_1).(X_2 - \mu_2)$ ,

$$\text{Cov}(X_1, X_2) = E[(X_1 - \mu_1).(X_2 - \mu_2)]$$

- 
1. « Considérations à l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés », *C.r. Acad. Roy. Sci. Paris*, t. 37, p. 310.
  2. *Journal de Liouville*, 1, 17, 1852, p. 33-78 et *Mémoires Ac. Sc. Institut de France*, tome XV, 1858, p. 615-663.

*de lausier. Tardor*  
*73*

# MÉMOIRE

SUR

## PROBABILITÉ DES ERREURS

D'APRÈS LA MÉTHODE DES MOINDRES CARRÉS,

PAR M. BIENAYME,

*gal*  
INSPECTEUR DES FINANCES.

Nos sequimur probabilia, nec ultra quam id, quod  
verisimile occurrerit, progredi possumus; et refellere  
sine pertinacia, et refelli sine iracundia, parati sumus.

(Tusculan. 2.)



PARIS.

IMPRIMERIE IMPÉRIALE.

M DCCC LVIII.

On appelle Coefficient de Corrélation de Bravais-Pearson des variables  $(X_1, X_2)$ ,

$$\rho = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}}$$

Le coefficient de Corrélation est la version standardisée de la Covariance et ne dépend pas des unités dans lesquelles  $X_1$  et  $X_2$  sont exprimées.

La loi normale de distribution bivariable [variables  $(X_1, X_2)$  dont le coefficient de corrélation est  $\rho$ ] est la distribution dont la fonction densité de probabilité est

$$P(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{z}{2(1-\rho^2)}}$$

où

$$z = \frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} - \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2}.$$

Une distribution multivariée normale pour  $p$  variables aléatoires est une généralisation de la distribution normale bivariable. Une telle distribution avec un vecteur moyen  $\mu$  et une matrice de covariance  $\Sigma$  est notée  $N_p(\mu, \Sigma)$ .

On appelle distribution marginale la distribution individuelle d'une variable aléatoire sans tenir compte de la covariation d'autres variables aléatoires. À partir de la distribution conjointe d'un ensemble de variables aléatoires on définit la distribution marginale de toutes les variables qui le composent. Il faut souligner que les lois marginales déduites d'une distribution multinormale sont normales, de mêmes espérances et de mêmes variances.

Par définition, la distribution du Chi - 2 (notée  $\chi^2$ ) est celle de la somme des carrés des observations issues de la distribution  $N(0,1)$ .

Plus précisément, si  $\{X_1, X_2, \dots, X_n\}$  sont  $n$  variables indépendantes et toutes  $\sim N(0,1)$ , alors la distribution  $\chi^2_n$  est définie comme celle de la somme  $X_1^2 + X_2^2 + \dots + X_n^2$ .

### Loi normale multivariée

On peut généraliser la distribution normale univariée de densité de probabilité égale à :

$$\varphi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

pour la paire de valeurs des paramètres  $\mu$  et  $\sigma$ , en écrivant, pour un vecteur



aléatoire  $\vec{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$ , la densité de probabilité comme une fonction où inter-

viennent le vecteur des espérances mathématiques des composantes de  $X$ ,

$\vec{\mu} = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_n) \end{pmatrix}$  et une matrice définie positive  $\Omega$ , dite matrice de covariance,

où

$$|\Omega|_{ii} = V(X_i) = E(X_i - \mu_i)^2 \text{ et}$$

$$|\Omega|_{ij} = \text{Cov}(X_i, X_j) = E(X_i - \mu_i)(X_j - \mu_j).$$

### **Méthode des moindres carrés**

Dans la théorie des moindres carrés, on cherche à estimer un vecteur  $\vec{\beta}$   $[1 \times r]$ , ayant ses  $r$  composantes inconnues, à partir de  $N$  observations  $\vec{Y}$  reliées linéairement à  $\vec{\beta}$ , mais comportant des erreurs  $\vec{\varepsilon}$ .

On a  $\vec{Y} = M\vec{\beta} + \vec{\varepsilon}$  où

$M = \{x_{ij}\}$  est une matrice connue de  $N \times r$  éléments avec  $N \geq r$ , la colonne de rang  $r$  étant pleine.

Le système des multiplicateurs  $k_{ij}$  utilisés par Legendre, Gauss et Laplace pour traiter la méthode des moindres carrés constituait en fait la recherche d'une matrice  $K$  de  $r \times n$  éléments  $K = \{k_{ij}\}$ , telle que l'on obtienne  $KM = I$ ,  $I$  étant la matrice unité.

L'estimation de la moyenne du vecteur  $\vec{\beta}$  noté  $\bar{\vec{\beta}}$ , consiste alors à choisir la matrice  $K$  selon une approche probabiliste ou une approche algébrique afin d'obtenir l'estimation

$$\bar{\vec{\beta}} = \vec{\beta} + K\vec{\varepsilon} = K\vec{Y}$$

L'approche de Legendre, purement algébrique consiste à considérer la relation  $\vec{Y} = M\vec{\beta} + \vec{\varepsilon}$  comme un jeu surdéterminé d'équations linéaires ; la matrice  $K$  est alors établie en agissant sur les erreurs elles-mêmes. L'approche probabiliste consiste à établir  $K$  en imposant les distributions de probabilité sur les erreurs.

La méthode des moindres carrés, est, en fait, justifiée par les deux démarches, algébrique et probabiliste, et correspond au choix de  $K$  tel que  $K = (M' M)^{-1} M'$ . D'où

$$\bar{\beta} = (M' M)^{-1} M' \bar{Y}$$

En effet, si on a  $n$  observations  $x_1, x_2, \dots, x_n$ , leur moyenne arithmétique est

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n}; \text{ et la valeur de } \mu \text{ qui minimise la somme des carrés } \sum_{i=1}^n (x_i - \mu)^2$$

$$\text{est } \bar{x} = \sum_{i=1}^n \frac{x_i}{n}.$$

En 1809, Gauss a choisi de partir d'un postulat : « *On a coutume de regarder comme un axiome l'hypothèse que si une quantité a été obtenue par plusieurs observations immédiates, faites avec le même soin dans des circonstances semblables, la moyenne arithmétique des valeurs observées sera la valeur la plus probable de cette quantité, sinon en toute rigueur, du moins avec une grande approximation, de telle sorte que le plus sûr soit toujours de s'y arrêter* ».

Gauss introduit en fait le principe du maximum de vraisemblance car il cherche la valeur de la quantité inconnue qui rendra maximale la probabilité d'observation des mesures observées  $x_1, \dots, x_n$  et donc des erreurs  $\varepsilon_1, \dots, \varepsilon_n$ . Avec l'ajout du postulat de la moyenne arithmétique, il obtient à la fois la loi de distribution normale et le principe des moindres carrés. La justification de ce principe des moindres carrés est donc la maximalisation de la distribution *a posteriori* des erreurs. Les erreurs étaient supposées indépendantes et ayant toutes une densité de probabilité identique  $f(x)$ . La démarche de Gauss établit cette densité comme étant la loi normale  $N(0, \sigma^2)$

$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}. -\infty < x < \infty.$$

En 1821-1823, Gauss adopte toujours des erreurs supposées indépendantes et ayant toutes une densité de probabilité identique  $f(x)$  non spécifiée de variance définie  $\sigma^2$ , mais telle que  $f(x) = f(-x)$ . Il démontre que la méthode des moindres carrés correspond aux variances minimales des estimations.

Entre les deux publications de Gauss, Laplace a publié lui aussi deux papiers relatifs aux moindres carrés. Dans celui de 1811, Laplace montre que la distribution normale multivariée apparaît comme la distribution limite de formes linéaires des erreurs, pour  $n$  erreurs indépendantes et de lois de distribution identiques, lorsque  $n$  tend vers l'infini.

Il ajoute : « L'analyse exposée (précédemment) peut être étendue à la correction d'un nombre quelconque d'éléments par les observations. Elle conduit toujours à ce résultat : savoir que la méthode des moindres carrés des erreurs

des observations est celle qui donne sur la correction des éléments la plus petite erreur moyenne à craindre. »

Dans son 1<sup>er</sup> supplément à la Théorie Analytique, vers 1816, Laplace établit une méthode d'estimation indépendante des lois de distribution des erreurs (qu'il considère quelconques mais possédant une variance finie) et non pas limitée à des distributions normales. Pour  $m$  variables, il élimine  $m - 1$  paramètres dans ses équations de condition et applique la théorie de l'estimation pour le paramètre restant.

Il est à noter que toutes les tendances vers la loi normale, dans les travaux de Laplace sont asymptotiques pour un nombre d'observations infini.

Revenons à Bienaymé. Voici les considérations initiales de son papier de 1852 :

« On sait que la méthode inventée par Legendre, il y a environ cinquante ans, et publiée pour la première fois dans ses *Nouvelles méthodes pour la détermination des orbites des comètes*, in-4°, 1805, se réduit à multiplier chacune des équations du premier degré formées entre la grandeur observée et les inconnues par le coefficient de chacune de ces inconnues successivement ; à ajouter les produits donnés par les coefficients de la même inconnue, ce qui fournit autant de nouvelles équations que d'inconnues ; enfin à résoudre ces équations de la manière ordinaire. Les solutions ainsi obtenues jouissent de la propriété de ne renfermer que les moindres erreurs possibles pour une probabilité donnée. Ce n'est pas un minimum absolu, comme on a semblé le croire assez souvent, mais c'est un minimum relatif aux erreurs des observations et au mode choisi pour la combinaison de toutes les équations qu'elles fournissent. Il pourrait se trouver d'autres combinaisons plus avantageuses, et il y aurait à les discuter suivant les cas.

Le calcul de la grandeur des erreurs subsistantes, et de la probabilité qu'elles peuvent avoir, se fait suivant les règles ordinaires de la théorie, et n'entraîne que des difficultés analytiques. Quand il n'entre qu'une inconnue dans les équations, ce calcul est exact, du moins quant aux fondements théoriques ; mais quand il y a plusieurs inconnues, les règles données pour calculer l'erreur et la probabilité de chacune d'elles ne fournissent que l'erreur et la probabilité qu'elle pourrait avoir si elle était seule, et quelque grandes que fussent les erreurs des autres.

Or, un des premiers principes de la théorie des probabilités, c'est que, quand plusieurs événements arrivent simultanément, la probabilité du concours de ces événements est le produit des probabilités de chacun. De sorte que la probabilité de ce concours est inférieure à la probabilité de chaque événement pris à part, et elle est d'autant plus petite qu'il y a plus d'événements.

Évidemment, il en est de même des erreurs de plusieurs inconnues, la probabilité que ces erreurs restent, toutes à la fois, dans certaines limites, ne peut être que le produit des probabilités séparées que chacune ne s'écarte pas de ses limites propres, et par conséquent, cette probabilité du concours des erreurs de grandeur limitée doit être notablement inférieure à la probabilité des limites de chaque erreur considérée isolément, quelles que puissent être les autres.

C'est donc une défectuosité que d'assigner comme probabilité de l'erreur d'une inconnue, faisant partie d'un système à déterminer, celle qu'elle aurait si elle était seule, au lieu de donner des règles pour calculer la probabilité de l'ensemble des erreurs du système qui ne peuvent, en réalité, être isolées les unes des autres.

Ce défaut devient plus évident encore lorsqu'on fait attention que la probabilité relative à une limite d'erreur est d'autant plus grande que l'étendue bornée par cette limite est plus grande, de sorte que, si l'on veut conserver à l'ensemble des erreurs une probabilité un peu élevée, il faudra assigner à chaque erreur des limites plus écartées que celles qu'exigerait la même probabilité si l'inconnue était seule. Lors donc qu'on s'arrête à ces dernières limites, on n'a réellement qu'une très faible probabilité qu'elles ne sont pas dépassées. On verra, en effet, dans ce qui suit, qu'il suffit de deux inconnues pour que l'étendue des erreurs soit presque doublée pour une même probabilité. Les règles usuelles font naître par là des idées fort inexactes sur la connaissance des grandeurs inconnues qu'on s'est proposé de déduire des observations, et elles égarent même parfois sur la valeur de ces observations, ainsi que sur le nombre de bonnes observations qu'il serait indispensable de faire pour parvenir à un résultat assigné. [...]

La défectuosité une fois indiquée, on voit qu'il ne s'agit plus que de modifier convenablement la démonstration de la méthode, afin de calculer, non plus la probabilité de l'erreur d'une inconnue, mais la probabilité de l'ensemble des erreurs. Il y avait ici à choisir parmi les démonstrations.

À proprement parler, Legendre n'en a point donné. Seulement il appuie son procédé assez solidement sur les avantages qu'il fait ressortir ; et il insiste principalement sur ce que la moyenne arithmétique des observations, que l'on prend d'ordinaire avec confiance, n'est précisément qu'un cas particulier de la méthode des moindres carrés.

Ce fut M. Gauss qui rattacha le premier cette méthode au calcul des probabilités, quelques années après la publication de Legendre, et ce fut aussi dans un ouvrage d'astronomie : *Theoria motus corporum celestium*, in-4°, 1809. La démonstration que M. Gauss y donne repose sur la réciproque de la remarque de Legendre, que la moyenne arithmétique se déduit de sa méthode quand cette moyenne peut avoir lieu. Réciproquement, si la moyenne arithmétique adoptée généralement est nécessaire, dit M. Gauss, il en résulte qu'une certaine loi de probabilité est nécessaire, et que le seul procédé à suivre pour combiner les équations fournies par l'observation est le procédé qui rend un minimum la somme des carrés de ces équations. Mais on doit convenir que l'hypothèse de la nécessité de prendre la moyenne arithmétique d'une masse de faits, pour obtenir le résultat le plus exact possible, est tout à fait gratuite, et ne saurait pas plus être admise *a priori* que l'hypothèse même de la nécessité du minimum des carrés. Il n'y a donc là qu'une preuve restreinte aux cas particuliers où la loi de probabilité des erreurs, qui conduit toujours à la moyenne arithmétique dans les observations, et c'est ce qui arrive très souvent par la nature des choses. Mais comme on l'ignore le plus ordinairement, il n'existe pas là de démonstration vraiment solide. Aussi dans un ouvrage bien plus récent, consacré exclusivement à la méthode des moindres carrés, *Theoria combinationis observationum minimis erroribus obnoxiae*, in-4°, 1823, M. Gauss a-t-il fondé sur d'autres considérations l'emploi de cette méthode. Ce ne sont, toutefois, que des considérations et non des

preuves, et l'on ne trouve de démonstration réelle que dans les travaux antérieurs de Laplace.

C'est dans un mémoire, publié en 1811, que Laplace fait voir que, quand le nombre des observations est assez grand, les erreurs les plus restreintes, probablement, sont données par la méthode des moindres carrés. Ce mémoire se trouve parmi ceux de l'Académie des sciences pour 1811, et il a été reproduit dans la *Théories analytique des probabilités*, qui parut en 1812. Les principes de la démonstration de Laplace sont à l'abri de toute objection, les moyens analytiques peuvent seuls en soulever quelques-unes, ils n'ont point toute la rigueur désirable, et peut-être dans les applications conviendrait-il de bien discuter les cas particuliers qui pourraient se présenter. Mais lorsqu'on réfléchit qu'il suffirait le plus souvent de quelques observations ajoutées ou retranchées pour rendre aux expressions analytiques toute leur valeur, on reconnaît que l'analyse de Laplace satisfait complètement aux démonstrations générales. C'est cette analyse qui sera employée dans ce qui va suivre, pour suppléer l'omission de la probabilité du concours des diverses erreurs dans le cas de plusieurs inconnues, qui se présente presque toujours. Seulement il y sera fait quelques simplifications, tirées surtout des beaux travaux de Laplace et de Gauss. Si l'un et l'autre ont négligé cette considération du concours des événements, qui accroît si fortement la grandeur des erreurs possibles avec une probabilité déterminée, ils ont du moins fourni tous les moyens de la calculer. »

Heyde et Seneta<sup>1</sup> ont écrit un important ouvrage détaillant les travaux de Bienaymé. Par ailleurs, Lancaster<sup>2</sup> avait également exposé les calculs de Bienaymé.

Il résulte de ces publications le fait que les calculs du paragraphe 2 du papier de Bienaymé de 1852 anticipent les résultats de Pearson sur la distribution du  $\chi^2$  :

Si  $\{X_1, X_2, \dots, X_n\}$  sont  $n$  variables indépendantes et toutes  $\sim N(0,1)$ , alors la distribution  $\chi^2_n$  (avec  $n$  degrés de liberté) est définie comme celle de la somme  $X_1^2 + X_2^2 + \dots + X_n^2$ .

Introduite par Pearson en 1900, la loi du  $\chi^2$  avait en fait été préalablement étudiée par le géodésien allemand Friedrich Robert Helmert (1843-1917), auteur d'une *Théorie mathématique et physique de géodésie supérieure* (1880) relative à la théorie des erreurs.

La loi probabilité du  $\chi^2$  a pour densité :

$$f(x) = \frac{1}{2^{n/2} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}$$

où  $\Gamma$  désigne la fonction gamma d'Euler,  $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ , ( $\alpha > 0$ )

1. C. C. Heyde and E. Seneta, *I. J. Bienayme: Statistical Theory Anticipated*, Berlin, Springer, 1977.
2. Lancaster, H. O., *Forerunners of the Pearson Chi 2*, Aust. J. Stat., 8, 1966, p. 117-126.

# ANALYSE MATHÉMATIQUE.

SUR

## LES PROBABILITÉS

DES ERREURS DE SITUATION D'UN POINT;

PAR A. BRAVAIS.

La détermination du lieu occupé par un point dans l'espace dépend d'un certain nombre d'éléments au moins égal à trois; il en est de même lorsque le point est assujéti à se trouver sur un plan donné, et, dans ce cas, le nombre des éléments déterminants doit être au moins égal à deux. L'observation de ces éléments consiste essentiellement en des mesures d'angles ou de longueurs, et quelquefois de durées. Les erreurs inévitables de ces diverses mesures devant réagir sur la situation du point, j'ai recherché l'influence qu'exerce sur ces *erreurs possibles de situation* la possibilité des erreurs élémentaires, concurremment avec les coefficients de la méthode employée: toutefois, je me suis restreint au cas où les éléments sont déterminés par un grand nombre d'observations.

Si nous nommons  $a, b, c, \dots$  les éléments observés;  $x, y, z$ , les coordonnées du point, nous avons en général

$$\left. \begin{aligned} x &= \varphi(a, b, c, \dots) \\ y &= \psi(a, b, c, \dots) \\ z &= \chi(a, b, c, \dots) \end{aligned} \right\} \quad (1)$$

Les fonctions  $\varphi, \psi, \chi$  sont la traduction analytique des rela-

Sa fonction de répartition est

$$F(x) = \int_0^x f(t)dt, (x \geq 0)$$

Son espérance est  $E(X) = n$ , sa variance  $\text{Var}(X) = 2n$ .

Son usage permet de confirmer ou infirmer avec un seuil de sûreté choisi (exprimé en termes de pourcentages), une hypothèse faite sur un phénomène aléatoire.

$$\Pr(\chi^2 < \alpha) = \frac{1}{2^{n/2} \Gamma\left(\frac{n}{2}\right)} \int_0^\alpha t^{\left(\frac{n}{2}-1\right)} e^{-t/2} dt.$$

Par ailleurs, on peut démontrer que si  $(N_1(n), N_2(n), \dots, N_k(n))$  est un *vecteur multinomial* de paramètres  $n$  et  $(p_1, p_2, \dots, p_k)$ , alors la loi de la variable aléatoire

$\sum_{i=1}^k \frac{(N_i - np_i)^2}{Np_i}$  converge vers une loi du  $\chi^2$  à  $k - 1$  degrés de liberté.

## 20.2. Les travaux antérieurs de Bravais

Il est à noter qu'Auguste Bravais avait soumis à l'Institut, en 1838, un mémoire intitulé *Analyse mathématique sur les probabilités des erreurs de situation d'un point*. Un rapport concernant ce papier fut alors rédigé par Poisson et Savary<sup>1</sup> mais la publication<sup>2</sup> du mémoire de Bravais n'eut lieu qu'en 1846. Cette publication contient la première exposition de la théorie de la corrélation et donne la surface normale de probabilité pour deux variables.

Le papier de Bravais traite le problème de la détermination d'un point dans un plan, comme l'avait entrepris Cotes. Il fait l'hypothèse que les coordonnées du point sont des fonctions connues d'un nombre fini d'éléments et que chaque élément a été estimé par un grand nombre d'observations si bien que, selon le théorème limite central de Laplace, chaque élément possède une distribution normale. Il écrit<sup>3</sup> :

« La détermination du lieu occupé par un point dans l'espace dépend d'un certain nombre d'éléments au moins égal à trois ; il en est de même lorsque le point est assujéti à se trouver sur un plan donné, et, dans ce cas, le nombre des éléments déterminants doit être au moins égal à deux. L'observation de ces éléments consiste essentiellement en des mesures d'angles ou de longueurs, et quelquefois de durées. Les erreurs inévitables de ces diverses mesures devant réagir sur la situation du point, j'ai recherché l'influence qu'exerce sur ces *erreurs possibles de situation* la possibilité des erreurs élémentaires, concurremment avec les coefficients de la

1. CRAS du 9 juillet 1838, p. 77-78.

2. Bravais A., *Analyse mathématique sur les probabilités des erreurs de situation d'un point*, Mém. Acad. R. Sci. Inst. Fr., 9, 1846, p. 255-332.

3. Page 255.

méthode employée : toutefois, je me suis restreint au cas où les éléments sont déterminés par un grand nombre d'observations. »

En négligeant les quantités du second ordre ou carrés des erreurs, Bravais linéarise les relations entre les coordonnées du point et les éléments et obtient donc un jeu d'équations où les erreurs d'estimation des coordonnées sont des fonctions linéaires des  $n$  erreurs indépendantes, normalement distribuées et ayant des variances connues.

Bravais démontre que les erreurs d'estimation des coordonnées ont des distributions multivariées normales, du moins pour les distributions bi et tri variées. Il ajoute<sup>1</sup> :

« D'après cette loi de formation, qui se trouve vérifiée pour deux et trois variables, il est probable que la même loi se continuerait dans le cas de quatre et même d'un nombre quelconque de variables. [...] Mais la démonstration générale de cette loi de formation m'est inconnue. »<sup>2</sup>

Un travail sur le même sujet avait été réalisé par Giovanni Plana<sup>3</sup> en 1818, mais Bravais l'ignorait.

### **20.3. Considérations à l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés**

Dans ce mémoire, Bienaymé a écrit<sup>4</sup> en 1853 :

« Il semble, au premier abord, qu'il y ait un lien mystérieux entre les probabilités et cette intégrale qui s'est présentée à M. Gauss dans ses premières recherches, comme une conséquence du principe de la moyenne arithmétique qu'il adoptait gratuitement, et qui est ensuite ressortie de l'analyse de Laplace, basée uniquement sur ce que les observations sont en grand nombre. Aussi, plus d'un savant a-t-il pensé qu'une connaissance mieux approfondie de la théorie des probabilités amènerait à reconnaître que la loi de probabilité des erreurs est représentée par la

fonction  $\frac{1}{\sqrt{\pi}} e^{-t^2}$ . Mais c'est là une de ces vues inexactes qui conduisent à des

conséquences erronées : et celle-ci en a produit en grand nombre, tant au point de vue théorique, que dans les résultats pratiques. Il suffira de dire ici qu'il n'y a pas de liaison nécessaire entre cette exponentielle et les lois de probabilités ; qu'elle

---

1. Page 301.

2. Cf. A. Hald, *A History of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 504-506.

3. Plana G.A.A., *Allgemeine Formeln, um nach der Methode der kleinsten Quadrate die Verbesserungen von 6 Elementen zu berechnen, and zugleich das jeder zukommenden Gewicht zu bestimmen*, Z. Astron. Verw. Wiss., 6, 1818, p. 249-264.

4. Bienaymé J., *Considérations à l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés*, CRAS, 37, 1853, p. 309-324.



(Extrait des COMPTES RENDUS DES SÉANCES DE L'ACADÉMIE DES SCIENCES, tome XXXVII, séance du 29 août 1853.)

# CONSIDÉRATIONS

A L'APPUI DE LA DÉCOUVERTE DE LAPLACE

SUR LA LOI DE PROBABILITÉ DANS LA MÉTHODE DES MOINDRES CARRÉS ;

PAR M. BIENAYMÉ.

Dans la séance du 8 août 1853, l'Académie a entendu des remarques, des réflexions faites par M. Cauchy et par M. Le Verrier sur la méthode des moindres carrés. Ces réflexions n'ont pas été insérées dans le *Compte rendu*, à mon grand regret, car elles m'avaient paru d'un intérêt véritable. Il n'a été publié que le *Mémoire* de M. Cauchy, dont la lecture avait fourni l'occasion de cette espèce de discussion ; et je n'ai point vu reproduite une opinion qui m'avait semblé le diriger dans ses observations verbales. M. Cauchy avait nié l'exactitude du résultat si remarquable, découvert et démontré par Laplace, et qui consiste en ce que la méthode des moindres carrés s'applique aux données des observations, quelle que soit la loi de probabilité des erreurs. J'ai cru pouvoir annoncer, alors, que j'aurais quelques bonnes raisons peut-être à présenter à l'appui de l'opinion de Laplace. Je viens les exposer le plus brièvement possible ; quoique je sois en droit de faire remarquer cependant que les divers *Mémoires* ou morceaux de *Mémoires* publiés par notre savant confrère, dans les nos 5, 6 et 7 du *Compte rendu*, ne justifient nullement son assertion. Loin de là, selon moi, il y a telles parties de cette analyse si féconde, et parfois très-ingénieuse, qui, avec bien peu de changements, démontreraient pleinement la découverte de Laplace. Et si je me borne à l'indiquer, c'est que je ne veux pas même avoir l'air d'en faire la critique.

n'est qu'un moyen d'approximation très commode, mais qui pourrait être remplacé par d'autres formules ; que ce qui la ramène sans cesse, c'est qu'elle est très propre à représenter une fonction aux environs du maximum, mais qu'elle ne la représente qu'à peu près ; et que même dans les questions où elle offre le plus de facilité comme approximation, elle donne fréquemment des résultats dont la fausseté est manifeste dès qu'on veut l'employer à des raisonnements un peu complexes, au lieu de la tenir pour ce qu'elle est réellement ; c'est-à-dire pour une pure machine arithmétique, bonne aux calculs numériques.

Mais il n'en est pas ainsi de la moyenne des carrés des différences des erreurs à leur moyenne. Ce n'est pas un élément arbitraire de l'approximation ; ni, comme le croyait M. Gauss, une mesure arbitraire de la précision, à laquelle on pourrait substituer toute autre moyenne de puissances de degré pair. Tout au contraire, la moyenne des carrés renferme la condition fondamentale du développement des probabilités à mesure que les observations se multiplient ; et s'il n'est pas permis de dire *a priori* qu'on ne saurait trouver quelque fonction plus avantageuse (autre que les moyennes de degré pair), on peut affirmer que dans l'hypothèse d'une telle découverte la fonction pourrait être remplacée par un emploi convenable de la moyenne des carrés. »

## 20.4. La controverse Bienaymé-Cauchy

Nous avons indiqué dans les rappels du paragraphe 20.1. que le système des multiplicateurs  $k_{ij}$  utilisés par Legendre, Gauss et Laplace pour traiter la méthode des moindres carrés constitue en fait la recherche d'une matrice  $K$  de  $r \times n$  éléments  $K = \{k_{ij}\}$ , telle que l'on obtienne  $KM = I$ ,  $I$  étant la matrice unité.

Cauchy avait abordé le problème de la minimisation des erreurs en 1831<sup>1</sup>. En 1853, il entreprend l'étude du système de multiplicateurs afin de montrer qu'il y a des situations où le choix des moindres carrés n'est pas optimal. Il écrit<sup>2</sup> :

« Si l'on nomme  $\nu$  la limite au-dessus de laquelle on veut abaisser l'erreur  $\xi$  de l'inconnue  $x$ , et  $P$  la probabilité de la coïncidence de cette erreur avec une quantité comprise entre limites  $-\nu$  et  $+\nu$ , le système de facteurs correspondant à la plus grande valeur de  $P$  sera ordinairement, pour des valeurs de  $\nu$  suffisamment grandes, très différent de celui qui donnerait la méthode des moindres carrés, quel que soit d'ailleurs le nombre  $n$  des quantités fournies par l'observation, et quelle que soit la limite  $\chi$  assignée aux erreurs que comportent ces mêmes quantités. »

Bienaymé considéra que les travaux de Cauchy étaient une critique des résultats de Laplace et une controverse naquit dans les pages des Comptes

1. A. L. Cauchy, *Mémoire sur le système de valeurs qu'il faut attribuer à divers Éléments, déterminés par un grand nombre d'observations, pour que la plus grande des erreurs, abstraction faite du signe, devienne un minimum*, Jal Polytechnique, cahier 20, p. 175-248, 1831.
2. CRAS, 37, p. 332.

Rendus à l'Académie des Sciences<sup>1</sup>. On trouvera dans le livre de Heyde et Seneta cité plus haut (en référence 2) une analyse complète du travail de Cauchy. La plupart des mémoires de Cauchy concernant les probabilités ont été publiés dans les Comptes Rendus de l'Académie des Sciences de Paris (CRAS) en 1853.<sup>2</sup>

## 20.5. L'inégalité de Bienaymé-Chebyshev

L'inégalité de Bienaymé apparaît dans le mémoire intitulé *Considérations à l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés*<sup>3</sup>.

En voici une démonstration :

Pour une variable aléatoire  $X$  dont l'espérance mathématique est  $m = E(X)$ , la variance  $\sigma^2$ , et la densité de probabilité  $f(t)$ , l'inégalité de Bienaymé-Tchebychev s'écrit :

$$P(|X - m| > \lambda\sigma) \leq \frac{1}{\lambda^2}.$$

Elle peut se démontrer de diverses manières, dont celle-ci :

$$\sigma^2 = E((X - m)^2) = \int_{-\infty}^{\infty} (t - m)^2 f(t) dt$$

Quel que soit  $\epsilon > 0$

$$\sigma^2 = \int_{-\infty}^{m-\epsilon} (t - m)^2 f(t) dt + \int_{m-\epsilon}^{m+\epsilon} (t - m)^2 f(t) dt + \int_{m+\epsilon}^{\infty} (t - m)^2 f(t) dt$$

1. I. Bienaymé, *Remarques sur les différences qui distinguent l'interpolation de M. Cauchy de la méthode des moindres carrés, et qui assurent la supériorité de cette méthode*, CRAS, 37, 5-13, 68-69, 197-198, 206.
2. Voici une liste des travaux de Cauchy (en plus de la référence 11 ci-dessus) concernant les probabilités :
  - *Mémoire sur l'interpolation*, J. Math. Pures Appl. 2, 193-205 (1837), OC, II, 2, 5-17, 1835 ; *Mémoire sur l'évaluation d'inconnues déterminées par un grand nombre d'équations approximatives du premier degré*, CRAS, 36, 1114-1122, 1853.
  - *Mémoire sur l'interpolation, ou remarques sur les remarques de M. Jules Bienaymé*, CRAS, 37, 64-69, 1853.
  - *Sur la nouvelle méthode d'interpolation comparée à la méthode des moindres carrés*, CRAS, 37, 100-109, 1853.
  - *Sur les résultats moyens d'observations de même nature, et sur les résultats les plus probables*, CRAS, 37, 198-206, 1853.
  - *Sur la probabilité des erreurs qui affectent des résultats moyens d'observations de même nature*, CRAS, 37, 264-272, 1853.
  - *Sur la plus grande erreur à craindre dans un résultat moyen, et sur le système de facteurs qui rend cette plus grande erreur un minimum*, CRAS, 37, 326-334, 1853.
  - *Mémoire sur les résultats moyens d'un très grand nombre des observations*, CRAS, 37, 381-385, 1853.
3. CRAS, 37, p. 309-324, 1853.

$$\begin{aligned}
 &\geq \int_{-\infty}^{m-\varepsilon} (t-m)^2 f(t)dt + \int_{m+\varepsilon}^{\infty} (t-m)^2 f(t)dt \\
 &\geq \varepsilon^2 \left( \int_{-\infty}^{m-\varepsilon} f(t)dt + \int_{m+\varepsilon}^{\infty} f(t)dt \right) \\
 &\geq \varepsilon^2 P(|X-m| > \varepsilon).
 \end{aligned}$$

En posant  $\varepsilon = \lambda\sigma$ , on a

$$P(|X-m| > \lambda\sigma) \leq \frac{\sigma^2}{\lambda^2 \sigma^2} = \frac{1}{\lambda^2}.$$

Cette inégalité peut s'écrire, pour tout  $a > 0$  :

$$P(|X-m| > a) \leq \frac{\sigma^2}{a^2}.$$

### **Égalité de Bienaymé**

L'Égalité de Bienaymé, qui, selon lui, justifie à elle seule la méthode des moindres carrés, spécifie que la variance d'une somme de variables aléatoires indépendantes est égale à la somme des variances de ces variables.

En effet :

$$\begin{aligned}
 Var(\sum_i X_i) &= E\left((\sum_i X_i - E(\sum_i X_i))^2\right) \\
 &= E\left((\sum_i X_i - E(X_i))^2\right) \\
 &= \sum_{i,j} E\left((X_i - E(X_i)) \cdot (X_j - E(X_j))\right) \\
 &= \sum_i Var(X_i).
 \end{aligned}$$

## **20.6. Biographie de Bienaymé (1796-1878)**

Irénée Jules Bienaymé est né à Paris le 28 août 1796. Les périples familiaux font qu'il est d'abord inscrit au lycée de Bruges. De retour à Paris, il entre au lycée Louis le Grand. Il réussit le concours d'entrée à l'École Polytechnique en 1814, puis participe lors de cette même année à la défense de Paris. Il intègre finalement l'École en 1815, mais il fait partie de la promotion qui est exclue l'année suivante par Louis XVIII à cause de ses sympathies bonapartistes. En 1818, Bienaymé devient professeur de mathématiques à l'Académie militaire de Saint-Cyr, puis, deux ans plus tard rejoint l'administration des finances. Il est rapidement nommé inspecteur et deviendra inspecteur général en 1834.

Après la révolution de 1848, il est écarté de l'administration « pour manque d'esprit républicain ». Il devient professeur de probabilités à la Sorbonne mais perd son poste en 1851. Il occupe alors un rôle de conseiller, expert en statistique, pour le gouvernement de Napoléon III. En 1852, Bienaymé est élu à l'Académie des sciences et sera pendant 23 ans examinateur pour la remise du prix de statistiques. Il est membre fondateur de la Société mathématique de France et en assure la présidence en 1875.

Bienaymé n'est pas un mathématicien très prolifique : il n'a publié au cours de sa carrière que 23 articles, dont certains sont parus dans des revues peu connues. Ses premiers travaux concernent la démographie et les tables de vie. Il étudie ensuite la taille des jurys et la majorité nécessaire à la conviction. En disciple de Laplace, il prend sa défense dans diverses polémiques. Bienaymé, très influencé par la *Théorie analytique des probabilités* de Laplace, généralise sa méthode des moindres carrés.

Bienaymé est également un linguiste distingué. Il traduit du russe au français les travaux de son ami Pafnuti Chebyshev. Bienaymé publie, en 1869, ce qui deviendra la fameuse « inégalité de Bienaymé-Tchebychev » qui fournit une démonstration simple et précise de la loi des grands nombres. Bienaymé est également en contact épistolaire avec Quetelet, pionnier des études démographiques, ainsi qu'avec Cournot et Lamé.

Bienaymé est mort à Paris le 19 octobre 1878.

## 20.7. Principaux mémoires de Bienaymé

- « Mémoire sur la probabilité des erreurs d'après la méthode des moindres carrés », in *Journal de Liouville* 1, 17, 1852, p. 33-78 et *Mémoires Ac. Sc.*, Institut de France, tome XV, 1858, p. 615-663.
- *Considérations à l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés*, CRAS, 37, p. 309-324. 1853.
- *Remarques sur les différences qui distinguent l'interpolation de M. Cauchy de la méthode des moindres carrés, et qui assurent la supériorité de cette méthode*, CRAS, 37, 5-13, 68-69, 197-198, 206, 1853.
- *Application d'un théorème nouveau du calcul des probabilités*, CRAS, 81, 417-423, 1875.

## 20.8. Références

- C. C. Heyde and E. Seneta, *I. J., Bienayme: Statistical Theory Anticipated*, Berlin, Springer 1977.
- E. Seneta, *Round the historical work on Bienaymé*, Austral. J. Statist., 21 (3), p. 209-220, 1979.
- A. Gatine, « Bienaymé », in *École Polytechnique. Livre du centenaire 1794-1894*, III (Paris, 1897), p. 314-316.
- E. Czuber, *Zur Theorie der Fehlerellipse. Sitzungsberichte der Kgl. Akademie der Wissenschaften in Wien*, Vienne, 82, 698-723, 1880.
- P. Nékrassov, *Détermination des inconnues par la méthode des moindres carrés dans les cas où le nombre d'inconnues est considérable*, Société Mathématique de Moscou, 12, 189-204, 1885.



# Boole

George Boole (1815-1864), est un mathématicien britannique surtout connu comme le créateur de la logique classique, fondée sur une structure algébrique, appelée aujourd'hui algèbre de Boole, mais il a également abordé le domaine des probabilités.

## 21.1. Les diverses publications de Boole

Dès 1842, Boole étudie la théorie des invariants, fort utile depuis en physique relativiste, et publie en deux parties dans le *Cambridge Mathematical Journal*<sup>1</sup> ses résultats sur ce sujet sous le titre *Exposition of a general theorie of Linear Transformations*.

Boole écrit aussi deux traités didactiques qui seront utilisés jusqu'à la fin du XIX<sup>e</sup> siècle et plusieurs fois réédités : Son *Treatise on differential equations* (Cambridge, 1859) et son *Treatise on the calculus of finite differences* (Cambridge, 1860).

En plus de deux grands ouvrages sur la logique<sup>2</sup>, Boole rédige une cinquantaine de mémoires scientifiques dont une dizaine sur les probabilités<sup>3</sup>. Boole considère que la probabilité d'une conclusion est relative aux fondements sur lesquels elle est basée. Il écrit dans son article *On a general method in the theory of probabilities*<sup>4</sup> « La définition de la probabilité implique qu'une probabilité est toujours liée à notre niveau actuel d'information et varie avec ce niveau ». Plus tard il affirme dans *On the application of the theory of probabilities to the question of the combination of testimonies or judgements*<sup>5</sup> : « Notre estimation de la probabilité d'un événement ne varie pas absolument avec

1. Vol. 3, p. 1-20 et 106-119.

2. Les deux ouvrages majeurs de Boole relatifs à la logique mathématique sont : *The mathematical analysis of Logic, being an essay towards a calculus of deductive reasoning* (Cambridge, 1847) et *An investigation of laws of thought...* (Londres, 1854).

3. Sans compter *An investigation of the laws of thought*, dont l'édition originale comporte cent cinquante-cinq pages sur la théorie des probabilités, sur un total de quatre cent vingt-quatre.

4. Phil. Mag. 4th Series, 1854.

5. Edin. Phil. Trans., vol. XXI, p. 597-602, 1857.



*Boole*

**AN INVESTIGATION**  
**OF**  
**THE LAWS OF THOUGHT,**  
**ON WHICH ARE FOUNDED**  
**THE MATHEMATICAL THEORIES OF LOGIC**  
**AND PROBABILITIES.**

**BY**  
**GEORGE BOOLE, LL.D.**

PROFESSOR OF MATHEMATICS IN QUEEN'S COLLEGE, COBURG.

**LONDON:**  
**WALTON AND MABERLY,**  
**UPPER GOWER-STREET, AND IVY-LANE, PATERNOSTER-ROW.**  
**CAMBRIDGE: MACMILLAN AND CO.**

**1854.**

les circonstances qui affectent son occurrence mais avec notre connaissance de ces circonstances. »

En 1851 Boole précise sa recherche dans son article *On the theory of probabilities, and in particular on Mitchell's problem of the distribution of fixed stars*<sup>1</sup> : « Il y a deux conditions pour la construction d'une méthode parfaite pour le calcul des probabilités : 1) la construction préalable d'une méthode générale pour déterminer la dépendance logique de toute proposition par rapport à une autre proposition donnée ou à un ensemble de propositions. 2) La déduction, à partir de cette expression de la relation entre leurs probabilités. »

Dans *The laws of thought*, Boole donne deux définitions de l'indépendance de deux événements. La première est adéquate : « Deux événements sont dits être indépendants quand la probabilité d'occurrence de chacun d'entre eux n'est pas affectée par notre espérance d'occurrence ou de non survenue de l'autre. » Mais dans une deuxième définition, il dit que deux événements sont indépendants sauf si nous savons qu'il y a une connexion invariable entre eux. Les deux définitions n'étant pas consistantes, l'erreur est signalée par H. Wilbraham dans son article *On the theory of chances developed in Professor Boole's Laws of thought*<sup>2</sup>. John Maynard Keynes (1883-1946), dans son livre *A Treatise on Probability*<sup>3</sup> déclare que les solutions données par Boole aux problèmes I à VI du chapitre XX des *Laws of thought* (que De Morgan a lui aussi contestées dans sa correspondance avec Boole) sont inexactes. Il faut cependant noter que l'exposition de la théorie des probabilités sous la plume de Boole a évolué avec le temps et les publications. La dernière d'entre elles *On the application of the theory of probabilities to the question of the combination of testimonies or judgements* de 1857, déjà citée plus haut, est la plus correcte et la meilleure. Boole présente son papier à la Royal Society d'Édimbourg et il obtient le prix Keith, la plus haute distinction décernée par cette institution.

Dans le domaine des probabilités, le nom de Boole reste cependant attaché à l'inégalité suivante<sup>4</sup> : Pour des événements  $A_i$  ( $1 \leq i \leq n$ )

$$P(\bigcup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i)$$

En utilisant la complémentation, cette inégalité peut s'écrire :

$$P(\bigcap_{i=1}^n \bar{A}_i) \geq 1 - \sum_{i=1}^n (1 - P(\bar{A}_i))$$

- 
1. London, Edinburg and Dublin Phil. Mag, suppl to vol. 1, 4th series, 1851, p. 521-530.
  2. Phil. Mag. 4th series, vol. VII, p. 465, 1854.
  3. Londres, 1921.
  4. Cf. C.C. Heyde, E. Seneta, *Statisticians of the centuries*, Springer N.Y. 2001.

Cette inégalité reste valable sans que les événements soient mutuellement exclusifs ou indépendants. Elle est la première d'une suite d'inégalités de complexités croissantes appelées inégalités de Bonferroni (1892-1960), bien que la première soit due à Boole. Stigler observe avec humour, en appelant sa loi : *loi de Stigler*, que le nom d'une découverte est rarement attaché à son auteur<sup>1</sup>.

Pour revenir aux prolongements des *Laws of thought*, on sait que la logique binaire de Boole est universellement utilisée en informatique. Les contributions ultérieures de John von Neumann (1903-1957) et Alan Turing (1912-1954) seront décisives dans ce domaine de même que celle de Claude Shannon (1916-2001). Shannon publie son mémoire *A Mathematical Theory of Communication* dans le *Bell System Technical Journal* en 1948. Cet article pose les bases de la théorie de l'information et propose de transmettre des mots, des sons ou des images sous forme d'une suite de signaux d'amplitudes 0 ou 1 sur une liaison électrique filaire. Il définit dans ce papier les notions de quantité d'information, de taux de transmission, de codes de communication et de fiabilité avec laquelle l'information peut être transmise en présence de bruit.

Enfin, la logique des classes (ou ensembles), conçue par Boole aura un énorme développement ultérieur : « L'algèbre de Boole, qui trouve en grande partie son origine dans cet ouvrage classique, a eu une influence de plus en plus grande dans toutes les branches des mathématiques. De nos jours, les généralisations qu'elle a connues jouent un rôle important en topologie générale, en géométrie projective, dans la théorie des algèbres abstraites, en analyse fonctionnelle et en théorie ergodique générale, ainsi que dans la constitution de relais électriques, sans oublier les applications logiques qui en étaient le but initial. Qui peut dire quels autres usages lui seront trouvés dans le prochain siècle ! » écrivait Garrett Birkhoff en 1954, à l'occasion du centenaire des *Laws of thought*.

Une algèbre de Boole est définie par un ensemble  $B$ , auquel sont associées deux opérations binaires  $\cup$  et  $\cap$  (union et intersection). Cette algèbre vérifie par définition les axiomes suivants :

- $\cup$  et  $\cap$  sont des opérations binaires commutatives. Pour deux éléments  $x$  et  $y$  de  $B$ , on a :

$$x \cup y = y \cup x \quad \text{et} \quad x \cap y = y \cap x$$

- $\cup$  et  $\cap$  sont deux opérations distributives l'une par rapport à l'autre. Pour trois éléments  $x, y$  et  $z$  de  $B$ , on a :

$$x \cup (y \cap z) = (x \cup y) \cap (x \cup z) \quad \text{et} \\ x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$$

Chacune des opérations  $\cup$  et  $\cap$  possède un élément neutre dans  $B$ , qui sont notés 0 et 1, tels que :  $0 \cup x = x$  et  $1 \cap x = x$

1. S.M. Stigler, *Stigler's law of eponymy*, Transactions of the New York Academy of Sciences, 39, 1980, p. 147-157.

- À tout élément  $x$  de l'ensemble  $B$ , on associe un élément appelé complément de  $x$  que l'on peut noter  $x'$ , tel que  $x \cup x' = 1$  et  $x \cap x' = 0$ .

Boole est à l'origine de la notion d'ensemble et du calcul sur ces ensembles qu'il appelle classes d'objets.

## 21.2. Le *challenge Problem* de Boole et la controverse avec Cayley

Le problème a été proposé par Boole en 1851 dans le *Cambridge and Dublin math. Journal*<sup>1</sup>. Voici son énoncé :

« Les probabilités *a priori* de deux causes  $A_1$  et  $A_2$  sont  $c_1$  et  $c_2$  respectivement. La probabilité que si la cause  $A_1$  se produit, un événement  $E$  l'accompagnera est  $p_1$  et la probabilité que  $E$  accompagne  $A_2$ , si  $A_2$  se présente est  $p_2$ . De plus  $E$  ne peut se produire en l'absence des causes  $A_1$  et  $A_2$ . On demande la probabilité de l'événement  $E$ . »

Un résultat fut donné par Arthur Cayley (1821-1895) en 1853, dans le *Philosophical magazine*<sup>2</sup> quelques mois avant la publication des *Laws of Thought*. Boole déclara que les résultats de Cayley étaient faux et il développa lui-même une solution après avoir précisé qu'il avait reçu nombre de solutions de mathématiciens éminents, toutes différentes les unes des autres, et toutes fausses. La solution de Boole, après de longs calculs est que  $u$ , la probabilité de l'événement  $E$  est la racine de l'équation

$$\frac{[1 - c_1(1 - p_1) - u][1 - c_2(1 - p_2) - u]}{1 - u} = \frac{(u - c_1 p_1)(u - c_2 p_2)}{c_1 p_1 + c_2 p_2 - u}$$

qui est plus grande que  $c_1 p_1$  et  $c_2 p_2$  et plus petite que  $1 - c_1(1 - p_1)$ ,  $1 - c_2(1 - p_2)$  ou  $c_1 p_1 + c_2 p_2$ .

J. M. Keynes<sup>3</sup> a montré que la solution de Boole est fausse et il a donné la sienne.

Finalement Cayley publia en 1862 un papier accompagné d'une lettre de Boole faisant le point de leurs divergences<sup>4</sup>.

## 21.3. Biographie

George Boole naît le 2 novembre 1815 à Lincoln, dans l'est de la Grande-Bretagne. Il est le fils de John Boole et de Mary Ann Joyce. Son père exerce le métier de cordonnier mais s'intéresse à la fabrication d'instruments d'optique

1. 4<sup>th</sup> series, Suppl to vol. I, p. 521-530.
2. Phil. Mag., 4<sup>th</sup> series, vol. VI.
3. *A Treatise of probability*, London, Macmillan, 1921, page 187.
4. *On a Question in the Theory of Probabilities (Including a letter from Professor Boole)*, Philosophical Magazine, Series 4, volume 23, n° 155, 1862, pages 361-365.

et aux sciences. George est initié dans ces domaines par son père et fréquente une école primaire à Lincoln puis une école commerciale. Il apprend le latin auprès d'un libraire local et acquiert par lui-même des connaissances en Grec, Français, Allemand. À seize ans, il devient l'assistant d'un instituteur à Doncaster, situé à environ quarante kilomètres de Lincoln, puis, en 1835, il crée son propre établissement d'enseignement. Faute de moyens financiers, il ne peut aller à l'Université et il approfondit, en autodidacte, ses connaissances en mathématiques. Sa réputation locale devient telle, qu'alors qu'il n'a que dix-neuf ans, on lui demande de prononcer une allocution lors de l'inauguration d'un buste de Newton, natif lui aussi du Comté de Lincoln. Ce sera la première publication de Boole. Elle est intitulée *An address on the Genius and Discoveries of Sir Isaac Newton*. En 1841, il fait insérer dans le *Cambridge Mathematical Journal* qui vient d'être créé, ses *Researches on the theory of analytical transformations, with a special application to the reduction of the general equation of the second order*<sup>1</sup>. Le jeune Duncan Gregory (1813-1844) est l'un des éditeurs du journal de Cambridge et il va influencer Boole dans ses premiers travaux d'algèbre. En 1842, Boole a son premier contact avec Augustus De Morgan (1806-1871), qui est déjà un mathématicien renommé. Ils deviennent amis et le resteront toute leur vie comme le prouvent leurs correspondances. En 1844, Boole publie dans les *Philosophical Transactions of the Royal Society of London* une étude sur les équations différentielles et la théorie des opérateurs intitulée *On a general method in Analysis*<sup>2</sup> et une *Royal Medal* lui est décernée pour cet important travail. Boole et De Morgan rendent publics simultanément deux travaux relatifs à la logique fin novembre 1847 : *The mathematical analysis of Logic, being an essay towards a calculus of deductive reasoning* (Cambridge, 1847) de Boole et *Formal Logic, or the calculus of inference, necessary and probable* (Londres, 1847), de De Morgan. Tandis que De Morgan reste plutôt attaché à la tradition classique du syllogisme, Boole présente un traitement algébrique des concepts.

Bien qu'il ne possède aucun diplôme universitaire, Boole, en 1849, postule pour le poste de professeur de mathématiques au *Queen's College* de Cork en Irlande, nouvellement créé, et il l'obtient. Il occupera ce poste durant toute sa vie et gagnera rapidement la réputation d'être un enseignant exceptionnel. En 1855, il épouse Marie Everest (1832-1916), la nièce de Sir George Everest (dont le sommet de l'Himalaya porte le nom) qui lui donnera cinq filles. En 1857, il est élu Fellow à la Royal Society de Londres. Les Universités de Dublin et d'Oxford lui décernent aussi des diplômes honorifiques.

En 1864, il prend froid sous une averse en se rendant à son Collège et meurt le 8 décembre 1864 à Ballintemple près de Cork, d'une pneumonie, à l'âge de quarante-neuf ans.

1. Vol. 2, p. 64-73.

2. Vol. 134, p. 225-282.

## 21.4. Publications de Boole relatives aux probabilités

- *Proposed Questions in the Theory of Probabilities*, Cambridge and Dublin Math. Journal, 1852,
- *On the Theory of Probabilities, and in particular on Michell's Problem of the distributton of the fixed Stars*, Phil. Mag., 1861.
- *On a General Method in the Theory of Probabilities*, Phil. Mag., 1852.
- *On the Solution of a Question in the Theory of Probabilities*, Phil. Mag., 1854.
- *Reply to some Observations published by Mr. Wilbraham in the Phil. Mag.*, vii. p. 465, on *Boole's Laws of Thought*, Phil. Mag., 1854.
- *Further Observations in reply to Mr. Wilbraham*, Phil. Mag., 1854.
- *On the Conditions by which the Solutions of Questions in the Theory of Probabilities are limited*, Phil. Mag., 1854.
- *On certain Propositions in Algebra connected, with the Theory of Probabilities*, Phil. Mag., 1855.
- *On the Application of the Theory of Probabilities to the Question of the Combination of Testimonies or Judgments*, Edin, Phil. Trans, vol. XXL, p. 597-652, 1857.
- *On the Theory of Probabilities*, Roy. Soc. Proc., vol. XII, p. 179-184, 1362-1863.

## 21.5. Références

- T. A. Broadbent, *Dictionary of Scientific Biography*, Charles Scribner's sons, New York, vol. 2, p. 293-298, 1970.
- D. MacHale, *George Boole, His Life and Work*, Boole Press Dublin, 1985.
- R. Harley, *George Boole: an essay, biographical and expository*, London, 1866.
- E. T. Bell, *Men of mathematics*, New York, 1986.



# Chebyshev

## 22.1. Mémoire sur *Des valeurs moyennes*

Les deux publications les plus importantes de Chebyshev dans le domaine des probabilités sont :

*O Srednikh velichinakh* (1867) traduite en français sous le titre *Des valeurs moyennes*<sup>1</sup> et *O dvukh teoremakh otnositelno veroyatnostey*, (lue le 10 mars 1887) traduite sous le titre *Sur deux théorèmes relatifs aux probabilités*<sup>2</sup>.

La première de ces publications donne une démonstration précise de la loi généralisée des grands nombres.

Nous donnons ci-dessous les énoncés des théorèmes (sans la démonstration assez simple mais assez longue que l'on peut par exemple consulter dans les *Œuvres complètes*<sup>3</sup>) :

« Si nous convenons d'appeler *espérance mathématique* d'une grandeur quelconque, la somme de toutes les valeurs qu'elle est susceptible de prendre, multipliées par leurs probabilités respectives, il nous sera aisé d'établir un théorème très simple sur les limites entre lesquelles restera renfermée une somme de grandeurs quelconques.

### Théorème 1

Si l'on désigne par  $a, b, c, \dots$  les espérances mathématiques des quantités

$$x, y, z, \dots,$$

et par  $a_1, b_1, c_1, \dots$  les espérances mathématiques de leurs carrés

$$x^2, y^2, z^2, \dots$$

la probabilité que la somme

$$x + y + z + \dots$$

est renfermée entre les limites

$$a + b + c + \dots \pm \alpha \sqrt{a_1 + b_1 + c_1 + \dots - a^2 - b^2 - c^2 - \dots}, \text{ et}$$

1. *Jal de Liouville*, 2<sup>e</sup> série, 12, 1867, p. 177-184.

2. *Acta Mathematica*, t. XIV, 1890-1891, p. 805-815.

3. *Œuvres* de P. L. Tchebychef, A. A. Markov et N. Y. Sonin éd., 2 vols, St-Petersbourg : Académie impériale des sciences, 1899-1907.



$a + b + c + \dots - \alpha \sqrt{a_1 + b_1 + c_1 + \dots - a^2 - b^2 - c^2 - \dots}$  sera toujours plus grande que

$$1 - \frac{1}{\alpha^2}, \text{ quel que soit } \alpha . »$$

Indépendamment de Bienaymé, Chebyshev prouva donc, en 1867, l'inégalité qui porte leurs deux noms et que l'on écrit, avec les notations  $x_i, \mu_i, \sigma_i^2$  respectivement pour les variables, les valeurs moyennes et les variances :

$$P\{|\sum x_i - \sum \mu_i| < \alpha(\sum (\sigma_i^2)^{1/2})\} \geq 1 - \frac{1}{\alpha^2}, \text{ pour tout } \alpha > 0.$$

En 1874, Chebyshev participa au congrès de l'AFAS (Association française pour l'avancement des sciences) à Lyon et reconnut l'antériorité de Bienaymé. Dans le texte de son intervention publiée dans le *Journal de Liouville*<sup>1</sup>, il écrit :

« La démonstration simple et rigoureuse de la loi de Bernoulli, que l'on trouve dans ma note, sous le titre *Des valeurs moyennes*, n'est qu'un des résultats que l'on tire aisément de la méthode de M. Bienaymé, et d'après laquelle il est parvenu lui-même à démontrer une proposition sur les probabilités, d'où la loi de Bernoulli découle directement. »

Rappelons que l'inégalité de Bienaymé apparaît dans son mémoire intitulé *Considérations à l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés*<sup>2</sup>.

### **Théorème 2**

Si les espérances mathématiques des quantités

$$x, y, z, \dots, x^2, y^2, z^2, \dots$$

sont respectivement

$$a, b, c, \dots a_1, b_1, c_1, \dots,$$

la probabilité que la différence entre la moyenne arithmétique des  $N$  quantités  $x, y, z, \dots$  et la moyenne arithmétique des espérances mathématiques de ces quantités ne surpassera pas

$$\frac{1}{t} \sqrt{\frac{a_1 + b_1 + c_1 + \dots}{N} - \frac{a^2 + b^2 + c^2 + \dots}{N}}$$

sera toujours plus grande que

$$1 - \frac{t^2}{N},$$

quel que soit  $t$ .

### **Théorème 3**

Si les espérances mathématiques des quantités

$$U_1, U_2, U_3, \dots$$

et de leurs carrés

1. Sé, 19, p. 157-158, 1874.

2. CRAS, 37, p. 309-324, 1853.

$$U_1^2, U_2^2, U_3^2, \dots$$

ne dépassent pas une limite finie quelconque, la probabilité que la différence entre la moyenne arithmétique d'un nombre  $N$  de ces quantités et la moyenne arithmétique des espérances mathématiques sera moindre qu'une quantité donnée, se réduit à l'unité, quand  $N$  devient infini.

Dans l'hypothèse particulière que les quantités

$$U_1, U_2, U_3, \dots$$

se réduiront à l'unité ou à zéro, selon qu'un événement  $E$  a ou n'a pas lieu dans la 1<sup>re</sup>, 2<sup>e</sup>, ... Même épreuve, nous remarquerons que la somme

$$U_1 + U_2 + U_3 + \dots + U_N$$

donnera le nombre de *répétition* de l'événement  $E$  en  $N$  épreuves, et la moyenne arithmétique

$$\frac{U_1 + U_2 + U_3 + \dots + U_N}{N}$$

représentera le rapport du nombre de *répétition* de l'événement  $E$  au nombre des épreuves ».

L'application du théorème 3 amène Chebyshev à la conclusion suivante :

« Lorsque le nombre des épreuves devient infini, on obtient une probabilité, aussi rapprochée que l'on veut de l'unité, que la différence entre la moyenne arithmétique des probabilités de cet événement, pendant ces épreuves, et le rapport du nombre des répétitions de cet événement au nombre total des épreuves, est moindre que toute quantité donnée.

Dans le cas particulier où la probabilité de l'événement reste la même pendant toutes les épreuves, nous avons le théorème de Bernoulli. »

## 22.2. Mémoire *Sur deux théorèmes relatifs aux probabilités*

L'autre mémoire *O dvukh teoremakh otnositelno veroyatnostey*, traduit sous le titre *Sur deux théorèmes relatifs aux probabilités* est relatif à un rappel de la loi des grands nombres (rappel du théorème 3 donné plus haut) et à l'application de la méthode des moments pour l'extension aux séquences de quantités aléatoires indépendantes du théorème de la limite centrale de De Moivre et Laplace.

Chebyshev établit ce qu'il appelle un théorème « sur les probabilités à l'aide duquel la détermination des valeurs les plus sûres des inconnues, quand on a un grand nombre d'équations qui contiennent des erreurs accidentelles plus ou moins considérables, se ramène à la méthode des *moindres carrés*. »

Il énonce le théorème comme suit :

« Si les espérances mathématiques des quantités

$$u_1, u_2, u_3, \dots$$

sont toutes nulles et si les espérances mathématiques de toutes leurs puissances ne dépassent pas une limite finie quelconque, la probabilité que la somme

$$u_1 + u_2 + u_3, \dots + u_n$$

d'un nombre  $n$  de ces quantités, divisée par la racine carrée de la double somme des espérances mathématiques de leurs carrés, sera comprise entre deux limites quelconques  $t$  et  $t'$  se réduit à

$$\frac{1}{\sqrt{\pi}} \int_t^{t'} e^{-x^2} dx . »$$

## 22.3. Biographie de Chebyshev (1821-1894)

Pafnuty Lvovich Chebyshev est né le 16 mai 1821 à Okatovo, une petite ville de l'ouest de la Russie, située dans la province de Kaluga. Son père, Lev Pavlovich est un militaire qui a fait la campagne de Russie contre Napoléon. À la naissance de son fils, il est à la retraite et mène une vie paisible de propriétaire aisé. La mère de notre futur mathématicien, Agraфena Ivanova Pozniakova, s'occupe de sa nombreuse progéniture. La famille comptera, au total, neuf enfants. Pafnuty Lvovich est élevé dans le domaine familial et apprend la lecture avec sa mère.

En 1832, alors que Pafnuty n'a que onze ans, sa famille déménage pour Moscou. Il n'est toujours pas inscrit à l'école, mais son éducation se renforce par la venue d'un excellent précepteur en mathématique et physique, P.N. Pogorelski<sup>1</sup>. Lorsque Chebyshev entre à l'université de Moscou, en 1837, il vient tout juste d'avoir dix-sept ans, mais il possède déjà une solide formation scientifique. De plus, il a la chance d'avoir comme professeur de mathématiques appliquées Nikolai Dmitrievich Brashman. Ce dernier, en plus de ses cours sur la construction mécanique et sur l'hydraulique, enseigne à ses étudiants la théorie de l'intégration des fonctions algébriques et le calcul des probabilités. Chebyshev reconnaîtra toujours la grande influence que Brashman a eue sur lui durant ses années d'université et dans le choix de ses axes de recherches et se référera souvent à leurs « précieux entretiens privés ». Chebyshev est un étudiant brillant dans le département de physique et mathématiques. Il rédige en 1840 un mémoire sur le calcul des racines des équations par une méthode itérative originale. En 1841, il obtient l'équivalent de la licence. Mais cette même année sa situation financière change drastiquement. La famine en Russie éclate, ses parents sont obligés de quitter la ville et ne peuvent plus aider leurs fils. Néanmoins, Pafnuty décide de poursuivre ses études et s'inscrit au Master, sous la direction de son maître Brashman.

1. P.N. Pogorelski est l'auteur de certains des livres de mathématiques élémentaires les plus populaires en Russie.

Vers la fin de 1842, Chebyshev écrit son premier article (qu'il rédige en français) sur les intégrales multiples. Il le soumet à Liouville et le papier paraît l'année suivante dans le *Journal des mathématiques pures et appliquées*. Son second article, toujours en français, porte sur des problèmes de convergence de la série de Taylor et est édité en 1844 dans le journal de Crelle, *Journal für die reine und angewandte Mathematik*. Ses articles commencent à lui apporter une reconnaissance internationale.

En 1846 Chebyshev soutient sa thèse de Master intitulée : *Un essai sur une analyse élémentaire de la théorie des probabilités*. Il espère obtenir un poste à l'université de Moscou, mais hélas les opportunités sont rarissimes. Finalement en 1847, toujours dans l'espoir d'obtenir un poste, il soutient une autre thèse, cette fois sur *Les méthodes d'intégration utilisant les logarithmes*. Grâce à ce travail il est nommé à l'université de Saint-Petersbourg. En 1850 il obtient un poste de professeur extraordinaire, c'est-à-dire sans chaire.

En 1852, il voyage à Paris, Londres et en Allemagne. À Paris, Liouville et Hermite lui demandent de généraliser la méthode exposée dans sa thèse, ce qu'il fait, mais ses résultats seront publiés beaucoup plus tard. Il rencontre également Bienaymé, Lebesgue et Poncelet<sup>1</sup>. À Londres, il travaille avec Cayley et Sylvester, tandis qu'à Berlin il rencontre Dirichlet. Chebyshev réalise des travaux importants dans plusieurs domaines : théorie des nombres où il développe la notion de congruence (il sera récompensé par un prix de l'Académie des sciences) ; nombres premiers<sup>2</sup> etc. Il aime également combiner les aspects théoriques et appliqués des mathématiques et, très habile de ses mains, il conçoit plusieurs machines arithmétiques ou autres structures mécaniques. C'est d'ailleurs dans un article consacré à la mécanique qu'il a introduit les polynômes dits de Chebyshev<sup>3</sup>. Ce goût le porte même, en parallèle de ses cours à l'université<sup>4</sup>, à faire, entre 1852 et 1856, un enseignement pratique de mécanique au lycée Alexandre, à Tsarskoe Selo<sup>5</sup>, à

1. Tchebychev se rendra souvent en France, il y retournera en 1856 et 1864 et interviendra entre 1873 et 1882 aux sessions de l'*Association Française pour l'Avancement des Sciences*, ce qui le conduira en particulier à Lyon en 1873, à Clermont-Ferrand en 1876, à Paris en 1878 et La Rochelle en 1882. Il retournera en France en 1884 et en 1893.
2. Il démontre en particulier, la conjecture de Joseph Bertrand, suivant laquelle il existe toujours un nombre premier compris entre  $n$  et  $2n$ , dès que  $n > 2$ . Il étudie également le nombre  $P(n)$  de premiers inférieurs à  $n$ . Confortant une conjecture de Gauss, il obtient que si  $(P(n)\log(n))/n$  admet une limite quand  $n$  tend vers l'infini, alors cette limite est nécessairement 1.
3. Cet article s'intitule : *Théorie des mécanismes connus sous le nom de parallélogrammes*, 1854.
4. Chebyshev est nommé en 1860, titulaire d'une chaire de mathématiques, poste qu'il occupera pendant plus de vingt ans. En 1882 il quittera l'université pour se consacrer entièrement à la recherche.
5. Tsarskoe Selo signifie « village royal » : c'était la résidence d'été des Tsars. Depuis 1937 le nom de Tsarskoe Selo a été transformé en celui de Pushkin, du nom du célèbre poète russe.

25 km au sud de Saint-Petersbourg. En 1856, Chebyshev devient membre du Comité d'artillerie. Sa contribution y sera importante, en particulier dans le domaine de la balistique. Des modèles et schémas mécaniques de notre mathématicien ont été exposés à l'époque au Conservatoire National des Arts et Métiers de Paris, et certains s'y trouvent encore<sup>1</sup>.

Ses travaux le font admettre à l'Académie des sciences impériale de Saint-Petersbourg, d'abord comme adjoint (1856) puis comme membre (1858). L'Académie des sciences de Paris l'accueille en son sein en 1860 en tant que correspondant et le nommera associé étranger en 1874.

Chebyshev ne s'est jamais marié et a vécu seul dans une grande maison de dix pièces ! Il était riche, mais ses seules dépenses concernaient sa propriété. Il a eu une fille qu'il n'a jamais reconnue officiellement mais dont il a assuré le confort matériel.

Chebyshev est mort le 8 décembre 1894 à Saint-Petersbourg.

Les travaux de Chebyshev sur l'intégration des fonctions algébriques et sur la théorie des meilleures estimations des fonctions sont bien connus.

Lyapunov, qui a suivi les conférences de Chebyshev dans les années 1870 a déclaré : « Ses cours n'étaient pas volumineux, et pour lui le bon critère n'était pas la quantité de connaissance fournie ; il aspirait plutôt à rendre clair les aspects les plus importants des problèmes qu'il traitait. Ses conférences étaient animées et passionnantes ; les remarques sur la signification et l'importance de certains problèmes, ainsi que sur les méthodes scientifiques pour les résoudre, étaient toujours abondantes [...]. Ses conférences étaient très stimulantes, les étudiants bénéficiaient à chaque fois de quelque chose de nouveau et d'essentiel ; il enseignait avec un grand recul et avait des points de vue inhabituels. »

Les études de Chebyshev sur les théorèmes limites des probabilités ont été poursuivies et développées par ses successeurs bien connus, Markov, Lyapunov et Bernstein.

## 22.4. Publications de Chebyshev

- *Œuvres* de P. L. Tchebychef publiées par les soins de MM. A. Markoff et N. Sonin, Saint-Petersbourg, Académie des sciences, 1899-1907.
- *Démonstration élémentaire d'une proposition générale de la théorie des probabilités*, Crelle's J. Reine Angew. Math., 33, 259-267, 1846.

---

1. Chebyshev était passionné par les machines à calculer. En particulier, un de ses articles, *Une machine arithmétique à mouvement continu*, est resté fameux.

- *Sur l'interpolation par la méthode des moindres carrés*, Mém. Acad. Sci., Saint-Pétersbourg, (7) 1(15), 1-24, 1859.
- *Sur le développement des fonctions à une seule variable*, Bull. Phys. Math. Acad. Sci., Saint-Pétersbourg, 1, 193-200, 1859.
- *Sur l'interpolation*, Proc. Acad. Sci., Saint-Pétersbourg, 4, n° 5. 1864.
- *Des valeurs moyennes*, J. de Liouville, Math. Pures Appl. [2] 12, 177-184, 1867.
- *Sur l'interpolation des valeurs équidistantes*, Proc. Acad. Sci., Saint-Pétersbourg, 25, n° 5, 1875.
- *Sur deux théorèmes relatifs aux probabilités*, Bull. Phys.-Math. Acad. Sci., Saint-Pétersbourg, 55, 1887.

## 22.5. Références

- A. P. Yousschkevitch, *Chebyshev*, Dict. of Sci. Biog., tome 3, Scribners, New York, 1970, p. 222-232.
- O. B. Sheynin, *Chebyshev's lectures on the theory of probability*, AHES, vol. 46, p. 321-340, 1994.



# Bertrand

## 23.1. Le cours de *Calcul des probabilités*

Le premier travail de Bertrand relatif aux probabilités fut la traduction, en 1855, de l'ensemble des publications de Gauss consacrées aux moindres carrés. Ce travail parut sous le titre *Méthode des moindres carrés par C. F. Gauss. Traduite par J. Bertrand*<sup>1</sup>. Il est extrêmement utile pour la compréhension des mémoires de Gauss, car la plupart avaient été publiés en Latin.

Le principal ouvrage de J. Bertrand sur les probabilités, résumé de leçons faites au Collège de France, ne fut publié que beaucoup plus tard, en 1888<sup>2</sup> sous le titre *Calcul des probabilités*. Ce livre s'adresse, plus ou moins, à des lecteurs ayant déjà quelques notions de l'histoire des probabilités, car Bertrand critique en permanence les travaux de ses prédécesseurs, les contredit à plaisir et ne cite pas toujours ses sources.<sup>3</sup>

Bernard Bru déclare fort justement :

« Ce livre, hautement apprécié des maîtres de l'École mathématique française de la Belle époque, Darboux, Poincaré ou Borel, et sans doute l'ouvrage "didactique" le plus commenté et le plus copié entre 1888 et 1940, a subi après la guerre un double discrédit, celui très général de presque tous les traités mathématiques français, dont le style trop littéraire ne correspondait plus aux rigueurs algébriques du temps, aggravé de celui de presque tous les traités statistiques préfishériens, peu conformes à la méthode statistique ou aux méthodes statistiques qui se sont imposées depuis la guerre. »

À la page VI de la préface de son livre, Bertrand refuse, dès l'abord, de définir le hasard :

« Comment oser parler des lois du hasard ? Le hasard n'est-il pas l'antithèse de toute loi ? En repoussant cette définition, je n'en proposerai aucune autre. »

Mais en fait, par des exemples, il précise sa conception des *lois du hasard* dans sa préface.

1. Paris, Mallet-Bachelier, 1855.
2. J. Bertrand, *Calcul des probabilités*, Paris, Gauthier-Villars, 1888
3. Cf. B. Bru, *Les leçons du Calcul des probabilités de Joseph Bertrand*, Jal électronique d'histoire des probabilités et de la statistique, vol. 2, n° 2, décembre 2006.



« Sur le Pont-Neuf », écrit-il, « pendant une journée ou pendant une heure, on peut prédire résolument que les passants de taille inférieure à deux mètres l'emporteront par le nombre. Le pont écarte-t-il les géants ? Quand, au jeu de dés, on annonce quelles combinaisons prévaudront, c'est, comme pour les passants du Pont-Neuf, une question d'arithmétique : les combinaisons qu'on ose exclure forment, dans le nombre total, si les épreuves sont nombreuses, une proportion beaucoup moindre que, parmi les Parisiens, les hommes de six pieds de haut. »

Par ailleurs, les fondements des probabilités dites continues ou géométriques semblent mal définis pour Bertrand ; car la définition classique de la probabilité, c'est-à-dire le rapport du nombre de cas favorables au nombre de cas possibles est claire si les nombres qui y figurent sont finis ; elle l'est encore comme limite d'une suite convergente lorsque les deux nombres tendent vers l'infini. Mais elle ne l'est plus si, au départ, les nombres sont infinis. Pour démontrer son propos, Bertrand pose le problème des cordes d'un cercle, appelé depuis paradoxe de Bertrand.

Voici l'énoncé par Bertrand<sup>1</sup> de ce que nous appelons son paradoxe :

« On trace au *hasard* une corde dans un cercle. Quelle est la probabilité pour qu'elle soit plus petite que le côté du triangle équilatéral inscrit ?

On peut dire : si l'une des extrémités de la corde est connue, ce renseignement ne change pas la probabilité ; la symétrie du cercle ne permet d'y attacher aucune influence, favorable ou défavorable à l'arrivée de l'événement demandé.

L'une des extrémités de la corde étant connue, la direction doit être réglée par le hasard. Si l'on trace les deux côtés du triangle équilatéral ayant pour sommet le point donné, ils forment entre eux et avec la tangente trois angles de  $60^\circ$ . La corde, pour être plus grande que le côté du triangle équilatéral, doit se trouver dans celui des trois angles qui est compris entre les deux autres. La probabilité pour que le hasard entre trois angles égaux qui peuvent le recevoir le dirige dans celui-là semble, par définition, égale à  $1/3$ .

On peut dire aussi : si l'on connaît la direction de la corde, ce renseignement ne change pas la probabilité. La symétrie du cercle ne permet d'y attacher aucune influence, favorable ou défavorable à l'arrivée de l'événement demandé.

La direction de la corde étant donnée, elle doit, pour être plus grande que le côté du triangle équilatéral, couper l'un ou l'autre des rayons qui composent le diamètre perpendiculaire, dans la moitié la plus voisine du centre. La probabilité pour qu'il en soit ainsi semble, par définition, égale  $1/2$ .

On peut dire encore : choisir une corde au hasard, c'est en choisir au hasard le point milieu. Pour que la corde soit plus grande que le côté du triangle équilatéral, il faut et il suffit que le point milieu soit à une distance du centre plus petite que la moitié du rayon, c'est-à-dire à l'intérieur d'un cercle quatre fois plus petit en surface. Le nombre des points situés dans l'intérieur d'une surface quatre fois moindre est quatre fois moindre. La probabilité pour que la corde dont le milieu est choisi au hasard soit plus grande que le côté du triangle équilatéral semble, par définition, égale à  $1/4$ ,

1. *Calcul des probabilités*, Paris, Gauthier-Villars, 1888, p. 4-5.

Entre ces trois réponses, quelle est la véritable ? Aucune des trois n'est fausse, aucune n'est exacte, la question est mal posée. »

Ce « paradoxe », dit « de Bertrand », souligne la dépendance de la probabilité au mode de sélection d'une corde « au hasard ». Lorsque le mode est spécifié, le problème possède une solution bien définie. En l'absence d'une définition du mode de sélection, le terme « au hasard », dans « tracer une corde au hasard », est ambigu. Les trois solutions correspondent à des modes distincts et valables, et sont correctes dans chaque cas.

Nous avons déjà, dans divers chapitres, cité ou analysé des commentaires de Bertrand. Nous ajouterons simplement deux citations, l'une relative à l'application des probabilités aux décisions judiciaires<sup>1</sup> et l'autre, à nouveau, au Paradoxe de Saint-Petersbourg déjà abordé au chapitre Daniel Bernoulli :

« L'application du calcul aux décisions judiciaires est, dit Stuart Mill, le scandale des Mathématiques. L'accusation est injuste. On peut peser du cuivre et le donner pour or, la balance reste sans reproche. Dans les travaux sur la théorie des jugements, Condorcet, Laplace et Poisson n'ont pesé que du cuivre. La réunion, quelle qu'elle soit, qui peut juger bien ou mal, est remplacée dans leurs études par des urnes où l'on puise des boules blanches ou noires. [...] On peut assurément supposer le nombre de boules noires égal à celui des jugements mal rendus, les deux problèmes n'en restent pas moins fort différents et, pour tout dire, sans analogie. »

En ce qui concerne le paradoxe de Saint-Petersbourg, Bertrand écrit, avec raison<sup>2</sup>, dans son ouvrage de 1888 :

« Quant au problème de Saint-Petersbourg, il faut approuver absolument et simplement la réponse réputée absurde. Pierre possède, je suppose un million d'écus et les donne à Paul en échange des promesses convenues. Il est fou ! dira-t-on. Le placement est aventureux mais excellent ; l'avantage infini est réalisable. Qu'il joue obstinément, il perdra 1 partie, 1 000, 1 000 millions, un million de milliards peut-être ; qu'il ne se rebute pas, qu'il recommence un nombre de fois que la plume s'userait à écrire, qu'il diffère surtout le règlement des comptes, la victoire pour lui est certaine, la ruine de Paul inévitable. Quel jour ? quel siècle ? On l'ignore ; avant la fin des temps certainement, le gain de Pierre sera colossal.

Une fourmi transporte un grain de poussière de la cime du mont Blanc dans la plaine, retourne sur la hauteur, descend une nouvelle charge et recommence toujours. Après combien de voyages aura-t-elle comblé les vallées et nivelé la chaîne des Alpes ? Le premier écolier, sans consulter l'arénaire d'Archimède, fera le calcul sans erreurs. Le dessein de la fourmi dépasse ses forces, s'écrieront des gens sages : elle mourra à la peine. Condorcet et Poisson ne sont pas moins sages. Pierre est un imprudent ; il entreprend au-delà de son crédit, une opération beaucoup trop longue ; il est aussi certain pourtant de ruiner Paul que la fourmi de niveler la Suisse. »

---

1. Page XLIII.

2. Page XI.

En ce qui concerne le problème de la ruine du joueur, Bertrand lui consacre le chapitre 6 de son traité.

B. Bru<sup>1</sup> rappelle qu'Ampère, dans sa toute première publication en 1802<sup>2</sup>, avait calculé avant Bertrand la probabilité qu'un joueur de fortune finie  $m$ , soit ruiné face à un adversaire infiniment riche au jeu de pile ou face, précisément à la  $n$ ième partie avec nécessairement  $n = m + 2p$ , pour un entier  $p$  donné. Cette probabilité est  $\frac{m}{p} C_{n-1}^{p-1} \cdot \frac{1}{2^n}$ .

Bertrand montre, de plus, que si  $n$  est très grand, la probabilité que le joueur soit ruiné avant la  $n$ ième partie est

$$1 - \frac{2}{\sqrt{\pi}} \int_0^{m/\sqrt{2n}} e^{-t^2} dt .$$

On retrouve la fonction  $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\zeta^2} d\zeta$ , définie dans notre chapitre 12 et qui est tabulée dans de nombreux ouvrages.

Signalons également le célèbre problème du scrutin de ballottage<sup>3</sup>. Bertrand écrit dans les CRAS :

« On suppose que deux candidats A et B soient soumis à un scrutin de ballottage. Le nombre des votants est  $\mu$ . A obtient  $m$  suffrages et est élu, B en obtient  $\mu - m$ . On demande la probabilité pour que, pendant le dépouillement du scrutin, le nombre des voix de A ne cesse pas une seule fois de surpasser celles de son concurrent.

La probabilité demandée est  $\frac{2m - \mu}{\mu}$ . La démonstration est fondée sur la formule

suivante qu'il est aisé de rendre évidente :

Si  $P_{m,\mu}$  désigne le nombre de combinaisons qui, dans le dépouillement du scrutin, sont favorables à l'événement demandé, on a

$$P_{m+1,\mu+1} = P_{m,\mu} + P_{m+1,\mu} .$$

L'expression générale de  $P_{m,\mu}$  se déduit de cette formule, mais il semble vraisemblable qu'un résultat aussi simple pourrait se démontrer d'une manière plus directe.

Si le nombre des votants est 60, il faut, que le candidat élu obtienne 45 voix pour que la probabilité de conserver la majorité pendant toute la durée du scrutin soit égale à  $1/2$ . »

Bertrand donne une démonstration complète de ce résultat au paragraphe 18 de son livre.

1. B. Bru, *Les leçons du Calcul des probabilités de Joseph Bertrand*, Jal électronique d'histoire des probabilités et de la statistique, vol. 2, n° 2, décembre 2006., p. 14.
2. A. M. Ampère, *Considérations sur la théorie mathématique du jeu*, Lyon, Perisse, 1802.
3. CRAS, vol. 105, p. 369, 1887.

Le chapitre VIII du *Traité de Bertrand*<sup>1</sup> est relatif à la *Loi des erreurs d'observations*.

Bertrand critique tout d'abord la théorie de Gauss de 1809. Il écrit :

« La loi rigoureuse de probabilité des erreurs d'observation varie sans doute avec la grandeur mesurée, comme avec le choix de l'instrument et l'habileté de l'observateur ; elle est inaccessible aux géomètres.

Euler, Bernoulli, Lagrange et Laplace ont fait des hypothèses démenties par les faits et mal justifiées par des preuves sans vraisemblance.

Gauss, plus heureux, a déduit d'un raisonnement fort simple une loi que la démonstration laisserait douteuse, mais que les conséquences justifient.

Lorsque plusieurs mesures d'une grandeur inspirent une confiance égale, la valeur la plus probable est la moyenne de celles qu'il a obtenues.

Tel est le postulatum de Gauss. [...] À la condition énoncée, il faut malheureusement, en ajouter plusieurs autres qu'on ne dit pas. Nous devons signaler d'abord une différence essentielle entre la valeur la plus probable d'une grandeur et la meilleure valeur à adopter.

La valeur la plus probable est celle dont la probabilité est la plus grande. [...]

En disant : "En présence de plusieurs mesures d'une même grandeur, le parti le meilleur est la valeur la plus probable", et : "La moyenne entre plusieurs mesures est la valeur la plus probable", on énonce deux propositions différentes. On a eu tort de les confondre.

Supposons, pour donner un exemple, que l'on cherche l'origine probable d'une plante d'espèce connue cueillie en France. Dans la liste des 100 localités où l'espèce se rencontre, on en trouve trois dans la même commune du Cantal et les autres dans 97 communes différentes du Finistère.

La commune la plus probable est en Auvergne ; l'origine probable de la plante, la Bretagne.

Après avoir proposé le postulatum qui a été admis sans difficulté, Gauss représente par  $\varphi(\Delta)d\Delta$  la probabilité pour que l'erreur d'une mesure soit comprise entre  $\Delta$  et  $\Delta + d\Delta$  ; la fonction  $\varphi(\Delta)$  est l'inconnue qu'il veut déterminer.

Ici encore s'élève une objection. La probabilité d'une erreur  $\Delta$  est-elle une fonction de  $\Delta$  ? »

Bertrand déclare donc que Gauss aurait dû chercher, non pas la condition pour que la moyenne soit la valeur la plus probable de la véritable valeur de la grandeur mesurée, mais la condition pour que la moyenne soit la valeur probable de cette véritable valeur.

Nous indiquerons au chapitre 24, la méthode utilisée par Poincaré<sup>2</sup> pour généraliser la démonstration de Gauss en tenant compte des critiques de Bertrand.

1. Pages 175-225.

2. H. Poincaré, *Calcul des probabilités*, Paris, Gauthier-Villars, 1912, chapitre X, p. 169-188,

Par ailleurs<sup>1</sup>, Bertrand critique également la méthode du maximum de vraisemblance.

In fine, Bertrand adhère<sup>2</sup> à la méthode décrite par Gauss en 1821-1823 dans *La théorie de la combinaison des observations* (*Theoria combinationis observationum erroribus minimis obnoxiae*)<sup>3</sup>.

On trouve, sous la plume d'Oscar Sheynin, une analyse<sup>4</sup> détaillée de l'ensemble des travaux de Bertrand.

## 23.2. Biographie de Bertrand (1822-1900)

Joseph Louis François Bertrand est né à Paris le 11 mars 1822. Son père, polytechnicien, est médecin et journaliste. Enfant prodige, Joseph satisfait, à l'âge de 11 ans, à toutes les épreuves d'entrée à l'École Polytechnique, mais ne peut y être admis, l'âge minimum étant de 16 ans. Entre onze et dix-sept ans, il obtient ses deux baccalauréats, sa licence et soutient une thèse de doctorat ès sciences sur la théorie mathématique de l'électricité. Il entre major à l'École Polytechnique en 1839, à 17 ans<sup>5</sup>, puis à l'École des Mines de Paris. Il est ensuite reçu premier au concours d'agrégation de mathématiques et devient professeur de mathématiques au Lycée Saint-Louis. Enfin il est nommé examinateur, puis professeur d'analyse à l'École Polytechnique (en 1852, à 30 ans) et maître de conférences de calcul différentiel et intégral à l'École Normale Supérieure. En 1862, il remplace Jean-Baptiste Biot à la chaire de physique mathématique du Collège de France. Entre-temps, il est élu membre de l'Académie des sciences en 1856, la même année que le mathématicien Charles Hermite, également né en 1822, qui est son beau-frère<sup>6</sup>. En 1874, à la mort d'Élie de Beaumont, il est élu secrétaire perpétuel de l'Académie des sciences et remplace, la même année, Jean-Baptiste Dumas à l'Académie française.

Trois des fils de Joseph Bertrand sont également entrés à Polytechnique : Marcel (1847-1907) qui intégra le corps des mines et qui deviendra un géologue de grande renommée, Joseph Désiré (1853-1929) et Léon Gratien (1858-1951).

Bertrand publie, dès 1843, à l'occasion de ses recherches d'analyse pure, de mécanique rationnelle et de physique mathématique, divers mémoires sur la théorie des surfaces (*Courbure des surfaces ; Normales principales des*

1. Pages 181-183.

2. Chapitre X, Théorie des moyennes ?, p. 247-258.

3. Cf. notre chapitre 16.

4. O. B. Sheynin, *Theory of probability, an historical essay*, N. G. Verlag, Berlin, 2005, § 11.

5. Fait unique dans les annales de l'École Polytechnique il y entre avec déjà le titre de docteur ès sciences.

6. L'une des deux filles de Charles Hermite épousera le grand scientifique Émile Picard.

*courbes à double courbure ; Tautochrones, etc.*). Sa publication *Surface des ondes* est importante pour l'optique. On lui doit aussi un monumental *Traité de Calcul différentiel et intégral*, des leçons sur le *Calcul des probabilités*, une *Théorie mathématique de l'électricité*, des leçons sur la *capillarité*, et un cours de *Thermodynamique*. Il est également l'auteur de plusieurs biographies de grands physiciens. En 1845, en analysant une table de nombres premiers jusqu'à 6 000 000, il fait la conjecture qu'il y a toujours au moins un nombre premier entre un nombre  $n$  et son double (en fait  $2n - 2$ ) pour tout  $n$  plus grand que 3. Pafnuti Chebyshev démontrera cette conjecture en 1850.

Joseph Bertrand est mort à Paris le 3 avril 1900, à l'âge de 77 ans.

### 23.3. Principaux ouvrages de Bertrand

- *Traité d'arithmétique*, Paris, Hachette, 1849.
- *Traité de calcul différentiel et intégral*, Gauthier-Villars, 1864-1870, (2 volumes).
- *Méthode des moindres carrés*, Mallet-Bachelier, 1855.
- *Les fondateurs de l'astronomie moderne : Copernic, Tycho Brabé, Kepler, Galilée, Newton*, J. Hetzel, 1865.
- *Arago et sa vie scientifique*, J. Hetzel, 1865.
- *Thermodynamique*, Gauthier-Villars, 1887.
- *Calcul des probabilités*, Gauthier-Villars, 1888.
- *Leçons sur la théorie mathématique de l'électricité professées au Collège de France*, Gauthier-Villars, 1890.
- *Blaise Pascal*, C. Lévy, 1891.

### 23.4. Références

- D. J. Struik, *Joseph Bertrand*, Dictionary of Scientific Biography, Charles Scribner's sons, New York, vol. 2, p. 87-89, 1970.
- O. B. Sheynin, *Bertrand's work on probability*, AHES, vol. 48, p. 155-199, 1994.



# Poincaré

Henri Poincaré a enseigné le calcul des probabilités à la Sorbonne de 1886 à 1896. Le cours qu'il a professé pendant l'année scolaire 1893-1894 a été publié en 1896, puis le livre a connu une deuxième édition augmentée<sup>1</sup> en 1912, juste avant le décès de son auteur. En fait, Poincaré a également abordé ce domaine des probabilités à plusieurs reprises dans des revues et des livres de philosophie des sciences ou de vulgarisation.

## 24.1. Contenu du *Calcul des Probabilités*

Une analyse complète du livre de Poincaré a été faite par Sheynin<sup>2</sup>. Nous exposons ici quelques définitions, méthodes ou résultats de Poincaré. Nos citations sont extraites de l'édition de 1912.

Rappelons que Poincaré a introduit dans ce traité le terme *fonction caractéristique*. L'usage qu'il en fait correspond à ce que nous appelons aujourd'hui la fonction génératrice des moments<sup>3</sup>.

Dans les trois premiers chapitres de son ouvrage intitulé *Calcul des probabilités* on trouve quelques remarques concernant les concepts ainsi que les théorèmes des probabilités totales et des probabilités composées.

En ce qui concerne le concept de probabilité, Poincaré estime « qu'on ne peut guère en donner une définition satisfaisante » et qu'il est lié à notre degré d'ignorance<sup>4</sup>. Il écrit :

« Une question de probabilité ne se pose que par suite de notre ignorance : il n'y aurait place que pour la certitude si nous connaissions toutes les données du problème. D'autre part, notre ignorance ne doit pas être complète, sans quoi nous ne pourrions rien évaluer. Une classification s'opérerait donc suivant le plus ou moins de profondeur de notre ignorance. [...] »

Notre ignorance est plus grande dans les problèmes de Physique ; il s'agit de prévoir un événement, c'est-à-dire un phénomène conséquent qui dépend d'une part d'un phénomène antécédent, et d'autre part de la loi qui unit l'antécédent au conséquent. Il

1. *Calcul des probabilités*, Gauthier-Villars, 1912.

2. O. B. Sheynin, *Poincaré's work on probability*, AHES, vol. 42, p. 131-171, 1991.

3. Cf. le paragraphe 12.6.1. de notre chapitre 12.

4. *Calcul des probabilités*, Gauthier-Villars, 1912, p. 24 et 30-32.



COURS DE LA FACULTÉ DES SCIENCES DE PARIS

---

COURS DE PHYSIQUE MATHÉMATIQUE

---

CALCUL  
DES  
**PROBABILITÉS**

PAR  
**H. POINCARÉ,**  
MEMBRE DE L'INSTITUT,

RÉDACTION DE  
**A. QUIQUET,**  
ANCIEN ÉLÈVE DE L'ÉCOLE NORMALE SUPÉRIEURE.

---

DEUXIÈME ÉDITION, REVUE ET AUGMENTÉE PAR L'AUTEUR.



PARIS,  
GAUTHIER-VILLARS, IMPRIMEUR-LIBRAIRE  
DU BUREAU DES LONGITUDES, DE L'ÉCOLE POLYTECHNIQUE,  
Quai des Grands-Augustins, 55.

1912

peut se faire que nous connaissions la loi, mais non le phénomène antécédent : quelle est la probabilité pour que se produise le phénomène conséquent ?

Nous connaissons, par exemple, la loi du mouvement des molécules ; si nous connaissons exactement leur position initiale, nous serions capables de dire où elles seront à un moment donné ; la probabilité pour que ces molécules occupent telle position finale dépendra donc de la probabilité que nous attribuerons *par convention* à telle ou telle position initiale. Dans chaque cas, une hypothèse particulière est nécessaire.

Ainsi, quand on cherche la probabilité pour que les comètes aient des orbites elliptiques, on est obligé de faire une convention, on suppose qu'à une grande distance du Soleil ces astres sont uniformément distribués dans l'espace ainsi que les directions de leurs vitesses. »

Quant au problème de la probabilité géométrique, Poincaré l'aborde dans le cours de l'ouvrage en traitant le paradoxe de Bertrand et le *jeu du Franc-Carreau* de Buffon.

Au chapitre II, Poincaré démontre le théorème des probabilités totales et celui des probabilités composées.

Si A et B sont des événements indépendants, écrit Poincaré,

« la somme des probabilités pour que A se produise et pour que B se produise est égale à la somme des probabilités pour que l'un des deux au moins se produise et pour que tous les deux se produisent

$$(A) + (B) = (A \text{ ou } B) + (A \text{ et } B)$$

La probabilité pour que A et B se produisent tous deux est égale à la probabilité pour que B se produise, multipliée par la probabilité pour que A se produise, quand on sait que B s'est produit.

Ou, inversement, elle est égale à la probabilité pour que A se produise, multipliée par la probabilité pour que B se produise, quand on suppose que A doit se produire.

$$(A \text{ et } B) = (B) (A \text{ si } B) = (A) (B \text{ si } A) »$$

Le chapitre 3 traite de façon classique du concept d'espérance mathématique et de la ruine d'un joueur.

Poincaré pose le problème comme suit : Deux joueurs, dont la fortune de l'un est  $m$ , celle de l'autre étant  $n$ , jouent l'un contre l'autre jusqu'à ce que l'un des deux soit ruiné. On se propose de calculer la probabilité que l'événement (ruine d'un des joueurs) se produise. La conclusion de Poincaré est que si la somme des fortunes des joueurs est très grande, l'une d'entre elles étant cependant finie,

« on a la certitude d'être ruiné dans un jeu équitable ou favorable à l'adversaire. Mais si le jeu est favorable au joueur, la probabilité d'être ruiné devient d'autant plus petite que sa fortune est plus grande. »

Au chapitre suivant, Poincaré aborde ce qu'il appelle « les théories qui se rapportent à la formule de Stirling, au théorème de Bernoulli et aux probabilités des causes déduites d'épreuves répétées. »

Poincaré étudie, pour des événements contradictoires de probabilités  $p$  et  $q$ , la propriété des termes du binôme  $(p + q)^m$ . Il écrit :

« Si l'on veut que  $m$  épreuves donnent, dans un ordre quelconque,  $\alpha$  événements A et  $m - \alpha$  événements B, en vertu du principe de la probabilité totale, la probabilité cherchée sera la somme d'autant de termes égaux à  $p^\alpha q^{m-\alpha}$  qu'il y a d'unités dans le nombre de permutations avec répétitions de  $\alpha$  lettres A et de  $m - \alpha$  lettres B ; ce nombre est

$$\frac{m!}{\alpha!(m-\alpha)!},$$

La probabilité pour que les événements se succèdent dans un ordre quelconque est

$$u_\alpha = \frac{m!}{\alpha!(m-\alpha)!} p^\alpha q^{m-\alpha} ;$$

c'est l'un des termes du développement de  $(p + q)^m$ . »

Poincaré cherche alors la valeur la plus probable de  $u_\alpha$  et il note alors que, dans ce cas « si  $m$  est très grand, le rapport  $\alpha/m$  est compris entre  $p + 1/m$  et  $p - 1/m$  ; donc  $\alpha/m$  sera voisin de  $p$ . C'est une forme d'établissement du *théorème de Bernoulli*. Si je choisis  $\alpha$  de façon que  $u_\alpha$  soit le plus grand possible, le rapport du nombre des événements A au nombre des événements B sera à peu près celui des probabilités  $p$  et  $q$ . »

En ce qui concerne la probabilité des causes, traitée au chapitre IX de son livre, Poincaré utilise un exemple intéressant. Il écrit :

« *Formule de Bayes*. Soient  $n$  causes différentes qui peuvent être mises en jeu,  $C_1, C_2, \dots, C_n$  ; la probabilité pour que la cause  $C_i$  si elle est mise en jeu, produise l'événement A est  $p_i$ .

Si nous savions que  $C_i$  est en jeu, nous pourrions affirmer que la probabilité de A est  $p_i$ .

Il faut supposer que deux causes ne peuvent être mises en jeu simultanément.

Avant l'événement, chacune de ces causes avait une probabilité *a priori* que je suppose donnée : la probabilité que la cause  $C_i$  soit mise en jeu était  $\varpi_i$ .

L'événement A a eu lieu : quelle est la probabilité que ce soit la cause  $C_i$  qui l'ait produit ?

Énumérons les cas possibles et les cas favorables, et, pour fixer les idées, considérons un exemple particulier.

M urnes contiennent chacune Q boules ; il y a MQ boules, soit MQ cas possibles, que je suppose également probables.

Les urnes sont réparties en catégories  $C_1, C_2, \dots, C_n$ .

Les urnes de la catégorie  $C_1$  seront au nombre de  $\varpi_1 M$  ; de la catégorie  $C_2$ , au nombre de  $\varpi_2 M$  ; de la catégorie  $C_n$ , au nombre de  $\varpi_n M$ .

La probabilité *a priori* pour que la cause  $C_n$  soit en jeu sera

$$\frac{\varpi_n M}{M} = \varpi_n.$$

Dans les urnes, les boules sont noires ou blanches. L'événement A est, par exemple, la sortie d'une boule blanche. La probabilité pour tirer une boule blanche de la première catégorie sera  $p_1$ .

Dans la catégorie  $C_1$ , il y aura  $p_1 Q$  boules blanches ; dans la catégorie  $C_2$ , il y en aura  $p_2 Q$ , ..., dans la catégorie  $C_n$ , il y en aura  $p_n Q$ .

On a tiré une boule blanche : on demande la probabilité pour que l'urne qu'on a choisie appartienne à la catégorie  $C_i$ .

Le nombre des cas favorables est le nombre des boules blanches de la catégorie  $C_i$ , soit  $\varpi_i p_i MQ$ .

Le nombre total des cas possibles est celui des boules blanches

$$\varpi_1 p_1 MQ + \varpi_2 p_2 MQ + \dots + \varpi_n p_n MQ$$

Le rapport de ces deux nombres est, par définition, la probabilité cherchée

$$\frac{\varpi_i p_i}{\varpi_1 p_1 MQ + \varpi_2 p_2 MQ + \dots + \varpi_n p_n MQ} .$$

Le chapitre X est intitulé *La théorie des erreurs et la moyenne arithmétique*.

Nous avons vu au chapitre 16 du présent ouvrage les hypothèses de départ de la théorie des erreurs de 1809 de Gauss.

Poincaré généralise dans son livre la démonstration de Gauss en partant d'hypothèses moins arbitraires. Voici son texte :

« Je suppose qu'on ait effectué différentes mesures d'une même grandeur : quelle est la probabilité pour que la véritable valeur soit comprise entre  $z$  et  $z + dz$  ?

Il faut introduire une *loi des erreurs*. Je suppose que la véritable valeur de la grandeur à mesurer soit  $z$  ; quelle est la probabilité pour que le résultat de l'observation soit compris entre  $x_1$  et  $x_1 + dx_1$  ? Je pourrai dans tous les cas représenter cette probabilité par

$$dx_1 \varphi(x_1, z) .$$

Cette loi des erreurs étant admise par convention, quelle est la probabilité pour que  $z$  soit compris entre  $z$  et  $z + dz$  ?

C'est un problème de probabilité des causes, et nous allons calculer

$$\frac{\varpi_i p_i}{\sum \varpi_i p_i} .$$

$\varpi_i$  est la probabilité *a priori* pour que  $z$  soit compris entre  $z$  et  $z + dz$  ; cette probabilité sera représentée par

$$\varpi_i = \Psi(z) dz .$$

$\Psi$  étant une fonction qui dépendra de ce que nous savons sur  $z$ .

$p_i$  est la probabilité pour que, à supposer que la quantité observée soit  $z$ , les observations aient donné des résultats compris entre

$$x_1 \text{ et } x_1 + dx_1, x_2 \text{ et } x_2 + dx_2, \dots, x_n \text{ et } x_n + dx_n .$$

La probabilité respective de ces événements est

$$dx_1\varphi(x_1,z), dx_2\varphi(x_2,z), \dots, dx_n\varphi(x_n,z).$$

$p_i$  est la probabilité pour que tous ces événements se soient produits à la fois ; comme ces événements sont indépendants, c'est une probabilité composée

$$p_i = dx_1dx_2\dots dx_n\varphi(x_1,z)\varphi(x_2,z)\dots\varphi(x_n,z).$$

La probabilité *a posteriori* cherchée a pour numérateur

$$dzdx_1dx_2\dots dx_n\Psi(z)\varphi(x_1,z)\varphi(x_2,z)\dots\varphi(x_n,z).$$

Pour obtenir le dénominateur  $\sum \varpi_i p_i$  il faut intégrer cette expression par rapport à  $z$  seulement. Dans le quotient,  $dx_1, dx_2, \dots, dx_n$  sont des constantes qui disparaîtront, et il restera pour la probabilité

$$\frac{dz\Psi(z)\varphi(x_1,z)\varphi(x_2,z)\dots\varphi(x_n,z)}{\int_{-\infty}^{+\infty} dz\Psi(z)\varphi(x_1,z)\varphi(x_2,z)\dots\varphi(x_n,z)}.$$

Cela ne nous apprendrait pas grand-chose si nous n'avions aucune donnée sur  $\varphi$  et  $\Psi$ . On a donc fait une hypothèse sur  $\varphi$ , et cette hypothèse a été appelée *loi des erreurs*.

Elle ne s'obtient pas par des déductions rigoureuses ; plus d'une démonstration qu'on a voulu en donner est grossière [...].

Voici comment Gauss y est arrivé.

Lorsque nous cherchons la meilleure valeur à donner à  $z$ , nous n'avons pas d'autre ressource que de prendre la moyenne entre  $x_1, x_2, \dots, x_n$  en l'absence de toute considération qui justifierait un autre choix. Il faut donc que la loi des erreurs s'adapte à cette façon d'opérer. Gauss cherche quelle doit être  $\varphi$  pour que la valeur la plus probable soit la valeur moyenne.

Si  $dz$  est constant, la probabilité pour que  $z$  soit compris dans l'intervalle  $dz$  est

$$\Psi(z)\varphi(x_1,z)\varphi(x_2,z)\dots\varphi(x_n,z)dz.$$

La valeur la plus probable sera celle pour laquelle cette fonction sera maximum. Supposons ce maximum atteint quand  $z$  est la moyenne.

Gauss a d'abord égalé la fonction  $\Psi$  à 1, puis il a admis que  $\varphi(x_1, z)$  était de la forme  $\varphi(z - x_1)$ .

Quelle doit être alors la fonction  $\varphi$  pour que

$$\varphi(z - x_1)\varphi(z - x_2)\dots\varphi(z - x_n)$$

soit maximum avec cette valeur de  $z$  ?

Égalons à zéro la dérivée logarithmique de l'expression précédente par rapport à  $z$ ,

$$\frac{\varphi'(z - x_1)}{\varphi(z - x_1)} + \frac{\varphi'(z - x_2)}{\varphi(z - x_2)} + \dots + \frac{\varphi'(z - x_n)}{\varphi(z - x_n)} = 0$$

Je pose

$$\frac{\varphi'(z - x_1)}{\varphi(z - x_1)} = F(x_1).$$

L'équation devient

$$F(x_1) + F(x_2) + \dots + F(x_n) = 0$$

Cette condition devra être réalisée toutes les fois que

$$x_1 + x_2 + \dots + x_n = nz$$

Je vais donner à  $x_1, x_2, \dots, x_n$  des accroissements  $dx_1, dx_2, \dots, dx_n$  ;  $z$  restant constant, la somme des  $x$  doit rester constante et l'on doit avoir

$$F'(x_1)dx_1 + F'(x_2)dx_2 + \dots + F'(x_n)dx_n = 0$$

$$dx_1 + dx_2 + \dots + dx_n = 0$$

Ces deux équations doivent être identiques, d'où

$$F'(x_1) = F'(x_2) = \dots = F'(x_n) ,$$

c'est-à-dire que  $F'(x_1)$  est une constante que je représente par  $-a$

$$F(x_1) = a(z - x_1) + b ,$$

et

$$\log \varphi(z - x_1) = \frac{a(z - x_1)^2}{2} + b(z - x_1) + c.$$

Déterminons les trois constantes  $a, b, c$ .

$$F(x_1) + F(x_2) + \dots + F(x_n) = \sum a(z - x_i) + nb = 0$$

$$x_1 + x_2 + \dots + x_n - nz = -\sum (z - x_i) = 0$$

Comme ces deux équations doivent être identiques, on a

$$b = 0,$$

et l'on peut écrire

$$\varphi(z - x_1) = e^c e^{\frac{a(z-x_1)^2}{2}} ;$$

$c$  se détermine par la condition

$$\int_{-\infty}^{+\infty} \varphi(z - x_1) dx_1 = 1 .$$

En posant

$$-a = 2h$$

$$z - x_1 = y ,$$

on trouve

$$\varphi(y) = \sqrt{\frac{h}{\pi}} e^{-hy^2} . »$$

Poincaré obtient donc, comme Gauss, la loi normale.

Il ajoute :

« J. Bertrand présente les objections suivantes :

La fonction  $\varphi$  a été prise sous la forme  $\varphi(z - x_i)$ , tandis qu'en réalité elle devrait être  $\varphi(z, x_i)$ . De plus, on a fait  $\Psi(z) = 1$ , et l'on ne peut l'affirmer *a priori*. Autre objection : La moyenne est-elle la valeur *la plus probable* ou la valeur *probable* ? Ce n'est pas la même chose. [...].

Dans le cas du problème des erreurs, la valeur probable de  $z$  est représentée par le rapport

$$\frac{\int_{-\infty}^{+\infty} z \Psi(z) \varphi(x_1, z) \varphi(x_2, z) \dots \varphi(x_n, z) dz}{\int_{-\infty}^{+\infty} \Psi(z) \varphi(x_1, z) \varphi(x_2, z) \dots \varphi(x_n, z) dz}$$

J. Bertrand dit que Gauss aurait dû chercher, non pas la condition pour que la moyenne soit la valeur la plus probable de  $z$ , mais la condition pour que la moyenne soit la valeur probable de  $z$ .

On peut chercher à s'affranchir des hypothèses que nous avons faites, à savoir que  $\varphi(x_1, z)$  était de la forme  $\varphi(z - x_1)$  que  $\Psi(z)$  était égale à 1 ; on peut se demander quelle forme on pouvait donner à ces deux fonctions pour que la moyenne arithmétique de  $x_1, x_2, \dots, x_n$  fût bien la valeur la plus probable de  $z$ .

En d'autres termes, cette moyenne arithmétique, comme nous l'avons déjà dit doit rendre maximum

$$\Psi(z) \varphi(x_1, z) \varphi(x_2, z) \dots \varphi(x_n, z)$$

Quand il y a maximum, la dérivée logarithmique est nulle c'est-à-dire que si l'on pose

$$\frac{\varphi'(x_1, z)}{\varphi(x_1, z)} = F(x_1, z),$$

$$\frac{\Psi'(z)}{\Psi(z)} = \chi$$

on doit avoir

$$F(x_1, z) + F(x_2, z) + \dots + F(x_n, z) + \chi = 0$$

Cette égalité doit être satisfaite par la valeur de  $z$  que définit l'équation

$$x_1 + x_2 + \dots + x_n = nz$$

Je vais donner à  $x_1, x_2, \dots, x_n$  des accroissements  $dx_1, dx_2, \dots, dx_n$ . Je suppose que  $z$  ne change pas et que la dernière égalité continue à être satisfaite ;  $\chi$  est alors une constante et

$$\frac{dF(x_1)}{dx_1} dx_1 + \frac{dF(x_2)}{dx_2} dx_2 + \dots + \frac{dF(x_n)}{dx_n} dx_n = 0$$

$$dx_1 + dx_2 + \dots + dx_n = 0$$

Ceci ne peut avoir lieu que si

$$\frac{dF(x_1)}{dx_1} = \frac{dF(x_2)}{dx_2} = \dots = \frac{dF(x_n)}{dx_n},$$

Donc

$$\frac{dF}{dx_1} = A',$$

où  $A'$  est fonction de  $z$  seulement ; et

$$F = A'x_1 + B'$$

$B'$  étant aussi fonction de  $z$  seulement.

La condition à remplir devient

$$A'(x_1 + x_2 + \dots + x_n) + nB' + \chi = 0,$$

c'est-à-dire

$$n(A'z + B') + \chi = 0$$

Cette relation ne peut être satisfaite quels que soient  $z$  et  $n$  ;  
donc

$$\chi = 0,$$

c'est-à-dire que  $\Psi(z)$  est constant et que

$$A'z + B' = 0$$

Voilà ce que deviendrait l'analyse de Gauss si l'on voulait la reprendre en tenant compte de la première observation de J. Bertrand.

De

$$F(x_1, z) = A'x_1 + B'$$

on déduit aisément

$$\log \varphi(x_1, z) = FAx_1 + B + \log \theta(x_1)$$

$\log \theta(x_1)$  représente une fonction de  $x_1$  seulement.  $A$  et  $B$  sont des fonctions de  $z$ , admettant des dérivées  $A'$  et  $B'$  telles que

$$A'z + B' = 0$$

Ainsi

$$\varphi(x_1, z) = \theta(x_1)e^{Ax_1 + B}.$$

Tel serait le résultat sans autre condition que le postulat de Gauss sur les valeurs moyennes. Il entre encore deux fonctions arbitraires,  $\theta$  et  $A$  ;  $B$  est lié à  $A$  par une relation. »

A. Hald<sup>1</sup> fait remarquer que la distribution  $\varphi(x_1, z) = \theta(x_1)e^{Ax_1 + B}$  est appelée aujourd'hui la famille de distributions exponentielles à un paramètre. La loi normale de variance 1 est obtenue pour  $A(z) = z$ ,  $B(z) = -\frac{z^2}{2}$  et  $\log \vartheta = -\frac{x^2}{2}$ .

1. *A history of mathematical statistics from 1750 to 1930*, Wiley, 1998, p. 373-376.



Poincaré analyse également au chapitre XIII de son livre la méthode des moindres carrés et il calcule les erreurs à craindre au chapitre suivant. Il applique ensuite la méthode des moindres carrés à une question nouvelle, la recherche d'une fonction inconnue à partir d'un nombre fini de valeurs connues de cette fonction (Théorie de l'interpolation).

Le dernier chapitre du livre *Questions diverses* contient l'introduction par Poincaré de la notion de groupe en probabilités, notion qu'il applique au problème du battage des cartes. Il montre ainsi que si le jeu de cartes a été battu assez longtemps, toutes les permutations possibles sont équiprobables.

Poincaré traite également dans ce chapitre la répartition des décimales dans une table numérique et le problème du mélange des liquides.

Terminons en citant les phrases de Poincaré relatives à l'application des probabilités aux sciences morales<sup>1</sup> :

« Les lois du hasard ne s'appliquent pas à ces questions [...]. Nous sommes tentés d'attribuer au hasard des faits de cette nature, parce que les causes en sont obscures ; mais ce n'est pas là le vrai hasard [...]. Quand des hommes sont rapprochés, ils ne se décident plus au hasard et indépendamment les uns des autres ; ils réagissent les uns sur les autres. Des causes multiples entrent en action, elles troublent les hommes, les entraînent à droite et à gauche, mais il y a une chose qu'elles ne peuvent détruire, ce sont leurs habitudes de moutons de Panurge. Et c'est cela qui se conserve. »

## 24.2. Biographie de Poincaré (1854-1912)

Henri Poincaré naît à Nancy le 29 avril 1854. Son père, Léon Poincaré est médecin neurologue et professeur à la faculté de médecine dans cette ville. Son oncle, Antoine Poincaré, qui a fait ses études à l'École Polytechnique, est inspecteur général des ponts et chaussées. Cet oncle d'Henri aura deux fils : Raymond Poincaré, futur président de la République française, et Lucien Poincaré, physicien qui deviendra recteur de l'Université de Paris puis directeur de l'enseignement secondaire au ministère de l'Instruction publique. En 1856, naît Aline, sœur d'Henri, qui épousera le philosophe Émile Boutroux.

Henri Poincaré a une enfance heureuse. Il révèle très tôt des dons exceptionnels. Poincaré fait ses études secondaires au lycée Impérial de Nancy<sup>2</sup> d'octobre 1862 à août 1873. Il se distingue au lycée par des bonnes notes, surtout en histoire et en géographie. Il obtient ses deux bacs (bac ès lettres et bac ès sciences) avec mention en 1871. Il poursuit ses études dans le même lycée et entre en mathématiques élémentaires en novembre 1871 où il est

1. *Calcul des probabilités*, Gauthier-Villars, 1912, p. 19-20.

2. Rebaptisé Lycée Henri Poincaré en 1913.

premier de sa classe. Il obtient en août 1872 le Prix d'honneur de mathématique au concours général, puis s'inscrit en novembre 1872 en mathématiques spéciales où il compte parmi ses camarades de classe Paul Appell (le futur mathématicien) ainsi que Clément Colson (qui deviendra un économiste réputé). Au concours général de mathématiques spéciales d'août 1873, Poincaré obtient la plus haute note de toute la France ainsi qu'une mention spéciale des examinateurs. Il se présente aux concours d'entrée de l'École normale (où il ne sera reçu que cinquième, à cause de l'épreuve de géométrie descriptive) et de l'École Polytechnique où il est reçu premier. Admis à la fois à l'École normale et à l'École polytechnique, Poincaré opte, sur les conseils de son oncle Antoine, pour cette dernière. Alors qu'il est encore à Polytechnique et qu'il n'a que vingt ans, Poincaré présente en 1874 sa première publication<sup>1</sup>. L'année suivante il sort second de l'École (à cause d'une mauvaise note en stéréotomie) et entre en octobre 1875 à l'École des Mines. Il continue à s'occuper de mathématiques, négligeant complètement les cours d'application. Lecornu, avec lequel il fait en 1877 un voyage d'études en Autriche-Hongrie indique : « Poincaré m'émerveillait, non seulement pas ses dons d'intuition, mais aussi par son imperturbable mémoire ; il connaissait toutes les dates importantes de l'histoire, tous les horaires de chemin de fer, etc. »

En 1878, Poincaré sort troisième de l'École des Mines et devient, en août 1879, docteur ès sciences mathématiques à la suite de la présentation d'une thèse intitulée *Sur l'intégration des équations aux dérivées partielles à un nombre quelconque d'inconnues*. Peu après, le 1<sup>er</sup> décembre 1879, il est nommé chargé du cours d'analyse mathématique à la faculté des sciences de Caen. À cette époque, il découvre les fonctions qu'il a appelées fuchsienues. En 1880, alors qu'il n'a que 27 ans, Poincaré présente pour la première fois sa candidature (sans succès) à l'Académie des sciences, il a alors déjà à son actif plusieurs dizaines de mémoires ou notes publiés dans les Comptes rendus de l'Académie, avec des résultats dans des domaines aussi divers que la théorie des équations différentielles, les fonctions fuchsienues, l'arithmétique, l'algèbre ou la théorie des fonctions elliptiques et abéliennes. Vers la fin de son séjour à Caen, le 20 avril 1881, il épouse Mlle Poulain d'Andecy, de la famille des Geoffroy Saint-Hilaire, et en octobre il devient Maître de conférences d'Analyse à la Faculté des Sciences de Paris<sup>2</sup>. Sa carrière académique évolue rapidement ; il est nommé répétiteur d'Analyse à l'École polytechnique en 1883 puis professeur de mécanique physique et expérimentale de la Faculté des sciences de Paris en mars 1885. De là datent ses premiers travaux en physique. Cette même année, Poincaré étudie de façon systématique les

1. *Démonstrations nouvelles des propriétés de l'indicatrice d'une surface*, Nouvelles annales mathématiques, 2<sup>e</sup> série, 13 (octobre 1874), pages 449-456.
2. Quand Poincaré est nommé, les fonctions de maître de conférences viennent d'être créées. L'idée est d'adjoindre au professeur titulaire un répétiteur dont la fonction est d'interroger les élèves et de s'assurer qu'ils ont compris le cours en leur donnant des devoirs que le maître de conférence est chargé de corriger.

conditions d'équilibre d'un fluide en rotation dans un champ de force<sup>1</sup>. Ceci l'amène d'une part à introduire la notion de bifurcation et d'autre part à montrer que les anneaux de Saturne ne peuvent qu'être composés d'une multitude de très petits satellites.

En 1886, Poincaré succède à Gabriel Lippmann<sup>2</sup> à la chaire de physique mathématique de la Sorbonne. Gaston Darboux, lors du discours qu'il prononcera le 21 janvier 1912 à l'occasion du jubilé de Poincaré, rappellera justement : « Avec des hommes tels que vous, la Faculté allait toute seule... Lorsque la considération des services m'a déterminé à vous demander de changer d'enseignement, vous l'avez fait sans hésitation, une première fois pour prendre la chaire de physique mathématique, une seconde fois pour passer à celle de mécanique céleste. Et ainsi, j'ai aujourd'hui la joie et l'orgueil de penser que j'ai pu avancer le moment où, en même temps que grand géomètre, vous avez été proclamé partout grand physicien et grand astronome ». En 1887, Poincaré est élu à l'Académie des sciences alors qu'il n'a que trente-deux ans. Après la mécanique des fluides, il travaille tour à tour sur l'optique (1887), la thermodynamique (1888) et l'électromagnétisme de Maxwell (1889). La famille Poincaré s'agrandit avec la naissance de trois filles (Jeanne, Yvonne et Henriette) et un fils (Léon) qui entrera à polytechnique en 1912. Henri Poincaré est lauréat en 1889 du Grand Prix du Roi de Suède pour sa contribution au problème des trois corps. Son mémoire *Sur le problème à trois corps et les équations de la dynamique*, contient les prémices de la théorie du chaos. Il développe et publie ses travaux sur *Les méthodes nouvelles de la mécanique céleste* dans trois volumes (1892, 1893, 1899). En tant que professeur de physique mathématique il poursuit son vaste tour d'horizon de la physique contemporaine, écrivant successivement sur les théories de Helmholtz et les expériences de Hertz (1889), la théorie de l'élasticité (1890), la théorie mathématique de la lumière et celle des tourbillons (1891), les oscillations électriques (1892), la théorie analytique de la propagation de la chaleur (1893), le calcul des probabilités (1893), le potentiel newtonien (1894), etc. C'est lui qui introduit la théorie électromagnétique de Maxwell en France. Plus tard il s'occupera des rayons cathodiques et des rayons X et mettra Henri Becquerel sur la voie de la découverte de la radioactivité. Ses travaux sur le problème des trois corps le font nommer en 1896 à la chaire de mécanique céleste de la Faculté des sciences de Paris, puis président du Bureau des longitudes en

1. Poincaré fait une étude très originale des figures d'équilibre d'une masse fluide en rotation. On admettait que ces figures étaient au nombre de deux : l'ellipsoïde aplati, et un ellipsoïde à trois axes inégaux, dit de Jacobi. Poincaré montra qu'avec l'accroissement de la vitesse de rotation la masse fluide devait prendre une troisième forme, comparable à celle de la poire, et finir par se partager en deux corps isolés (un phénomène analogue permet d'interpréter la fission d'un noyau dans le cadre du modèle de la goutte liquide).
2. Gabriel Lippmann (1845-1921) est l'inventeur en 1891 de la photographie en couleurs par une méthode interférentielle (utilisée de nos jours pour la réalisation d'hologramme couleur), ce qui lui vaut le Prix Nobel de physique en 1908.

1899. Il publie en 1902 son fameux ouvrage *La science et l'hypothèse* qui sera traduit dans plus de vingt-trois langues, puis un second ouvrage philosophique *La valeur de la science* en 1905. Poincaré devient en 1906 président de l'Académie des sciences, mais renonce en 1907 à se porter candidat au poste de secrétaire perpétuel. En 1908 il entre à l'Académie française. Le cousin germain d'Henri Poincaré, Raymond Poincaré (1860-1934), est élu à l'Académie française l'année suivante (1909). De 1909 à 1912 il y a donc, fait sans précédent, deux membres de la même famille à l'Académie française<sup>1</sup>. Henri Poincaré devient du 1<sup>er</sup> janvier au 1<sup>er</sup> avril 1912 directeur de l'Académie française et devait en devenir le secrétaire perpétuel, mais il meurt brusquement le 17 juillet 1912, à la suite d'une opération de la prostate. Il n'a que 58 ans. Il laissera à la postérité plus de trente volumes et près de cinq cents mémoires et a appartenu à divers titres à une quarantaine d'Académies ou de Sociétés savantes, françaises ou étrangères. Jean Dieudonné résume ainsi son œuvre mathématique : « Poincaré, génie égal à Gauss, et aussi universel. Il a dominé toutes les mathématiques de son temps », tandis que pour E.T. Bell, Poincaré est le « dernier universaliste ». Comme Gauss, Poincaré est également un très grand physicien et peut-être serait-il plus connu comme physicien s'il avait reçu le prix Nobel de physique, bien mérité, pour lequel il fut « nominé » en 1910 et 1911.

## Les œuvres mathématiques de Poincaré

Les travaux de Poincaré en mathématiques sont si nombreux que nous tenterons seulement d'en énumérer quelques-uns parmi les plus importants. Poincaré effectue des études sur les équations différentielles où il introduit la notion de fonction automorphe<sup>2</sup>. Il appelle ces fonctions, *fonctions fuchsienues*. Une dizaine de communications à l'Académie des sciences, en 1881, portent ce titre. En 1882, il publie deux mémoires importants sur le sujet dans le premier numéro des *Acta Mathematica*.<sup>3</sup> Poincaré raconte lui-même dans *Science et Méthode* (1908, p. 51) comment les liens de ces travaux avec la géométrie non-euclidienne lui sont venus à l'idée : « A ce moment, je quittais Caen, où j'habitais alors, pour prendre part à une course géologique entreprise par l'École des Mines. Les péripéties du voyage me firent oublier mes travaux mathématiques ; arrivés à Coutances, nous montâmes dans un

1. Raymond Poincaré deviendra président de la République française de 1913 à 1920. Raymond est de six ans le cadet d'Henri. Lorsque Raymond Poincaré est arrivé à Paris pour faire des études de droit, il logeait avec son cousin Henri dans un modeste hôtel du quartier latin, au 4 carrefour de l'Odéon.

2. Une fonction automorphe d'une variable complexe  $z$  est analytique, excepté pour ses pôles, dans un domaine  $D$  et est invariante sous un groupe de transformations du type  $z' = \frac{az + b}{cz + d}$ .

3. *Théorie des groupes fuchsienues*, *Acta Mathematica*, 1 (1882), pages 1-62. Et *Sur les fonctions fuchsienues*, *Acta Mathematica*, 1 (1882), pages 193-294.

omnibus pour je ne sais quelle promenade ; au moment où je mettais le pied sur le marchepied, l'idée me vint, sans que rien dans mes pensées antérieures parût m'y avoir préparé, que les transformations dont j'avais fait usage pour définir les fonctions fuchsiennes étaient identiques à celles de la géométrie non-euclidienne. »

Poincaré s'intéresse aux fonctions abéliennes et à la géométrie algébrique dès 1881 et jusqu'en 1911. Son mémoire le plus important sur le sujet traite des courbes contenues dans une surface algébrique<sup>1</sup>. Dans le domaine de la théorie des nombres il publie un article remarquable *Sur les propriétés arithmétiques des courbes algébriques*, en 1901<sup>2</sup>.

En Algèbre, Poincaré, grand admirateur de Sophus Lie, cherche une nouvelle démonstration du troisième théorème fondamental de Lie et publie un article résumant ses propres travaux en 1901<sup>3</sup>.

Les équations différentielles sont le domaine de prédilection des recherches de Poincaré. Il s'est en particulier proposé d'intégrer toutes les équations différentielles linéaires à coefficients algébriques. Il publie de 1881 à 1886 quatre longs mémoires sur le sujet<sup>4</sup>. Le cas de l'équation du premier ordre sur la sphère occupe les trois premiers mémoires. Les systèmes du second ordre font l'objet du quatrième et correspondent à l'étude générale des équations de la dynamique dont la mécanique céleste est un cas particulier. Le problème des trois corps en est le prolongement. Le dernier traitement original du problème des trois corps avant Poincaré était dû à Lagrange qui l'avait publié sous le titre *Essai d'une nouvelle méthode pour résoudre le problème des trois corps*<sup>5</sup>. L'un des résultats du traitement par Lagrange du problème des trois corps est la découverte de cinq points particuliers qui gardent une position relative constante des trois corps. Ces points sont appelés points de Lagrange.

En 1889, Poincaré présente un mémoire pour l'obtention du prix créé en 1887 par Oscar II, roi de Suède et de Norvège, en prévision de son soixantième anniversaire. Parmi les membres du jury figurent Hermite et Weierstrass et parmi les concurrents Henri Poincaré et son ami Paul Appel. Poincaré obtient

1. *Sur les courbes tracées sur les surfaces algébriques*, Annales scientifiques de l'École Normale Supérieure, 3<sup>e</sup> série, 27 (1910), pages 55-108.
2. *Journal de mathématiques pures et appliquées*, 5<sup>e</sup> série, 7(1901), pages 161-233.
3. *Quelques remarques sur les groupes continus*, Rendiconti del Circolo matematico di Palermo, 15 (1901), pages 321-368.
4. « Mémoire sur les courbes définies par une équation différentielle », I<sup>re</sup> partie, *Journal de mathématiques pures et appliquées*, 3<sup>e</sup> série, 7 (1881), pages 375-422 ; II<sup>e</sup> partie, *Journal de mathématiques pures et appliquées*, 3<sup>e</sup> série, 8, (1882), pages 251-296 ; III<sup>e</sup> partie, *Journal de mathématiques pures et appliquées*, 4<sup>e</sup> série, (1885), pages 167-244 ; IV<sup>e</sup> partie, *Journal de mathématiques pures et appliquées*, 4<sup>e</sup> série, (1886), pages 151-217.
5. Cet essai date de 1772, et a reçu le Prix de l'Académie royale des sciences.

ce prix relatif au problème de la stabilité du système solaire. Cependant, au moment de la publication du papier de Poincaré, une erreur est découverte. Elle a donné lieu à une importante correspondance avec Mittag-Leffler, mathématicien suédois, éditeur des *Acta Mathematica* et conseiller scientifique du souverain. Les recherches menées par Poincaré pour résoudre cette erreur donneront naissance à la théorie du chaos. Une version révisée du papier paraît en 1890 sous le titre *Sur le problème des trois corps et les équations de la dynamique*<sup>1</sup>.

Dans ce mémoire, Poincaré donne la première description mathématique du mouvement chaotique et est le premier à utiliser largement la notion d'invariant intégral. Toujours dans ce mémoire Poincaré écrit : « Je n'ai pu résoudre rigoureusement et complètement le problème de la stabilité du système solaire, en entendant ce mot dans un sens strictement mathématique. L'emploi des invariants intégraux m'a cependant permis d'atteindre certains résultats partiels, s'appliquant surtout au problème, dit restreint, où les deux corps principaux circulent dans des orbites sans excentricité, pendant que le corps troublé a une masse négligeable. Dans ce cas, si on laisse de côté certaines trajectoires exceptionnelles, dont la réalisation est infiniment peu probable, on peut démontrer que le système repassera une infinité de fois aussi près que l'on voudra de sa position initiale. C'est ce que j'ai appelé la stabilité à la Poisson. »

En 1908, dans *Science et Méthode*, Poincaré ajoute ce commentaire : « Une cause très petite, et qui nous échappe, détermine un effet considérable que nous ne pouvons pas ne pas voir, et alors nous disons que cet effet est dû au hasard. Si nous connaissions exactement les lois de la Nature et la situation de l'Univers à l'instant initial, nous pourrions prédire la situation de ce même Univers à l'instant ultérieur. Mais, lors même que les lois naturelles n'auraient plus de secret pour nous, nous ne pourrions connaître la situation initiale qu'approximativement. Si cela nous permet de prévoir la situation ultérieure avec la même approximation, c'est tout ce qu'il nous faut, nous disons que le phénomène a été prévu, qu'il est régi par des lois ; mais il n'en est pas toujours ainsi, il peut arriver que de petites différences dans les conditions initiales engendrent de très grandes dans les phénomènes finaux ; une petite erreur sur les premières produirait une erreur énorme sur les derniers. La prédiction devient impossible et nous avons le phénomène fortuit. »

L'étude des conditions moyennant lesquelles deux figures peuvent ou ne peuvent pas être ramenées l'une à l'autre par déformation continue sans soudure ni déchirure, s'appelle *Analysis situs* (ou topologie). Poincaré publie six grands mémoires sur la question de 1895 à 1904<sup>2</sup>.

1. *Acta Math.* 13 (1890), 1-270

2. *Analysis Situs*, Journal de l'École Polytechnique, 2<sup>e</sup> série, cahier 1 (1895), pages 1-121 ; *Complément à l'Analysis Situs*, Rendiconti del Circolo matematico di Palermo, 13 (1899),

En physique mathématique, Poincaré, qui occupe la chaire de Physique mathématique et de calcul des probabilités pendant dix ans à la faculté des sciences de Paris à partir de 1896, a laissé une douzaine de volumes de leçons, toutes transcrites par ses étudiants (parmi lesquels Émile Borel). Son travail le plus novateur est sans doute l'établissement du formalisme mathématique complet de la relativité restreinte à partir des travaux de son ami H.A. Lorentz.

En mécanique céleste, en plus de très nombreuses publications en revues, son très important ouvrage en trois volumes sur les *Méthodes nouvelles de la mécanique céleste*<sup>1</sup> a confirmé la réputation de Poincaré de *dernier universaliste*. Enfin, il semblerait que Perelman a pu démontrer récemment *la conjecture de Poincaré*, selon laquelle la sphère est la seule hypersurface à trois dimensions dépourvue de trous.

### 24.3. Références

- J. Dieudonné, *Dictionary of Scientific Biography*, Charles Scribner's sons, New York, vol. 2, p. 51-60, 1970.
- P. Appell, *Henri Poincaré*, Paris, 1925.
- A. Bellivier, *Henri Poincaré, ou la vocation souveraine*, Paris, 1956.
- E.T. Bell, *Men of mathematics*, Simon et Schuster, New York, 1986.
- G. Darboux, *Éloge historique d'Henri Poincaré*, Mémoires de l'Académie des sciences, 52 (1914), 81-148.
- J.-J. Samueli et J.-C. Boudenot, *Poincaré, physicien*, Ellipses, 2005 (cet ouvrage comporte une liste des publications de Poincaré).

---

pages 285-343 ; *Second complément à l'Analysis Situs*, Proceedings of the London Mathematical Society, 32 (1900), pages 277-308 ; *Sur certaines surfaces algébriques ; troisième complément à l'Analysis Situs*, Bulletin de la Société mathématique de France, 30 (1902), pages 49-70 ; *Sur les cycles des surfaces algébriques ; quatrième complément à l'Analysis Situs*, Journal de mathématiques pures et appliquées, 5<sup>e</sup> série, 8 (1902), pages 169-214 ; *Cinquième complément à l'Analysis Situs*, Rendiconti del Circolo matematico di Palermo, 18 (1904), pages 45-110.

1. *Les méthodes nouvelles de la Mécanique céleste*, Paris, Gauthier-Villars, tome 1, 1892, tome 2, 1893, tome 3, 1899.

# Épilogue : Kolmogorov et l'axiomatisation de la théorie des probabilités

L'émergence de la physique statistique est sans doute à l'origine de la réévaluation des fondements logiques des probabilités. Les définitions et concepts proposés par Laplace ayant montré leur insuffisance dans l'application à diverses disciplines, la méthode axiomatique a été élaborée car elle permet de prendre en compte la totalité des objets étudiés à l'aide d'une théorie mathématique. L'exemple d'axiomes de la géométrie différents de ceux d'Euclide, proposés par Lobachevsky avait montré que des géométries non-euclidiennes pouvaient être envisagées. Diverses tentatives furent donc tentées dans le domaine des probabilités.

Par ailleurs, lors du deuxième congrès international de mathématiques tenu à Paris en 1900, David Hilbert avait présenté une liste de vingt-trois problèmes qui tenaient jusqu'alors les mathématiciens en échec. Parmi ces problèmes, le sixième était relatif à l'axiomatisation, basée sur le modèle mathématique, de la physique. Compte tenu de l'apparition de la théorie de la relativité et de la mécanique quantique, le problème fut vite considéré comme obsolète. En axiomatisant la Théorie des probabilités, Kolmogorov a, cependant, résolu en partie ce problème.<sup>1</sup>

Dans ses *Fondations de la théorie des probabilités*<sup>2</sup>, Kolmogorov écrit<sup>3</sup>, en 1933 :

« La théorie des probabilités, en tant que discipline mathématique, peut et doit être développée à partir d'axiomes exactement par les mêmes méthodes que la géométrie et l'algèbre. Cela signifie qu'après avoir défini les éléments qui doivent être étudiés et leurs relations de base et établi les axiomes par lesquels ces relations

1. Cf. F. Terpe, *Andrej Nikolajevic Kolmogorov und die Hilbertschen Probleme*, Ernst-Moritz-Arndt-Univ. Math.-Natur. Reihe, 38 (4) (1989), p. 54-58.
2. A. Kolmogoroff, *Grundbegriffe der Wahrscheinlichkeit-rechnung*, Springer, Berlin, 1933.
3. Page 1.



# ERGEBNISSE DER MATHEMATIK UND IHRER GRENZGEBIETE

HERAUSGEGEBEN VON DER SCHRIFTLLEITUNG

DES

„ZENTRALBLATT FÜR MATHEMATIK“

ZWEITER BAND

---

3

---

## GRUNDBEGRIFFE DER WAHRSCHEINLICHKEITS- RECHNUNG

VON

A. KOLMOGOROFF



BERLIN  
VERLAG VON JULIUS SPRINGER  
1933

doivent être gouvernées, toute exposition ultérieure doit être exclusivement basée sur ces axiomes et rester indépendante de la signification usuelle et concrète de ces éléments et de leurs relations. [...] Toute théorie axiomatique abstraite admet, comme cela est bien connu, un nombre illimité d'interprétations concrètes en plus de celles à partir desquelles elle a été dérivée. »

Kolmogorov définit comme suit cinq axiomes<sup>1</sup> :

« Soit  $E$  un ensemble d'éléments  $\xi, \eta, \varsigma, \dots$ , que nous appellerons événements élémentaires, et  $\mathfrak{F}$  une famille de sous-ensembles de  $E$ , les éléments de  $\mathfrak{F}$  seront appelés événements aléatoires.

I.  $\mathfrak{F}$  est une famille d'ensembles

II.  $\mathfrak{F}$  contient  $E$

III. À chaque ensemble  $A$  de  $\mathfrak{F}$  est assigné un nombre non-négatif  $P(A)$ . Ce nombre  $P(A)$  est appelé probabilité de l'événement  $A$ .

IV.  $P(E) = 1$ .

V. Si  $A$  et  $B$  n'ont pas d'éléments en commun, alors

$$P(A + B) = P(A) + P(B).$$

Un système d'ensembles,  $\mathfrak{F}$ , associé à un assignement défini de nombres  $P(A)$ , satisfaisant les axiomes I à V est appelé un espace de probabilités.

Notre système d'axiomes I à V est consistant. Ceci est prouvé par l'exemple suivant.

Soit  $E$  consistant du seul élément  $\xi$  et considérons  $\mathfrak{F}$  constitué de  $E$  et de l'ensemble nul  $\emptyset$ .  $P(E)$  est alors égal à 1 et  $P(\emptyset)$  égal à 0.

Notre système d'axiomes, cependant n'est pas complet, car dans divers problèmes de la théorie des probabilités, différents espaces de probabilités sont à examiner.

*Construction d'espaces de probabilités.* Les plus simples sont construits comme suit : Nous prenons un ensemble fini arbitraire  $E = \{\xi_1, \xi_2, \dots, \xi_k\}$  et un ensemble arbitraire  $\{p_1, p_2, \dots, p_k\}$  de nombres non négatifs dont la somme  $p_1 + p_2 + \dots + p_k = 1$ ,  $\mathfrak{F}$  est pris comme l'ensemble de tous les sous-ensembles dans  $E$  et nous posons

$$P\{\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_\lambda}\} = p_{i_1} + p_{i_2} + \dots + p_{i_\lambda}.$$

Dans de tels cas,  $p_1, p_2, \dots, p_k$  sont appelés les probabilités des événements élémentaires  $\xi_1, \xi_2, \dots, \xi_k$ , ou simplement probabilités élémentaires. De cette façon sont dérivés tous les espaces finis de probabilité dans lesquels  $\mathfrak{F}$  est l'ensemble de tous les sous-ensembles de  $E$ . L'espace de probabilité est appelé fini si l'ensemble  $E$  est fini. »

Il est à noter que Kolmogorov fait référence à Sergei Natanovich Bernstein, dans sa toute première note en bas de page de son mémoire, comme un auteur ayant proposé un autre *Begründungssystem*, c'est-à-dire un autre ensemble de postulats. Il cite deux travaux de Bernstein, l'un est un mémoire de 1917 dont le titre pourrait être traduit par *Recherches d'un fondement axiomatique de la théorie des probabilités*<sup>2</sup> et l'autre est le livre *Théorie des Probabilités*, publié

1. Page 2.

2. Mitt. math. Ges. Charkow, vol. 15, 1917, p. 209-274.

en Russe à Moscou en 1927<sup>1</sup>. Est également cité, l'ouvrage de Richard von Mises<sup>2</sup> publié en 1931.

Bernstein avait introduit trois axiomes : l'axiome de la comparabilité des probabilités, celui des événements disjoints ou incompatibles, et celui de la combinaison des événements.

---

1. Éditions d'État, RSFSR, 1927.

2. *Wahrscheinlichkeit rechnung*, Leipzig et Vienne, Deuticke, 1931.

# Chronologie

## des événements constituant le contexte historique, politique et scientifique de l'élaboration du calcul des probabilités

Afin de permettre de mieux suivre l'évolution de la théorie des probabilités, nous donnons ci-dessous une chronologie des événements présentant le contexte politique et scientifique de l'élaboration du calcul de cette science. Pour en faciliter la lecture la présentation est faite par périodes de cinquante ans, à partir de 1650.

### 1650-1700

#### *Histoire politique*

Au plan politique, l'année 1654 est marquée par le sacre, le 7 juin, de Louis XIV à Reims. En 1657, une alliance franco-anglaise contre l'Espagne est conclue. La guerre qui s'en suit s'achèvera deux ans plus tard par le traité des Pyrénées. Au nord du pays, la France gagne des villes de Flandre et du Luxembourg. Louis XIV épouse Marie-Thérèse fille du roi Philippe IV d'Espagne et, à la mort de Mazarin, en 1661, le roi Soleil entame à 23 ans, un règne qui durera jusqu'en 1715.

En Angleterre, l'année 1658 est celle de la mort de Cromwell, son fils lui succède, mais il doit abdiquer dès l'année suivante. Cette année 1659 voit le rétablissement de la royauté par Charles II. En 1664, la guerre entre la Hollande et l'Angleterre aboutit à la prise par les Anglais de *Nieuw Amsterdam*, rebaptisée *New York*.

La guerre de Hollande débute en 1672 et oppose les Français et les Anglais aux Hollandais alliés aux Impériaux allemands. L'Empereur germanique envahit l'Alsace qui sera finalement libérée par l'armée française au début de l'année 1675. Par ailleurs la paix de Nimègue met fin à la guerre de Hollande en 1678. L'année suivante, l'*Habeas corpus* assure aux Anglais leur liberté individuelle, cependant qu'en France les Protestants sont persécutés. Tandis

que Louis XIV annexe Strasbourg en 1681, en Russie le règne de Pierre le Grand débute en 1682. Il durera jusqu'en 1725.

Les Turcs assiègent la ville de Vienne en 1683, mais ils sont vaincus par le roi de Pologne et ses alliés.

L'édit de Nantes par lequel Henri IV avait autorisé la liberté de culte aux protestants est révoqué en 1685. Cette révocation entraînera l'exil de beaucoup de huguenots, dont celui du français Abraham de Moivre à Londres.

De leur côté, les Turcs sont chassés de Hongrie en 1686.

Deux ans plus tard la guerre de neuf ans contre la ligue d'Augsbourg débute. Les Français remportent la victoire de Fleurus en 1690.

Notons que la naissance du libéralisme peut être datée de 1690, date à laquelle John Locke fait paraître son ouvrage *Two treatises of government*<sup>1</sup>. On assiste pendant cette même période à la disparition de plusieurs grands noms du théâtre et de la philosophie : Molière meurt en 1673 à l'âge de 51 ans ; le grand philosophe et universaliste hollandais Spinoza disparaît en 1677, à 45 ans ; Pierre Corneille s'éteint en 1682 à l'âge de 78 ans ; tandis que Jean Racine disparaît en 1699, à l'âge de 60 ans.

## Évolution scientifique

En science, on note la publication, en 1655, de l'*Arithmetica infinitorum* de John Wallis et en 1659, Pascal fait paraître ses *Lettres de A. Dettonville sur la géométrie*. En 1662 paraît l'ouvrage de Robert Boyle intitulé *New experiments physico-mechanical, touching the spring of air* où est formulée la fameuse loi qui porte son nom sur la relation entre la pression et le volume d'un gaz. L'année suivante l'ouvrage de Pascal, *Traitez de l'Équilibre des Liqueurs, et de la Pesanteur de la Masse de l'Air*, est publié de façon posthume (Pascal meurt en 1662, à 39 ans) ; ce traité marque le début de l'hydrostatique. Trois ans après la disparition de Pascal, Pierre de Fermat meurt à son tour, âgé de 64 ans. Son fils Samuel publiera, en 1670, *Arithmeticonum libri sex, et de numeris multangulis liber unus* de Diophante, avec les commentaires de son père. Il édite ensuite, en 1679, les *Varia opera mathematica*. Ce livre contient la correspondance échangée entre Pascal et Pierre Fermat sur les probabilités.

En 1667, Robert Hooke révèle l'usage du microscope dans son ouvrage *Micrographia : or, some physiological descriptions of minute bodies made by magnifying glasses*. Entre 1669 et 1672, Christian Huygens et son frère Louis échangent une correspondance relative aux rentes viagères et à l'espérance de vie. Otto de Guericke décrit en 1672 ses célèbres expériences de Magdebourg sur le vide dans un ouvrage intitulé *Experimenta nova (ut vocantur)*

1. London, Churchill, 1690.

*magdeburgica de vacuo spatio*. De janvier à mars 1673, Leibniz séjourne à Londres et rencontre Henry Oldenburg qui lui parle des découvertes de Newton concernant le calcul infinitésimal. La lettre d'Isaac Newton sur la décomposition spectrale de la lumière blanche « containing his new theory about light and colors » paraît dans les *Philosophical Transactions of the Royal Society* de Londres en 1672. Quant à Christian Huygens, il publie, en 1673, son livre *Horologium oscillatorium sive de motu pendulorum ad horologia aptato demonstrationes geometricae* qui servira de base à Isaac Newton pour les équations de la mécanique classique. Par ailleurs, les *Lectiones opticae et geometricae* d'Isaac Barrow sont éditées par Newton, en 1674. Le grand concurrent de Newton, Gottfried Wilhelm Leibniz, révèle en 1684 (sans démonstration) dans son article *Nova methodus pro maximis et minimis*, publié dans la revue scientifique *Acta Eruditorum*<sup>1</sup> de Leipzig, sa technique du calcul différentiel. Trois ans plus tard, en 1687, paraît le livre le plus célèbre de l'histoire de la Physique, voire même de l'histoire des sciences, les *Philosophiae naturalis principia mathematica*, de Newton.

De son côté, Huygens publie en 1690 le *Traité de la lumière où sont expliquées les causes de ce qui lui arrive dans la reflexion, & dans la refraction. Et particulièrement dans l'étrange refraction du cristal d'Islande. Par Monsieur Christian Huygens... Avec un discours de la cause de la pesanteur* : c'est l'acte de naissance de la théorie ondulatoire de la lumière.

Cette période se termine avec la parution, entre 1693 et 1699, des œuvres complètes de John Wallis, *Opera mathematica* ainsi que par celle de l'ouvrage de Guillaume de L'Hospital sur l'*Analyse des infiniments petits pour l'intelligence des lignes courbes* (1696) qui est le premier traité de calcul infinitésimal, rédigé avec l'aide des Bernoulli.

En sciences de la nature signalons que le sexe des plantes et la pollinisation sont découverts par Rudolph Camerarius en 1694<sup>2</sup>, tandis que l'inventeur du microscope, Anton van Leeuwenhoek, observe des microbes en 1696.

La communauté scientifique s'organise, et les académies succèdent à des cercles plus ou moins formels. La première académie scientifique l'*Accademia dei Lincei*, à Rome, est active entre 1603 et 1651<sup>3</sup>. À Florence, l'*Accademia del Cimento* (Académie de l'expérience), active entre 1657 et 1667, réunit essentiellement des expérimentateurs. La *Royal Society* de Londres est fondée en 1660<sup>4</sup>, tandis que l'*Académie des sciences* de Paris voit le jour en 1666. Elle doit son origine au projet de Colbert de créer une académie générale et s'inscrit

1. Le premier numéro de cette revue est sorti en 1682.

2. *De sexu plantarum Epistola*, Tübingen, 1694.

3. L'*Accademia dei Lincei* est maintenant l'Académie Pontificale des sciences.

4. Henry Oldenburg (1615-1677) est le premier secrétaire de la *Royal Society* tandis que John Collins (1625-1683), comme jadis Marin Mersenne (1588-1648) en Europe, correspond avec de nombreux savants.

également dans la lignée des divers cercles de savants qui se réunissaient au XVII<sup>e</sup> siècle, autour d'un mécène ou d'une personnalité érudite. Colbert a choisi un petit groupe de savants qui se sont rassemblés le 22 décembre 1666 dans la bibliothèque du roi, nouvellement installée rue Vivienne. Les séances de travail deviennent bi-hebdomadaires<sup>1</sup>.

L'*Académie de Berlin* est, quant à elle, créée en 1700 tandis que celle de *Saint-Pétersbourg* est fondée en 1720.

Avec les académies apparaissent les premiers journaux scientifiques. Le numéro 1 du *Journal des savans*, qui est la première publication scientifique périodique française, paraît le 5 janvier 1665. En Angleterre, le volume I des *Philosophical Transactions, giving some accompt of the present understanding, studies and labours of the ingenious in many considerable parts of the world*, paraît en 1667. Notons aussi que, pendant cette période, des jardins botaniques sont créés : celui de Paris a été ouvert en 1597, celui d'Oxford naît en 1621, tandis que celui de Berlin voit le jour en 1679.

### **Faits marquants en calcul des probabilités**

On peut situer le début de la théorie mathématique des probabilités vers 1565 avec Cardan, (mais son livre, *Liber de Ludo Aleae*, écrit vers cette date, ne sera publié qu'en 1663) ou citer le mémoire de Galilée sur les jeux de dés datant d'environ 1620, mais la plupart des commentateurs voient plutôt ce début des probabilités lié aux questions posées par le Chevalier de Méré à Pascal en 1654 : « Un problème relatif aux jeux de hasard, proposé à un austère janséniste par un homme du monde, a été à l'origine du calcul des probabilités » sont les premières lignes du livre qu'a consacré Poisson à cette discipline en 1837.

Entre juillet et octobre 1654, sept lettres sur le problème des partis ou des points dans les jeux de hasard sont échangées entre Pascal et Fermat. Cet échange peut être considéré comme la naissance du calcul des probabilités. Les lettres seront publiées en 1679 par le fils de Pierre Fermat, Samuel. En 1656 Huygens écrit son mémoire sur les probabilités *Van Rekeningh in Spelen van Geluck* et l'envoie à Frans van Schooten qui décide d'en inclure une traduction latine, *De ratiociniis in ludo aleae*, dans un ouvrage qui paraît un an plus tard sous le titre *Exercitationum mathematicorum libri quinque*.

---

1. Les trente premières années d'existence de l'Académie sont relativement informelles, la nouvelle institution n'ayant pas reçu de statuts. Le 20 janvier 1699 Louis XIV donne à la compagnie son premier règlement. L'Académie reçoit le titre d'Académie royale et est installée au Louvre. Composée de 70 membres, elle contribue au XVIII<sup>e</sup> siècle au mouvement scientifique de son temps par ses publications et joue un rôle de conseil auprès du pouvoir.

L'année 1662 voit la naissance de l'arithmétique politique avec le traité de John Graunt *Natural and political observations upon the bills of mortality*.

Le *Traité du triangle arithmétique, avec quelques autres petits traiteux sur la mesme matière*, le célèbre ouvrage posthume de Pascal relatif aux probabilités est publié trois ans plus tard, bien qu'il ait été écrit en 1654.

En 1680, Leibniz aborde, lui aussi, les questions de probabilité de survie dans son *Essai de quelques raisonnements nouveaux sur la vie humaine*.

Peu de travaux, en fait, furent consacrés aux probabilités durant les trois ou quatre décennies qui suivirent les échanges Pascal-Fermat et le mémoire de Huygens. Les mathématiciens, à cette époque, se consacraient essentiellement à la découverte et au développement du calcul infinitésimal.

## 1700-1750

### *Histoire politique*

Le XVIII<sup>e</sup> siècle a été appelé le *Siècle des Lumières* du fait des importantes transformations de la pensée philosophique en Europe (Le terme *Aufklärung* a été utilisé en Allemagne et celui *Enlightenment*, en Angleterre et aux États-Unis). Des philosophes comme Voltaire et Rousseau ont redéfini la place de l'homme dans la société à partir de principes naturels. Ces idées ont conduit à l'avènement de la démocratie, en Angleterre, aux États-Unis avec la révolution américaine, et en France avec la Révolution française. La devise *Liberté, Égalité, Fraternité* a eu un fort impact sur l'évolution du reste de l'Europe.

La période 1700-1750 commence avec la guerre de succession d'Espagne qui débute en 1701 et s'achèvera en 1713 par le traité d'Utrecht. Philippe d'Anjou, petit-fils de Louis XIV, deviendra alors Philippe V d'Espagne. Louis XIV meurt en 1715 à l'âge de 77 ans et la régence est confiée à Philippe d'Orléans. Deux ans plus tard, l'Angleterre, les Provinces Unies et la France font alliance contre l'Espagne. Finalement, la France déclare la guerre à l'Espagne en 1719. Ce n'est qu'en 1721 que le traité de Madrid sera signé entre l'Angleterre, la France et l'Espagne.

En 1719, la Russie devient, grâce au traité de Stockholm une grande puissance : elle reçoit les États Baltes mais rend la Finlande à la Suède.

En France, Law est nommé en 1720 contrôleur des finances mais six mois plus tard la faillite de son système basé sur le papier-monnaie, entraîne des émeutes. Toujours en France, la franc-maçonnerie est introduite en 1721. Louis XV est sacré à Reims en 1722, et il épouse Marie Leszczyńska trois ans



plus tard. Son règne personnel débutera en 1743 et Madame de Pompadour deviendra sa favorite en 1745.

À l'est de l'Europe, Pierre II devient tsar de Russie en 1727 ; la guerre de succession de Pologne survient en 1733 et durera jusqu'en 1738, date de la signature du traité de Vienne ; trois ans plus tard on assiste à la guerre de succession d'Autriche, qui durera jusqu'en 1748. Après son succès, en 1745, contre les Anglais (victoire de Fontenoy), la France occupe la même année les Pays-Bas.

Sur le plan des idées, George Berkeley pose, dans son ouvrage *A treatise concerning the principles of Human Knowledge*<sup>1</sup> de 1710, la question « Le monde extérieur existe-t-il ? ». De son côté, David Hume, dans son *Treatise of human nature* publié en 1740, affirme que l'existence est faite de perceptions. Montesquieu publie en 1750 *l'Esprit des lois* qui tente de dégager les principes fondamentaux et la logique des différentes institutions politiques.

## Évolution scientifique

Au plan scientifique, malgré le fait que la grande période de créativité de Newton soit passée, son renom grandit. Il publie, en 1704, son célèbre livre d'optique : *Opticks : or a treatise of the reflexions, refractions, inflexions and colours of light. Also two treatises of the species and magnitude of curvilinear figures*. Sept ans plus tard (1711) il fait paraître son traité d'analyse supérieure *Analysis per quantitatum series, fluxiones, ac differentias : cum enumeratione linearum tertii ordinis*. L'année suivante c'est le fameux *Commercium epistolicum*, livre destiné à démontrer la priorité de sa découverte du calcul infinitésimal par rapport à Leibniz, qui est rendu public. Isaac Newton meurt en 1727, à l'âge de 85 ans. Son traité sur les fluxions est traduit en anglais en 1736 sous le titre : *The method of fluxions and infinite series ; with its application to the geometry of curve-lines*.

En 1729, Pierre Bouguer fonde la photométrie, une nouvelle branche de l'Optique concernée par la mesure des intensités lumineuses, en rendant public son ouvrage *Essai d'Optique sur la gradation de la lumière*<sup>2</sup>.

Au cours d'une expédition qui s'étend de 1735 à 1744, La Condamine mesure, en Amérique du Sud, un arc de méridien près de l'équateur, précisément pour vérifier un aspect des travaux de Newton prédisant que la Terre est aplatie aux pôles et gonflée à l'équateur.

En 1738, Daniel Bernoulli fait paraître son fameux livre sur l'hydrodynamique : *Hydrodynamica, sive de viribus et motibus fluidorum commentarii*, qui anticipe d'un siècle les travaux en physique statistique. De son côté Jean d'Alembert

1. Dublin, 1710.

2. Paris, Jombert, 1729.

publie, en 1743, son *Traité de dynamique, dans lequel les loix de l'équilibre & du mouvement des corps sont réduites au plus petit nombre possible*, livre qui marque un premier pas vers la mécanique lagrangienne et hamiltonienne. En mathématiques, Leonhard Euler écrit deux grands livres d'analyse : *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici latissimo sensu accepti* (1744) et *Introductio in analysin infinitorum* (1748).

En sciences de la nature, la nutrition des plantes est décrite par Stephen Hales en 1731 et Linné publie 1735<sup>1</sup> sa classification naturelle (*Systema Naturae*).

### Faits marquants en calcul des probabilités

En probabilités, la première moitié du XVIII<sup>e</sup> siècle va être une période extrêmement importante, marquée par la parution de plusieurs traités. Le premier ouvrage totalement consacré aux probabilités, *Essai d'analyse sur les jeux de hazard*, de Pierre Remond de Montmort paraît à Paris en 1708. Mais l'ouvrage fondamental de Jakob Bernoulli, l'*Ars conjectandi* avait été rédigé, sinon publié, plusieurs années auparavant<sup>2</sup>. Sa parution en 1713 ouvre la voie au développement de la théorie du calcul des probabilités. Abraham de Moivre publie, quant à lui, dans les *Philosophical Transactions*, en 1712, un long mémoire intitulé *De mensura sortis* qui est la première version de son futur livre *The doctrine of chances*<sup>3</sup>.

En 1709, la thèse de Nicholas I Bernoulli *De usu arte conjectandi in jure* est une des premières applications des probabilités au domaine du droit.

Dans le cadre d'une application fort différente, la mise en évidence de la constance du rapport du nombre des naissances des garçons à celui des filles est utilisé par John Arbuthnot, en 1710, pour tenter de prouver la Divine Providence dans une publication intitulée *An argument for Divine Providence*.

En 1718, Abraham de Moivre fait paraître à Londres un important ouvrage intitulé : *The doctrine of chances : or, a method of calculating the probabilities of events in play*. Il fera paraître ensuite, en 1725, un livre relatif aux rentes et assurances sur la vie : *Annuities upon lives*. De Moivre rend ensuite public un travail sur les séries, qui permet d'envisager la formulation de la loi normale grâce au développement du binôme (formule de De Moivre-Stirling) dans

1. Leiden, de Groot, 1735.

2. Le grand traité de Jakob Bernoulli, mort en 1705, *Ars conjectandi opus posthumum. Accedit tractatus de seriebus infinitis, et epistola gallice scripta de ludo pilae reticularis* paraît en 1713 édité par les soins de son neveu Nicholas Bernoulli.

3. Simpson publiera, en 1740, un court traité intitulé *The nature and laws of chance*, où il reprend les résultats de De Moivre énoncés dans *Doctrine of Chances*.

un livre intitulé : *Miscellanea analytica de seriebus et quadraturis*, qui paraît à Londres en 1730, qu'il complète par un papier, daté du 13 novembre 1733, dont le titre est explicite : *Approximatio ad summam terminorum binomii  $(a + b)^n$  in seriem expansi*.<sup>1</sup> Il s'agit de la première publication de ce que nous appelons aujourd'hui la *Loi normale*, résultat qui est devenu l'outil le plus puissant de la théorie des probabilités pour les deux siècles qui ont suivi.

Signalons la parution, en 1722, du livre de Roger Cotes *Harmonia mensurarum, sive analysis & synthesis per rationum & angulorum mensuras promotae*, qui comporte la première étude de la combinaison des observations et de la minimisation des erreurs. En 1733, Buffon soumet à l'Académie des sciences un célèbre mémoire sur le jeu de Franc-carreau qui sera publié longtemps après (1777) dans son *Essai d'arithmétique morale*. En 1738, Daniel Bernoulli publie son mémoire *Specimen theoriae novae de mensura sortis* où il discute du célèbre *Paradoxe de Saint-Petersbourg* et propose son concept d'espérance morale. En 1746, Antoine Deparcieux publie son *Essai sur les probabilités de la durée de la vie humaine ; d'où l'on déduit la manière de déterminer les rentes viagères, tant simples qu'en tontines : précédé d'une courte explication sur les rentes à terme, ou annuités ; et accompagné d'un grand nombre de tables*, qui vise la détermination des rentes viagères et comporte la première table de mortalité française.

On peut donc affirmer qu'au milieu du XVIII<sup>e</sup> siècle, le calcul des probabilités est une discipline bien établie avec les travaux de Jakob Bernoulli, et ceux de De Moivre.

## 1750-1800

### *Histoire politique*

Cette période débute avec la naissance de Louis XVI à Versailles en 1754. En 1756 débute la guerre de sept ans (1756-1763) provoquée par l'affrontement entre la Prusse et l'Autriche. Cette guerre oppose d'abord la France et la Grande Bretagne d'une part, l'Autriche et la Prusse d'autre part, mais on assiste au renversement des alliances : la France se rapproche de l'Autriche et s'oppose à l'Angleterre maintenant alliée à la Prusse. Louis XV est victime d'un attentat le 5 janvier 1757, son auteur, Damiens, est condamné à l'écartèlement et au bûcher et la sentence sera exécutée dans des conditions particulièrement atroces. À l'Est de l'Europe, Catherine II devient impératrice de Russie en 1762, tandis qu'en Autriche, Joseph II, frère de Marie Antoinette, devient

1. Ce mémoire a été imprimé en quelques exemplaires pour les amis de De Moivre. On en connaît actuellement trois copies (Prussische Staatbibliothek de Berlin, University College de Londres et Bibliothèque Macclesfield, n° 1424 du catalogue).

empereur du Saint Empire Germanique en 1765. Son règne durera jusqu'en 1790. Entre-temps, la Pompadour, favorite de Louis XV, meurt en 1764.

La France achète la Corse à la république de Gênes le 15 mai 1768. Bonaparte, qui naît le 15 août 1769 à Ajaccio, est ainsi un sujet de Louis XV. En cette même année 1768 meurt Marie Lesczinska, reine de France. 1773 voit la naissance de Louis Philippe, tandis que l'année suivante meurt Louis XV à qui succède son petit-fils, Louis XVI, âgé de vingt ans.

Hors du royaume, on assiste, en 1772, au partage de la Pologne entre la Russie, la Prusse et l'Autriche. Le 4 juillet 1776, treize États d'Amérique du Nord déclarent leur indépendance ; ce jour sera choisi comme fête nationale des États-Unis d'Amérique. L'année suivante Lafayette, qui n'a que vingt ans, se joint aux insurgés américains contre l'Angleterre. Louis XVI envoie en Amérique, en 1780, un corps expéditionnaire de 6 000 hommes, dirigé par Rochambeau. Finalement, l'Angleterre reconnaît l'indépendance des États-Unis d'Amérique en 1783, par le traité de Versailles.

En France des émeutes secouent en 1784 la Normandie. Elles sont la conséquence de mauvaises récoltes conjuguées à une pression des impôts insupportable. Deux ans plus tard ce sont les canuts de Lyon qui se soulèvent (ce qu'ils avaient déjà fait une première fois en 1744), c'est en quelque sorte l'avant-garde du mouvement ouvrier en France.

La Révolution française débute en 1789 avec la réunion des États généraux et la prise de la Bastille et s'achève avec le coup d'État du 18 brumaire (9-10 novembre 1799) de Bonaparte. Elle marque la fin de l'Ancien Régime et le passage à une monarchie constitutionnelle puis à la première République. Son impact a été considérablement accru par les guerres de la Révolution française et de l'Empire, dont il résultera des « républiques sœurs » en Europe continentale et la transformation des frontières et des États de l'Europe.

Bonaparte, Général de la Révolution française à 24 ans, remporte une victoire spectaculaire en Italie en 1796. Il est envoyé ensuite en Égypte, où le succès politique ne sera pas au rendez-vous, par contre l'expédition d'Égypte réalisée par les 160 savants qu'il emmène avec lui, sera une réussite sans précédent. A son retour, il prend le pouvoir par le coup d'État du 18 brumaire an VIII (9 novembre 1799).

En marge de ces événements politiques notons la mort de Montesquieu en 1755, à 66 ans et celle de Diderot en 1784, à 71 ans.

Des événements scientifiques spectaculaires ont par ailleurs lieu pendant cette période, citons-en deux : le voyage autour du monde de Bougainville entre 1766 et 1769, et le premier vol en ballon des frères Montgolfier en 1783.

Sur le plan des idées, d'Holbach publie en 1770 son ouvrage *Système de la nature*, considéré depuis comme la Bible du matérialisme. L'année suivante paraît la première édition de l'*Encyclopedia Britannica*. Immanuel Kant publie, en 1781, son très important ouvrage *La critique de la raison pure*<sup>1</sup>. Dix ans plus tard, Thomas Paine rend public son livre sur les droits de l'Homme<sup>2</sup>, cependant que Mary Wollstoncraft revendique, un an après, des droits pour la Femme<sup>3</sup>. Enfin, Thomas Robert Malthus publie en 1798 son *Essay on the principle of population*.

## Évolution scientifique

En sciences, nous nous appuyons essentiellement sur les grandes publications pour illustrer la richesse de cette période. Cela commence avec la parution, en 1751, du premier volume de l'Encyclopédie de d'Alembert et Diderot (le dernier volume paraîtra en 1765). La même année, Benjamin Franklin publie son traité d'électricité : *Experiments and Observations on Electricity, Made at Philadelphia in America*. Deux ans plus tard (1753) le grand mathématicien et physicien Leonhard Euler écrit la première version de ses calculs de théorie lunaire<sup>4</sup>. En 1755 on peut lire, toujours sous sa plume, le premier traité complet de calcul différentiel<sup>5</sup>. Anticipons un peu en restant avec ce très grand savant : il écrira également, entre 1768 et 1770, un traité complet de calcul intégral<sup>6</sup>, puis, entre 1769 et 1771, un traité de dioptrique<sup>7</sup>. L'optique s'enrichit d'ailleurs, en 1760, de l'extension du domaine de la photométrie (inauguré en 1729 par Bouguer) par Johann Heinrich Lambert dans son ouvrage *Photometria sive de mensura et gradibus luminis, colorum et umbrae*<sup>8</sup>. En 1784, le français René Just Haüy ouvre un nouveau chapitre en physique, la cristallographie, grâce à la publication de son *Essai d'une Théorie sur la structure des cristaux, appliquée à plusieurs genres de substances cristallisées*<sup>9</sup>. En électricité, Charles Augustin Coulomb énonce en 1785 sa fameuse loi<sup>10</sup> : elle marque le sommet de l'électrostatique et sera toute de suite adoptée, malgré des conditions expérimentales extrêmement

1. *Critick der Reinen Vernunft*, Riga, 1781.

2. *Rights of Man*, London, 1791.

3. *A vindication of the rights of Woman*, London, 1792.

4. *Theoria motus lunae*, impensis Academiae imperialis scientiarum, 1753.

5. Leonhard Euler, *Institutiones Calculi Differentialis cum eius Usu in Analysisi Finitorum ac Doctrina Serierum*, Petersburg, Academiae Imperialis Scientiarum, 1755.

6. Leonhard Euler, *Institutionum Calculi Integralis*, Petersburg, Academia Imperialis Scientiarum, 1768-1769-1770.

7. Leonhard Euler, *Dioptricae*, Petersburg, Impensis Academiae Imperialis Scientiarum, 1769-1770-1771.

8. Augsburg C., *Detleffsen for the widow of Eberhard Klett*, 1760.

9. Paris, Gogué & Née de La Rochelle, 1784.

10. Charles Augustin Coulomb, *Mémoires sur l'électricité et le magnétisme*, Paris, Bachelier, 1785-1789.

déliçates, car c'est une loi newtonienne. En Allemagne, Ernst Florens Friedrich Chladni publie le premier traité important d'acoustique en 1787<sup>1</sup>. Joseph-Louis Lagrange, en publiant en 1788 sa *Mécanique analytique*<sup>2</sup>, fait faire un pas de géant à la mécanique, mais aussi à la physique puisque son formalisme (et le formalisme hamiltonien, qui en découle), servira de base à nombre de théories fondamentales. De leur côté, Luigi Galvani, avec son *De viribus electricitatis in motu musculari*, publié en 1791, puis son « opposant » Alessandro Volta, qui publie en 1800 *On the electricity excited by the mere contact of conducting substances of different kinds*<sup>3</sup>, introduisent la notion de courant électrique (comme l'appellera plus tard Ampère). Avec l'invention de la pile électrique de Volta l'électricité n'est plus limitée aux phénomènes statiques, l'électrocinétique fait son apparition. Cette période s'achève par la publication de quatre grands ouvrages en mathématiques et en astronomie. Le premier, *Essai sur la théorie des nombres*, est signé en 1798 d'Adrien Marie Legendre. Le second est une œuvre monumentale en 5 volumes, qui débute en 1798 et se termine en 1827, il s'agit du fameux *Traité de mécanique céleste* de Pierre Simon Laplace. Le troisième est l'énoncé, en 1799, du théorème fondamental de l'algèbre par Carl Friedrich Gauss dans : *Demonstratio nova theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse*. Le quatrième est la *Géométrie descriptive* de Gaspard Monge<sup>4</sup>.

Notons également, durant cette période, deux événements marquants : la naissance de l'anthropologie (~ 1775) avec l'ouvrage *De generis humani varietate natura* de J. F. Blumenbach et celle de la chimie moderne avec l'édition, en 1789, du *Traité élémentaire de chimie* d'Antoine Lavoisier.

### Faits marquants en calcul des probabilités

En probabilités, la période est marquée par la parution de plusieurs ouvrages importants. Thomas Simpson publie, en 1755, dans les *Philosophical Transactions* de Londres, *The Advantage of taking the Mean of a Number of Observations in practical Astronomy*, mémoire où il aborde le problème des observations astronomiques en se focalisant sur les erreurs de positionnement du corps observé. J. H. Lambert, que nous avons rencontré plus haut, est le premier à utiliser la méthode du maximum de vraisemblance dans son ouvrage *Photometria* de 1760. Thomas Bayes fait naître la notion d'inférence statistique en 1763 dans son *Essay towards solving a problem in the doctrine*

- 
1. Ernst Florens Friedrich Chladni, *Entdeckungen ueber die Theorie des Klanges*, Leipzig, 1787.
  2. Paris, Veuve Desaint, 1788.
  3. Trans. Royal Soc. London, 1800.
  4. *Leçons données aux écoles normales de l'an III de la République*, Paris, Baudouin, 1799.

*of chances*<sup>1</sup>. Pierre Simon Laplace écrit, entre 1774 et 1785, divers mémoires sur les probabilités qui seront repris dans son ouvrage de 1812. De son côté Joseph Louis Lagrange publie, en 1776, un mémoire intitulé *Sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations*. L'année suivante c'est au tour de Buffon de produire un *Essai d'arithmétique morale*. En 1785 on trouve, sous la plume de D'Alembert son célèbre article intitulé *Croix-Pile*, dans l'Encyclopédie méthodique 1785. La même année Condorcet publie son *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Vers 1790, il rédige ses *Elémens du calcul des probabilités et son application aux jeux de hasard* qui seront rendus publics, à titre posthume, en 1805.

## 1800-1850

### Histoire politique

Bonaparte devient Premier consul le 10 novembre 1799. La Constitution de l'an VIII (1799) lui donne l'essentiel du pouvoir. Il consolide les acquis de la Révolution, en s'appuyant sur des notables, souvent de riches propriétaires ; il installe un préfet qui le représente dans chaque département ; il crée une monnaie solide et stable<sup>2</sup> ; il fait rédiger le Code Civil<sup>3</sup> ; il rétablit la paix religieuse en 1801 en signant un Concordat avec le pape ; il met fin au conflit avec l'Angleterre en signant, en 1802, la paix d'Amiens. Malheureusement, la guerre contre l'Angleterre reprend à partir de 1803 et le 1<sup>er</sup> Consul tente d'isoler ce pays par un Blocus continental qui interdit tout commerce. Entre-temps, en 1802, Bonaparte est nommé consul à vie. Le 18 mai 1804, il devient Empereur des Français<sup>4</sup>, sous le nom de Napoléon I<sup>er</sup>. En 1805, c'est le désastre naval de Trafalgar. Les Anglais s'assurent une prépondérance sur les mers et Napoléon est contraint de se rabattre sur le continent. Mais à partir de la victoire d'Austerlitz, le 2 décembre 1805, il devient le maître de l'Europe et enchaîne les victoires : Iéna (1806) ; Eylau (1807) ; Wagram (1809), ce qui lui permet d'établir un Grand Empire. Toutefois la révolte des Espagnols en 1808 est une première alerte, mais le vrai déclin commence en 1812 avec la campagne de Russie qui est un désastre militaire total. À la suite de cette défaite, une coalition générale se forme contre la France. L'armée française est battue à Leipzig en octobre 1813, puis la France est envahie en 1814. Napoléon est obligé d'abdiquer le 6 avril 1814. Il est exilé sur l'île d'Elbe<sup>5</sup>. Il revient au pouvoir pendant les « Cent jours », du 20 mars au 22 juin 1815,

---

1. Philosophical Trans., vol. 53, p. 370-418, London, 1763.

2. Le Franc germinal en 1803 est stable jusqu'au début du XX<sup>e</sup> siècle.

3. Instauré en 1804.

4. Il est sacré à Notre Dame le 2 décembre 1804.

5. Entre la Corse et l'Italie.

mais, suite à la défaite de Waterloo, le 18 juin 1815, il abdique définitivement avant d'être exilé à Sainte Hélène où il meurt en 1821<sup>1</sup>. Pendant l'exil de Napoléon sur l'île d'Elbe, en 1814, le frère de Louis XVI, Louis XVIII, monte sur le trône : c'est la Première Restauration. La Seconde restauration débute après les cent jours. En 1815, le Traité de Paris ramène les frontières de la France à ce qu'elles étaient en 1790.

Outre Atlantique, Simon Bolivar bat les Espagnols en 1819 et crée la république de Colombie. En France Louis XVIII meurt en 1824 à l'âge de 69 ans. Son frère, le comte d'Artois, lui succède sous le nom de Charles X.

En Russie, le Tsar Nicolas I<sup>er</sup> règne de façon autocratique pendant trente ans, de 1825 à 1855.

En France, les 3 glorieuses (27, 28 et 29 juillet 1830), commencent avec les barricades et se terminent avec l'abdication de Charles X. Louis Philippe I<sup>er</sup> le remplace, et devient roi des Français (et non roi de France), c'est la monarchie de juillet qui s'étend de 1830 à 1848.

En Angleterre c'est l'époque du règne de Guillaume IV (1830 à 1837). Il sera suivi par long règne de la reine Victoria (1837-1901).

L'Europe est atteinte, en 1831, par une dramatique épidémie de choléra.

Varsovie est prise par les Russes en 1831 et la Pologne devient province Russe.

En 1836, Louis Napoléon Bonaparte, neveu de Napoléon I<sup>er</sup>, prétendant au trône, tente de soulever un régiment à Strasbourg, mais il est arrêté et expulsé au Brésil.

En 1845, une grande famine conduit deux millions d'irlandais à émigrer aux États-Unis.

En France, l'année 1848 est marquée, comme en 1830, par les barricades (février 1848). Notons que François Arago devient membre du gouvernement provisoire de 1848, il est nommé le 27 février 1848 ministre de la Marine et de la Guerre. Il institue, le 4 mars 1848, une commission chargée de l'émancipation des esclaves dans les colonies et signe le décret d'abolition de l'esclavage le 27 avril 1848. Arago exerce même, du 5 mai au 25 juin 1848, les fonctions de chef de l'État. Mais les émeutes de juin 1848 mettent fin à la révolution. Après l'insurrection de juin 1848, Arago, profondément affecté et malade se retire à l'Observatoire. Cela fera dire à Victor Hugo, dans les *Choses vues* : « Arago ne paraît plus à l'assemblée. Quand on a ces deux spécialités de regarder le ciel et de regarder la terre, je comprends qu'on préfère le ciel ».

---

1. Les cendres de Napoléon seront transférées aux Invalides en 1840.



Louis Napoléon Bonaparte devient, le 10 décembre 1848, le premier président de la république française. Il est élu (avec 74 % des voix) au suffrage universel.

Signalons, pendant cette période, la disparition de grands artistes : Beethoven meurt à Vienne en 1827, à l'âge de 57 ans ; Goya meurt l'année suivante à 82 ans ; tandis que Goethe disparaît en 1832 à 83 ans.

Sur le plan des idées le Socialisme apparaît, en 1814, avec l'ouvrage de Robert Owen *A new view of Society*, tandis que la philosophie du Pessimisme voit le jour en 1819 avec un ouvrage d'Arthur Schopenhauer<sup>1</sup>. En 1832, Karl von Clausewitz publie ses célèbres idées sur la guerre *Vom Kriege*, et à la même époque Auguste Comte publie son Cours de Philosophie Positive (1830-1842).

## Évolution scientifique

En sciences, nous allons faire, comme pour la précédente période une rapide revue des publications marquantes. En 1801, Carl Friedrich Gauss publie ses fameuses recherches arithmétiques<sup>2</sup>. En physique le savant anglais prodige Thomas Young, introduit en 1802 la théorie ondulatoire de la lumière<sup>3</sup>, et publie également, en 1807, un traité de physique et de mécanique<sup>4</sup>. Cette théorie ondulatoire de la lumière relance le vieux débat de l'existence de l'éther. Celui-ci, supposé nécessaire à la propagation de la lumière, est-il entraîné par la Terre dans son mouvement annuel autour du Soleil ? Augustin Fresnel dans un document célèbre de 1818 adressé (à sa demande) à François Arago : *Lettre à M. Arago, sur l'influence du mouvement terrestre dans quelques phénomènes d'optique*<sup>5</sup>, écrit, près d'un siècle avant la naissance de la relativité restreinte : « La vitesse avec laquelle se propagent les ondes [lumineuses] est indépendante du mouvement des corps dont elles émanent ». Le même Fresnel, peut-être le plus grand opticien de tous les temps, développe, en se fondant sur l'optique ondulatoire, la théorie de la diffraction et publie, en 1819, un mémoire qui fera date : *Sur la diffraction de la lumière*<sup>6</sup>. En cette même année un professeur d'université danois, Hans Christian Ørsted, remarque l'influence du passage d'un courant électrique sur la direction d'une aiguille aimantée. Au début de 1820, il informe les principaux savants

1. *Die welt als Wille und Vorstellung*, Leipzig, 1819.

2. Carl Friedrich Gauss, *Disquisitiones arithmeticae*, Lipsiae, In Com. apud Gerh., Fleischer June, 1801.

3. Presque un siècle et demi après Huyghens et sensiblement en même temps, et indépendamment, de Fresnel.

4. Thomas Young, *A Course of Lectures on Natural Philosophy and the Mechanical Arts*, London, J. Johnson, 1807.

5. *Annales de chimie et de physique*, tome 9, 1818.

6. *Annales de chimie et de physique*, tome 11, 1819.

européens de sa découverte dans un mémoire de 4 pages écrit en latin : *Experimenta circa effectum conflictus electrici in acum magneticam*. Arago refait l'expérience devant les membres de l'Académie des sciences de Paris. André Marie Ampère assiste à la séance, se passionne pour ce phénomène et crée, en quelques mois, l'électrodynamique. Il publie de nombreuses notes, en particulier celle faisant état de l'interaction entre deux courants<sup>1</sup>.

Par ailleurs, John Dalton en 1808<sup>2</sup> et Avogadro en 1811 élaborent la notion d'atome et de molécule afin d'expliquer les réactions chimiques, mais ce schéma n'est adopté officiellement que vers 1860.

En mathématiques, Augustin-Louis Cauchy introduit la rigueur en analyse, son cours d'analyse de l'École Polytechnique de 1821, fait référence<sup>3</sup>. L'année suivante, c'est un autre professeur à l'École Polytechnique, Joseph Fourier, qui publie l'ouvrage considéré comme fondateur de la physique mathématique : la *Théorie analytique de la chaleur*<sup>4</sup>. On y trouve la notion de « transformée de Fourier » si féconde dans un grand nombre de domaines. Deux ans plus tard, en 1824, le polytechnicien Sadi Carnot publie ses *Réflexions sur la Puissance Motrice du Feu*<sup>5</sup>, c'est l'acte de naissance de la thermodynamique. En 1826, Ampère synthétise, dans son dernier grand ouvrage de physique, ses travaux en électrodynamique<sup>6</sup>. Toujours dans le domaine de l'électricité, Georg Simon Ohm, en conduisant un raisonnement analogue à celui de Fourier en thermique, aboutira à la fameuse « loi d'Ohm » et en publiera le résultat en 1827<sup>7</sup>. Malheureusement sur l'instant ses travaux n'ont pas le succès légitimement dû. Le développement de ce domaine se fait aussi bien sous l'angle théorique qu'expérimental. Côté théorie, George Green publie en 1828 le mémoire pour lequel il est le plus connu : *An essay on the application of mathematical analysis to the theories of electricity and magnetism*<sup>8</sup>. Côté expérimental, Michael Faraday, l'un des plus grands expérimentateurs de tous les temps, publie, entre 1832 et 1856, ses recherches<sup>9</sup> qui seront le

1. André Marie Ampère, *Mémoires sur l'action mutuelle de deux courans électriques, sur celle entre un courant électrique et un aimant ou le globe terrestre, et celle de deux aimans l'un sur l'autre*, Annales de Chimie et de Physique, 15, 1820, Paris.
2. John Dalton, *A new system of chemical philosophy*, Manchester and London, 1808.
3. Augustin-Louis Cauchy, *Cours d'Analyse de l'école royale polytechnique*, Paris, de Bure, 1821.
4. Jean Baptiste Joseph Fourier, *Théorie analytique de la chaleur*, Paris, Firmin Didot, 1822.
5. Nicolas Leonard Sadi Carnot, *Réflexions sur la Puissance Motrice du Feu et sur les machines propres à développer cette puissance*, Paris, Bachelier, 1824.
6. André Marie Ampère, *Théorie des phénomènes électrodynamiques uniquement déduite de l'expérience*, Paris, Méquignon-Marvis, 1826.
7. Georg Simon Ohm, *Die Galvanische Kette mathematisch bearbeitet*, Berlin, T. H. Riemann, 1827.
8. George Green, *An essay on the application of mathematical analysis to the theories of electricity and magnetism*, Nottingham, 1828.
9. Michael Faraday, *Experimental researches in electricity*, Phil. Trans., London, 1832-1856.

point de départ de la synthèse maxwellienne. Gauss, qui collabore avec Weber sur le magnétisme, fait une étude, publiée en 1833, extrêmement précise du magnétisme terrestre<sup>1</sup> et introduit, toujours avec Weber, la théorie du potentiel<sup>2</sup>. En mathématiques il faut souligner, à la fin des années vingt et au début des années trente, les travaux en géométrie non-euclidienne de Lobachevski<sup>3</sup>, ainsi que ceux de Bolay<sup>4</sup> ; l'introduction des déterminants par Cayley<sup>5</sup> en 1843 ; ainsi que la richesse de l'œuvre, hélas si rapidement et si tragiquement interrompue, d'Évariste Galois<sup>6</sup>.

En physique, Joule réalise, entre 1843 et 1847 différentes mesures permettant de déterminer l'équivalent mécanique de la chaleur<sup>7</sup>. Après les expériences mettant en jeu la conversion d'électricité en chaleur, il met au point une méthode qui consiste à faire tourner une roue à ailettes à l'aide de la chute de poids dans un récipient d'eau, hermétique, dont il mesure l'élévation de température. Les travaux de Joule sont complétés par l'analyse d'Helmholtz qui établit, en 1847, les bases mathématiques de la conservation de l'énergie<sup>8</sup>. En chimie, naît en 1842 une nouvelle discipline, la chimie organique, avec le traité Justus Liebig.

Notons également, durant cette période, plusieurs inventions d'une très grande importance. En 1816, Nicéphore Niepce obtient la toute première photographie<sup>9</sup> grâce à sa chambre obscure, mais il s'agit d'une image positive, non fixée. La plus ancienne photo qui nous soit parvenue est de Niepce, il s'agit du *Point de vue du Gras* de 1827. Louis Daguerre améliorera le procédé, rapidement appelé daguerréotype, lequel sera présenté par Arago devant l'Académie des sciences le 19 août 1839.

Le tout premier train emportant des voyageurs date de 1804, mais les chemins de fer ne commencent à être vulgarisés que dans les années 1815. En France la première ligne relie Paris et Saint-Germain en 1837. Ampère annonce à l'Académie des Sciences le 2 octobre 1820 que « l'on pourrait former une

1. Carl Friedrich Gauss, *Intensitas vis magneticae terrestris ad mensuram absolutam revocata*, Göttingen, Dieterich, 1833.
2. Carl Friedrich Gauss, Wilhelm Weber, *Resultate aus den Beobachtungen des magnetischen Vereins im Jahre*, Göttingen, Dieterichschen Buchhandlung, 1836-1840.
3. Ivanovich Lobachevski, *Théorie des Parallèles*, Presses de l'Université Kazan, 1829.
4. Farkas et Janos Bolay, *Tentamen juventutem studiosam in elementa matheosos purae (Appendix) Scientiam spatii absolute veram exhibens*, Maros Vasarhelyini, 1832-1833.
5. Arthur Cayley, *On the theory of determinants*, Trans. Cambridge Phil. Soc. VIII, 1843.
6. Évariste Galois, *Œuvres mathématiques*, Journal de Liouville, Tome XI, Paris, 1846.
7. James Prescott Joule, *On the Calorific Effects of Magneto-Electricity, and the Mechanical Value of Heat*, 1843-1844.
8. Hermann Helmholtz, *Ueber der Erhaltung der Kraft, eine physikalische Abhandlung*, Berlin, Reimer, 1847.
9. Il écrit le 5 mai 1816 : « Le fond du tableau est noir, et les objets sont blancs, c'est-à-dire plus clairs que le fond » (!), il ajoute : « La possibilité de peindre de cette manière me paraît à peu près démontrée ».

sorte de télégraphe », mais le premier qui sera opérationnel (sur moins de 2 km) est celui de Gauss et Weber (1833). Le télégraphe commence vraiment son développement en Angleterre avec Charles Wheatstone en 1838, puis surtout avec Morse aux États-Unis qui réalise la première ligne télégraphique américaine entre Baltimore et Washington en 1844.

### **Faits marquants en calcul des probabilités**

En probabilités, Adrien Marie Legendre est l'auteur, en 1805, d'une première publication sur les moindres carrés<sup>1</sup>. Elle est suivie par celle de Gauss<sup>2</sup>, sur le même thème, en 1809. En 1812, Pierre Simon Laplace fait apparaître sa très importante *Théorie analytique des probabilités*<sup>3</sup>. Notons également le mémoire de Poisson, en 1837, sur ses *Recherches sur la probabilité des jugements*<sup>4</sup>. Quand à Auguste Bravais, il soumet à l'Institut, en 1838, un mémoire intitulé : *Analyse mathématique sur les probabilités des erreurs de situation d'un point*. Cournot publie, en 1843, son *Exposition de la théorie des chances et des probabilités* qui est une tentative d'interprétation des fondements du calcul des probabilités. Terminons, un peu en marge de cette période, par le *Mémoire sur la probabilité des erreurs d'après la méthode des moindres carrés*, publié par Bienaymé en 1851 et qui anticipe la théorie statistique et la loi du  $\chi^2$ .

Le premier quart du dix-neuvième a donc été, à l'évidence, marqué, dans le domaine des probabilités, par les travaux de Laplace et de Gauss et l'énoncé du théorème central limite.

## **1850-1900**

### **Histoire politique**

La seconde république se termine le 2 décembre 1851, elle n'aura duré qu'un peu plus de trois ans. Le coup d'État de Louis Napoléon Bonaparte lui permet de mener la restauration impériale à son profit. En janvier 1852, la France se dote d'une nouvelle constitution qui porte la durée du mandat du président à 10 ans. Le 7 novembre 1852, le régime impérial est rétabli. Louis-Napoléon Bonaparte devient Napoléon III, c'est le début du second Empire, qui durera jusqu'en 1870. Cette époque est marquée par la prospérité, la croissance et les conquêtes coloniales. Napoléon III, en inaugurant

1. Adrien Marie Legendre, *Nouvelles méthodes pour la détermination des orbites des comètes*, Paris, Firmin Didot, 1805.
2. Carl Friedrich Gauss, *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*, Hamburg, Frid. Perthes et I. H. Besser, 1809.
3. Pierre Simon Laplace, *Théorie analytique des probabilités*, Paris, Courcier, 1812 (2<sup>e</sup> éd. 1814 ; 3<sup>e</sup> éd. 1820).
4. Paris, Bachelier, 1837.

l'exposition universelle de Paris de 1855, veut montrer à l'Angleterre que la France peut également organiser une exposition aussi grandiose que celle de Londres de 1851. Il inaugurera également l'exposition universelle de Paris de 1867 qui ne comptera pas moins de 52 000 exposants pour 41 pays représentés. En 1854, après l'attaque par les Russes des territoires turcs de Moldavie et de Valachie, la France, alliée à l'Angleterre, déclare la guerre à la Russie. Cette guerre se terminera en 1856 avec la victoire de l'Alma. L'année suivante, en Crimée, Sébastopol est prise par les forces anglaises et françaises. Le Traité de Paris mettra fin à la guerre de Crimée, qui avait commencé en 1853. En 1855 la Russie couronne son nouveau Tsar, Alexandre II. Il régnera jusqu'en 1881 (il est assassiné par de jeunes monarchistes le 13 mars 1881, jour où il s'appropriait à donner une nouvelle constitution à son pays). Tandis qu'on assiste en 1857 à la guerre franco-anglaise contre la Chine, les Français s'établissent à Saigon en 1859. Cette même année, en accord avec le ministre italien Cavour, la France déclare la guerre à l'Autriche. L'armée franco piémontaise est victorieuse à Magenta. En Prusse, Guillaume I devient roi en 1861 et le restera jusqu'en 1888. En Italie, c'est Victor-Emmanuel II qui devient, en 1861, roi de l'Italie unifiée.

Aux États-Unis cette année est celle du début de la guerre de sécession. Toujours en Amérique, la France, l'Angleterre et l'Espagne interviennent au Mexique. Le monde bouge : Bismarck devient le premier ministre de la Prusse en 1862 ; le droit de grève est reconnu en France en 1864 ; aux États-Unis l'année 1865 est marquée par la victoire des nordistes et par l'assassinat du président Lincoln ; en Russie Marx publie, en 1867, *Le capital*.

En France, après un plébiscite approuvant des réformes libérales, l'Empire libéral est instauré en 1870. Par ailleurs c'est le début de la guerre franco-prussienne au cours de laquelle la France perdra l'Alsace-Lorraine. Après la défaite de Sedan Napoléon III se rend. Le siège de Paris commence. C'est en France, dans la galerie des glaces du château de Versailles, le 18 janvier 1871, que Guillaume I<sup>er</sup>, cinquième roi de Prusse (1861-1888), est proclamé empereur du II<sup>e</sup> Reich<sup>1</sup> d'Allemagne. Le 28 janvier de la même année Paris assiégé capitule. Le 18 mars c'est la Commune de Paris et l'incendie des Tuileries. La situation se rétablit ensuite peu à peu et le 31 août 1871, Thiers devient président de la République. La fin de l'occupation allemande n'est effective qu'en 1873, tandis que Napoléon III meurt le 9 janvier de cette année en exil en Angleterre.

Deux ans plus tard c'est la naissance de la III<sup>e</sup> République. En 1881, l'Égypte se révolte contre la présence franco-anglaise sur le Canal de Suez. En France Jules Ferry rend l'école primaire obligatoire en 1882. En 1887 on assiste à

---

1. Empire.

des changements à la tête des États : Sadi Carnot<sup>1</sup> devient président de la République française (il sera assassiné par l'anarchiste Caseris en 1894), tandis que Guillaume II devient Empereur d'Allemagne. La fin de cette période est marquée par l'affaire Dreyfus, qui débute en 1894.

Les décennies qui précèdent la première Guerre Mondiale correspondent en France à la « Belle Époque » souvenir d'une période dorée. Il s'agit toutefois d'une période politiquement turbulente, avec le militantisme de la classe ouvrière, les mouvements socialistes organisés, et l'affaire Dreyfus, qui s'étend de 1894 à 1906, et qui divise le pays entre la gauche et la droite antisémite. Les nouvelles inventions telles que l'électricité et la vaccination contre la maladie ont facilité la vie à tous les niveaux sociaux. Au début de cette période, en 1862, Jean Henri Dunant a créé la Croix-Rouge.

La scène culturelle a, pour sa part, prospéré et a pris de nouvelles formes avec l'impressionnisme et l'Art Nouveau, les romans réalistes de Gustave Flaubert et d'Émile Zola, le cabaret et le cancan et, en 1895, la naissance du cinéma. C'est également l'époque des expositions universelles, la première a lieu à Londres en 1851, mais trois d'entre elles auront lieu à Paris, en 1855, en 1867 et en 1900.

## Évolution scientifique

En sciences, nous allons faire, comme pour les précédentes périodes une rapide revue des publications marquantes. Michael Faraday achève l'édition de ses *Experimental researches in electricity*<sup>2</sup>, qui joueront un rôle essentiel dans les recherches de Maxwell. En publiant *Lectures on Quaternions : Containing a Systematic Statement of a New Mathematical Method*<sup>3</sup>, William Rowan Hamilton développe un outil qui, dans un cas simplifié, conduira à l'intro-

1. Sadi Carnot est issu d'une brillante lignée où même le prénom ne réussit pas à lever l'ambiguïté, ce qui peut entraîner quelques confusions. Tout commence avec Lazare Carnot (1753-1823), lui-même issu d'une famille de 18 enfants. Il est chargé en 1793 des questions militaires au Comité de Salut Public, c'est alors que le surnom d'« Organisateur de la victoire » lui a été attaché. Membre du Directoire en 1795, il a joué le rôle de Ministre de la Guerre et a envoyé Bonaparte en Italie. Passionné de sciences, il est entré à l'Institut en mars 1799. Un an plus tard, il est nommé officiellement Ministre de la Guerre par Bonaparte. Il a également été Ministre de l'Intérieur dans le Gouvernement des Cent Jours (1815). Sadi Carnot (1796-1832), le second fils de Lazare et Sophie Carnot, sera un très grand physicien et est considéré comme le père de la thermodynamique. Hippolyte Carnot (1801-1887), le 3<sup>e</sup> fils de Lazare, devient au lendemain de la révolution de 1848 ministre de l'Instruction publique et des Cultes. Hippolyte, frère du physicien Sadi Carnot, a lui-même un fils en 1837 qu'il appelle également. Sadi. C'est ce Sadi (II) Carnot qui deviendra en 1887 le 4<sup>e</sup> président la République sous la III<sup>e</sup> République.
2. Phil. Trans., London, 1832-1856.
3. Dublin, Hodges and Smith, London, Whittaker & Co., 1853.

duction des vecteurs en physique<sup>1</sup>. Bernhard Riemann fait paraître, en 1857, sa très importante *Theorie des Abel'schen Funktionen*<sup>2</sup>. Notons également la publication, dix ans plus tard, de son mémoire *Ueber die Hypothesen, welche der Geometrie zu Grunde liegen*, qui développe la géométrie non-euclidienne et les espaces multidimensionnels, qui seront d'une grande importance en relativité générale. Dans un tout autre domaine, Charles Darwin fait paraître en 1859 l'ouvrage fondamental sur l'origine des espèces<sup>3</sup>. En 1866, Gregor Mendel, dans son mémoire *Versuche in Pflanzenhybriden*, interprète l'hérédité en terme de caractères dominants et/ou récessifs. Gustav Kirchhoff introduit l'analyse spectrale en 1861<sup>4</sup> et la développe avec Robert Bunsen, ce qui leur permet de découvrir le césium et le rubidium. Ces éléments trouveront leur place dans le cadre de la loi périodique des éléments chimiques introduite par Dimitri Mendeleïev en 1869. Toujours en chimie, notons que le brevet concernant la dynamite est déposé en 1866 par Alfred Nobel. Rudolf Clausius donne des bases extrêmement solides à la thermodynamique<sup>5</sup> en introduisant l'entropie et en énonçant de façon synthétique les deux principes. Un autre axe se dégage : la construction, par Maxwell et Boltzmann de la thermodynamique statistique. Parmi leurs nombreuses publications sur ce sujet, signalons-en deux : celle de James Clerk Maxwell en 1867 : *On the Dynamical Theory of Gases*<sup>6</sup> et celle de Ludwig Boltzmann en 1877 : *Ueber die Betziehung zwischen dem zweitem Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respektive den Sätzen ueber das Wärmegleichgewicht*<sup>7</sup>. La synthèse de l'électromagnétisme apparaît en 1873 dans l'ouvrage fondamental de Maxwell : *A Treatise on Electricity and Magnetism*<sup>8</sup>. Il ouvre un nouveau champ en physique. Hendrik Antoon Lorentz sera un des tout premiers maxwelliens. Il publie, dès 1875, un mémoire relatif à la réflexion et la réfraction de la lumière dans le cadre de la théorie électromagnétique<sup>9</sup>. D'autres grands mémoires suivront, citons *La théorie électromagnétique de Maxwell et son application aux corps mouvants*<sup>10</sup> ainsi que le célèbre « Versuch »<sup>11</sup>. En Allemagne, un autre

1. Signalons en particulier la contribution fondamentale de Josiah Willard Gibbs : *Elements of Vector analysis*, New Haven, 1881-1884.
2. Berlin, Reimer, 1857.
3. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, Londres, John Murray, 1859.
4. Gustav Kirchhoff, *Das Sonnenspectrum*, Berlin, 1861-1863.
5. Rudolf Clausius, *Abhandlungen ueber die mechanische Wärmetheorie*, Braunschweig, Viewig, 1864.
6. *Philosophical Transactions of the Royal Society*, vol. 157, London, Taylor & Francis, 1867.
7. S. Ber. Akad. Wien., 76/3, 1877.
8. Clarendon Press, Oxford, 1873.
9. Hendrik Antoon Lorentz, *Over de theorie der terugkaatsing en breking van het licht*, academisch proefschrift, Arnhem, K. van der Zande, Firma Stenfort Kroese, 1875.
10. Hendrik Antoon Lorentz, *La théorie électromagnétique de Maxwell et son application aux corps mouvants*, Archives néerlandaises des sciences exactes et naturelles, t. XXV, 1892.
11. Hendrik Antoon Lorentz, *Versuch einer Theorie der elektrischen und optischen Erscheinungen in bewegten Körpern*, Leiden, Brill, 1895.

très grand maxwellien est Heinrich Hertz qui d'une part travaille sur le formalisme<sup>1</sup>, de façon à le rendre plus facilement manipulable, et d'autre part démontre expérimentalement l'existence des ondes électromagnétiques<sup>2</sup>. Ces « ondes hertziennes » vont révolutionner les communications alors en plein essor. Le télégraphe électrique trouve un complément « analogique » avec l'invention, en 1876, du téléphone par Alexander Graham Bell. Entre cette date et 1878, J.W. Gibbs, dans son papier *On the equilibrium of heterogenous substances*, indique que lorsque l'entropie d'un système a atteint un maximum, le système reste dans un état d'équilibre. En 1888, Nicola Tesla brevète la découverte du courant alternatif, tandis qu'en 1895, Guglielmo Marconi transmet un signal radio sur 1 500 mètres. 1895 marque un tournant dans la physique. Röntgen découvre à la fin de cette année les rayons X<sup>3</sup>, tandis que Joseph John Thomson découvre, en 1897, l'électron et tente d'élucider la nature des mystérieux rayons X<sup>4</sup>. L'année 1896 voit la découverte fortuite de la radioactivité spontanée par Henri Becquerel<sup>5</sup>, puis le 26 décembre 1898, Pierre et Marie Curie publient, avec Gustave Bémont, une note à l'Académie intitulée *Sur une nouvelle substance fortement radio-active contenue dans la pechblend*<sup>6</sup>. Ils y annoncent la découverte d'une nouvelle substance radioactive qu'ils appellent radium. En janvier 1899, Rutherford publie un article intitulé *Uranium Radiation and the Electrical Conduction produced by it*<sup>7</sup> dans lequel il conclut que le rayonnement de l'uranium est complexe et comporte au moins deux types distincts de rayonnements, l'un très rapidement absorbé qu'il appela  $\alpha$  et l'autre plus pénétrant qu'il nomme  $\beta$ . Cette période s'achève avec l'introduction, par Max Planck, d'une véritable révolution en physique : l'introduction des quanta<sup>8</sup>. Nous ne pouvons pas ne pas évoquer également l'immense contribution de Poincaré à la mécanique céleste<sup>9</sup>, ni

1. Ce qui est également le cas, de façon encore plus évidente d'Oliver Heaviside. Voir par exemple ses *Electrical papers*, London, Macmillan, 1892.
2. Parmi les publications de Hertz, notons : *Ueber sehr schnelle electrische Schwingungen [avec] Nachtrag zu der Abhandlung über sehr schnelle electrische Schwingungen*, p. 421-448 and 543-544 in *Annalen der Physik und Chemie*, vol. 31, Leipzig, Johann Barth, 1887.
3. Wilhelm Röntgen, *Eine neue Art von Strahlen*, Würzburg Stahel, 1896.
4. Joseph John Thomson, *On the charge of electricity carried by the ions produced by Röntgen rays*, in *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Fifth Series, Volume XLVI, July-December, 1898, London, Taylor & Francis, 1898.
5. *Sur les radiations émises par phosphorescence*. Note de M. Henri Becquerel, CRAS, tome 122, p. 420-421, 1896.
6. Pierre et Marie Curie et Gustave Bémont, CRAS, tome 127, 1215-1217 (1898).
7. *Philosophical magazine*, ser. 5, xlvii, p. 109-163.
8. Max Planck, *Ueber des Gesetzes der Energieverteilung im Normalspectrum*, In *Verhandlungen der Deutschen physikalischen Gesellschaft im Jahre, 1900, Zweiter Jahrgang*, p. 237-45, Leipzig, Johann Ambrosius Barth, 1900.
9. Henri Poincaré, *Les méthodes nouvelles de la Mécanique Céleste*, Paris, Gauthier-Villars, 3 volumes : 1892, 1893, 1899.



son postulat, indiqué en 1898 dans son papier *De la mesure du temps*, de la constance de la vitesse de la lumière.

Notons également que c'est durant cette période que Claude Bernard publie, en 1865, son *Introduction à l'étude de la médecine expérimentale*<sup>1</sup>.

### **Faits marquants en calcul des probabilités**

À partir de 1840 naît une littérature traitant des probabilités sur le plan philosophique. On peut citer le *System of Logic* de John Stuart Mill<sup>2</sup> de 1843 et la *Logic of chances* de John Venn<sup>3</sup> de 1866. Les *Principles of Science* de W. S. Jevons<sup>4</sup> datent de 1874 et la *Grammar of Science* de Karl Pearson<sup>5</sup> de 1892.

Les statistiques se développent entre 1880 et 1900, avec Karl Pearson, Francis Galton, Francis Edgeworth et George Yule comme figures dominantes.

En probabilités proprement dites, George Boole publie, en 1854, *An investigation of the laws of thought on which are founded the mathematical theories of logic and probabilities*<sup>6</sup>. Deux mémoires importants sont dus à Panufty Chebyshev, l'un est consacré aux valeurs moyennes<sup>7</sup>, l'autre à deux théorèmes relatifs aux probabilités<sup>8</sup>.

En France, deux livres portant le même titre *Calcul des probabilités* paraissent en 1888 et en 1896. Ils sont dus respectivement à Joseph Bertrand et Henri Poincaré. Ces deux ouvrages qui sont la transcription des cours magistraux des deux auteurs serviront dans l'enseignement durant des décennies et seront réédités plusieurs fois.

Enfin en 1900, David Hilbert propose ses célèbres vingt problèmes, dont le sixième est l'axiomatisation de la physique et des probabilités.

---

1. Paris, Baillière, 1865.

2. London, Parker, 1843.

3. London, Macmillan, 1866.

4. London, Macmillan, 1874.

5. London, Scott, 1892.

6. London, Macmillan and Co., 1854.

7. *O Srednikh velichinakh*, 1867.

8. *O dvukh teoremakh otnositelno veroyatnostey*, 1887.

# Quelques publications

## ou livres relatifs à l'Histoire de la période pré-axiomatique des Probabilités

(liste non exhaustive classée par ordre alphabétique)

- E. Barbin et Y. Marec (1987) : « Les recherches sur la probabilité des jugements de Siméon-Denis Poisson », *Histoire et mesure*, vol. II, n° 2, éditions de l'EHESS, p. 39-58.
- E. Barbin *et al.* (2004) : *Histoires de probabilités et de statistiques*, Ellipses, coll. « IREM – Épistémologie et Histoire des Mathématiques ».
- D. Bessot *et al.* (2005) : *L'espérance du Hollandais ou le premier traité du calcul du hasard*, Ellipses, coll. « IREM – Épistémologie et Histoire des Mathématiques ».
- B. Bru (1981) : *À propos de l'histoire des statistiques au début du XIX<sup>e</sup> siècle : probabilités et statistiques des jugements*, Séminaire d'histoire des mathématiques, Université de Rennes I, 24 p.
- B. Bru (1981) : « Poisson, le calcul des probabilités et l'instruction publique », dans Métivier, M., Costabel, P. et Dugac, P., *Siméon-Denis Poisson et la science de son temps*, éditions de l'École polytechnique, p. 51-94.
- B. Bru (1988) : *Estimations laplaciennes. Un exemple : la recherche de la population d'un grand empire (1785-1812)*, Journal de la Société de statistiques de Paris, vol. 129, n° 1-2, p. 6-45.
- B. Bru (1988) : *Statistique et bonheur des hommes*, Revue de Synthèse, tome CIX, n° 1, p. 69-95, Centre international de Synthèse – Albin Michel.
- P. Crépel (1988) : « Condorcet, la théorie des probabilités et les calculs financiers », dans R. Rashed (dir.), *Sciences à l'époque de la Révolution Française – Recherches historiques*, Paris, A. Blanchard, p. 267-325.

- E. Coumet (1970) : *La théorie du hasard est-elle née par hasard ?*, in *Annales : économies, sociétés, civilisations*, vol. 25, n° 3, mai-juin 1970, p. 574-598.
- D. Courgeau (2004) : *Probabilités, démographie et sciences sociales*, Math. & Sci. hum. 42<sup>e</sup> année, n° 167, 2004 (3), p. 27-50.
- A. I. Dale (1982) : *Bayes or Laplace ? An examination of the origin and early applications of Bayes' theorem*, AHES, 27, 23-47.
- A. I. Dale (1991) : *A history of inverse probability from Thomas Bayes to Karl Pearson*, Springer.
- L.J. Daston (1981) : « Mathematics and the moral sciences: The rise and fall of the probability of judgments, 1785-1840 », in *Epistemological and social problems of the sciences in the early nineteenth century* (édité par A.N. Jahnke et M. Otte), Dordrecht, D. Reidel, p. 287-309.
- L.J. Daston (1988) : *Classical probability in the Enlightenment*, Princeton University Press.
- R. Dedekind, (1901) : *Gauss in seiner Vorlesung über die Methode der kleinsten Quadrate*, Festschr Kgl. Ges. Wiss., Göttingen, 45-59.
- J. Dutka (1990) : *Robert Adrain and the method of least squares*, AHES, 41, 171-184.
- J. Dutka (1988) : *On the Saint-Petersburg paradox*, AHES, 39, n° 1, p. 13-39.
- C. C. Gillispie et al. (1997) : *Pierre-Simon Laplace, 1749-1827. A Life in Exact Science*, Princeton University Press.
- C. Gouraud (1848) : *Histoire du calcul des probabilités depuis ses origines jusqu'à nos jours avec une thèse sur la légitimité des principes et des applications de cette analyse*, Paris, Durand.
- A. Hald (1990) : *Evaluations of the Beta probability integral by Bayes and Price*, Arch. Hist. Ex. Sci., 41, 139-156.
- A. Hald (1990) : *History of probability and statistics and their applications before 1750*, Wiley.
- A. Hald (1998) : *A history of mathematical statistics from 1750 to 1930*, Wiley.
- A. Hald (2006) : *A History of Parametric Statistical Inference from Bernoulli to Fisher, 1713-1935*, Springer.

- C. C. Heyde and E. Seneta (1977) : *I. J. Bienayme, Statistical Theory Anticipated*, Berlin, Springer Verlag.
- C. C. Heyde and E. Seneta (éd.) (2001) : *Statisticians of the Centuries*. New York & Berlin, Springer Verlag.
- M. Kendall, and R. L. Plackett, éd. (1977) : *Studies in the History of Statistics and Probability*, vol. 2, Griffin, London.
- L. Krüger *et al.* (1987) : *The probabilistic revolution*, Cambridge, M.I.T. Press.
- D. Lanier et D. Trotoux (1995) : *La loi des grands nombres et le théorème de De Moivre-Laplace*, Presses universitaires de Franche-Comté.
- P. Lévy (1949) : *Les fondements du calcul des probabilités*, Dialectica, tome 3, p. 55.
- L. E. Maistrov (1974) : *Probability theory, a historical sketch*, translated and edited by Samuel Kotz, London, Academic press.
- T. Martin (éd.) (2003) : *Arithmétique politique dans la France du XVIII<sup>e</sup> siècle*, Paris, éditions de l'I.N.E.D.
- M. Métivier *et al* (1981) : *Siméon-Denis Poisson et la science de son temps* (édité par M. Métivier, P. Costabel, P. Dugac), Palaiseau, École Polytechnique.
- E. C. Molina (1931) : *Bayes' theorem, an expository presentation*, The Annals of Mathematical Statistics, 2, 1, p. 23-37.
- M. Paty (1988) : « D'Alembert et les probabilités », in *Sciences à l'époque de la Révolution Française – Recherches historiques* (sous la direction de R. Rashed), Paris, A. Blanchard, p. 203-265.
- K. Pearson (1892) : *The Grammar of Science*, W. Scott, London
- K. Pearson (1936) : *Method of moments and method of maximum likelihood*, Biometrika, 28, 34-59.
- K. Pearson (1948) : *Karl Pearson's Early Statistical Papers*, (E. S. Pearson, éd.), Cambridge University Press, Cambridge.
- E. S. Pearson and M. Kendall, éd. (1970) : *Studies in the History of Statistics and Probability*, vol. 1. Griffin, London.
- K. Pearson (1978) : *The history of statistics in the 17th and 18th centuries, against the changing background of intellectual, scientific and religious thought, Lectures by Karl Pearson given at University College London during the academic sessions 1921-1933*, London, Charles Griffin.

- P. Picard (2007) : *Hasard et probabilités, histoire, théorie et application des probabilités*, Vuibert.
- T.M. Porter (1986) : *The rise of statistical thinking, 1820-1900*, Princeton, Princeton University Press.
- R. Pressat (2001) : *Christian Huygens et la table de mortalité de Graunt*, Math et Sci. hum, 39<sup>e</sup> année, n° 153, p. 29-36.
- J.-M. Rohrbasser et J. Véron (2001) : *Leibniz et les raisonnements sur la vie humaine*, Paris, Ined.
- P. Ruffini (1821) : *Riflessioni critiche sopra il saggio filosofico intorno alle probabilità del Sig. Laplace*, Modena.
- E. Seneta (1984) : *The central limit problem and linear least squares in pre-revolutionary Russia: The background*, Math. Scientist, 9, 37-77.
- Oscar. B. Sheynin (1966) : *Origin of the theory of errors*, Nature, vol. 211, p. 1003-1004.
- (1968) : *On the early history of the law of large numbers*, Biometrika, vol. 55, p. 459-467.
- (1970) : *D. Bernoulli on the normal law*, Biometrika, vol. 57, p. 199-202.
- (1971) : *Newton and the classical theory of probability*, AHES, vol. 7, p. 217-243.
- (1971) : *Lambert's work in probability*, Ibid., p. 244-256.
- (1971) : *On the history of some statistical laws of distribution*, Biometrika, vol. 58, p. 234-236.
- (1972) : *On the mathematical treatment of observations by Euler*, AHES, vol. 9, p. 45-56.
- (1973) : *Boscovich's work on probability*, AHES, p. 306-324.
- (1973) : *Mathematical treatment of astronomical observations. Historical essay*, AHES, vol. 11, p. 97-126.
- (1974) : *On the prehistory of the theory of probability*, AHES, vol. 12, p. 97-141.
- (1976) : *Laplace's work on probability*, AHES, vol. 16, p. 137-187.
- (1977) : *Laplace's theory of errors*, AHES, vol. 17, p. 1-61.
- (1977) : *Early history of the theory of probability*, AHES, p. 201-259.

- (1978) : *Poisson's work in probability*, AHES, vol. 18, p. 245-300.
- (1979) : *Gauss and the theory of errors*, AHES, vol. 20, p. 21-72.
- (1981) : « Poisson and statistics », in Métivier et al (1981), p. 177-182.
- (1984) : *On the history of the statistical method in astronomy*, AHES, vol. 29, p. 151-199.
- (1984) : *On the history of the statistical method in meteorology*, AHES, vol. 31, p. 53-93.
- (1985) : *On the history of the statistical method in physics*, AHES, vol. 33, p. 351-382. (1986), *Quetelet as a statistician*, AHES, vol. 36, p. 281-325.
- (1989) : *Markov's work on probability*, AHES, vol. 39, p. 337-377 ; vol. 40, p. 387.
- (1991) : *Poincare's work in probability*, AHES, vol. 42, p. 137-172.
- (1991) : *On the works of Buniakovsky in the theory of probability*, AHES, vol. 43, p. 199-223.
- (1993) : *On the history of the principle of least squares*, AHES, vol. 46, p. 39-54.
- (1993) : *Treatment of observations in early astronomy*, AHES, p. 153-192.
- (1994) : *Chebyshev's lectures on the theory of probability*, AHES, p. 321-340.
- (1994) : *Bertrand's work on probability*, AHES, vol. 48, p. 155-199.
- (1998) : *The theory of probability: its definition and its relation to statistics*, AHES, 52, p. 99-108.
- S. Stigler (1982) : *Thomas Bayes's Bayesian inference*, J. R. Stat. Soc., A 145, 250-258.
- (1982) : *Poisson on the Poisson distribution*, Stat. Prob. Lett., 1, 33-35.
- (1983) : *Who discovered Bayes's theorem?*, Am. Stat., 37, 290-296.
- (1984) : *Boscovich, Simpson and a 1760 manuscript note on fitting a linear relation*, Biometrika, 71, 615-620.
- (1986) : *The History of Statistics : The Measurement of Uncertainty before 1900*, Belknap Press of Harvard University Press, Cambridge, MA.

## Références

- (1986) : *Laplace's 1774 memoir on inverse probability*, Stat. Sci., 1, 359-378.
- (1989) : *Francis Galton's account of the invention of correlation*, Star. Sci., 4, 73-86.
- M. Strömer (1731) : *Specimen historiae literariae de arte conjectandi*, Upsaliae, Literis Wernerianis.
- R. Taton (1959) : *Condorcet et Sylvestre-François Lacroix*, Revue d'histoire des sciences, 12, p. 127-158, 243-262.
- I. Todhunter (1865) : *A history of the mathematical theory of probability from the time of Pascal to that of Laplace*, Cambridge.

# Table des matières

|                     |          |
|---------------------|----------|
| <b>Introduction</b> | <b>3</b> |
|---------------------|----------|

## *Chapitre 1*

|                                                         |          |
|---------------------------------------------------------|----------|
| <b>Débuts du calcul des probabilités</b>                | <b>7</b> |
| 1.1. Les probabilités avant 1550.....                   | 7        |
| 1.2. Cardan, <i>Liber de ludo aleae</i> .....           | 9        |
| 1.3. Galilée, <i>Sopra le scoperte dei dadi</i> .....   | 14       |
| 1.4. Biographie de Cardan (1501-1576) .....             | 16       |
| 1.5. Principaux ouvrages de Cardan et références.....   | 19       |
| 1.6. Biographie de Galilée (1564-1642).....             | 20       |
| 1.7. Principaux ouvrages de Galilée et références ..... | 24       |

## *Chapitre 2*

|                                                                                                            |           |
|------------------------------------------------------------------------------------------------------------|-----------|
| <b>La correspondance entre Pascal et Fermat<br/>et le <i>Traité du triangle arithmétique</i> de Pascal</b> | <b>25</b> |
| 2.1. Pascal, Fermat et le problème des partis .....                                                        | 25        |
| 2.2. Biographie de Fermat (1601-1665).....                                                                 | 34        |
| 2.3. La naissance de l'analyse combinatoire<br>et le <i>Traité du triangle arithmétique</i> de Pascal..... | 36        |
| 2.4. Biographie de Pascal (1623-1662) .....                                                                | 46        |
| 2.5. Références.....                                                                                       | 49        |

## *Chapitre 3*

|                                                                              |           |
|------------------------------------------------------------------------------|-----------|
| <b>Huygens</b>                                                               | <b>51</b> |
| 3.1. La genèse du mémoire <i>De ratiociniis in ludo aleae</i> .....          | 51        |
| 3.2. Le mémoire de Huygens .....                                             | 53        |
| 3.3. La notion d'attente ou d'espérance de gain et celle de probabilité..... | 63        |
| 3.4. La vie de Huygens (1629-1695).....                                      | 65        |
| 3.5. Principales publications de Huygens.....                                | 69        |
| 3.6. Références.....                                                         | 70        |



## Chapitre 4

|                                                               |           |
|---------------------------------------------------------------|-----------|
| <b>Montmort</b>                                               | <b>71</b> |
| 4.1. L'Essai d'analyse des jeux de hasards .....              | 71        |
| 4.2. La ruine du joueur .....                                 | 74        |
| 4.3. Le jeu de rencontre et les équations de récurrence ..... | 75        |
| 4.4. La controverse avec de Moivre.....                       | 76        |
| 4.5. La correspondance Montmort-Nicolas Bernoulli .....       | 77        |
| 4.6. Biographie de Montmort (1678-1719) .....                 | 78        |
| 4.7. Publications de Montmort et références.....              | 80        |

## Chapitre 5

|                                                                                       |           |
|---------------------------------------------------------------------------------------|-----------|
| <b>Les Bernoulli</b>                                                                  | <b>81</b> |
| 5.1. Jacob Bernoulli .....                                                            | 81        |
| 5.2. Daniel Bernoulli.....                                                            | 95        |
| 5.3. Nicholas I Bernoulli .....                                                       | 106       |
| 5.4. Biographie de Jacob Bernoulli (1657-1705).....                                   | 111       |
| 5.5. Principales publications de Jacob Bernoulli .....                                | 114       |
| 5.6. Références relatives à la biographie de Jacob Bernoulli .....                    | 114       |
| 5.7. Biographie de Daniel Bernoulli (1700-1782) .....                                 | 115       |
| 5.8. Principales publications de Daniel Bernoulli<br>relatives aux probabilités ..... | 118       |
| 5.9. Références relatives à la biographie de Daniel Bernoulli .....                   | 118       |
| 5.10. Biographie de Nicholas I Bernoulli (1687-1759).....                             | 118       |
| 5.11. Ouvrages de Nicholas Bernoulli relatifs aux probabilités .....                  | 119       |
| 5.12. Références relatives à la biographie de Nicholas Bernoulli.....                 | 119       |

## Chapitre 6

|                                                                                                                  |            |
|------------------------------------------------------------------------------------------------------------------|------------|
| <b>De Moivre</b>                                                                                                 | <b>121</b> |
| 6.1. Le mémoire <i>De Mensura sortis</i> .....                                                                   | 121        |
| 6.2. <i>The Doctrine of Chances</i> et les <i>Miscellanea analytica<br/>de seriebus et quadraturis</i> .....     | 122        |
| 6.3. Rappel concernant la loi des grands nombres .....                                                           | 127        |
| 6.4. Le commentaire de Laplace.....                                                                              | 128        |
| 6.5. Extrait du mémoire de 1733<br><i>Approximatio ad summam terminorum binomii <math>(a + b)^n</math></i> ..... | 129        |
| 6.6. Les causes premières du hasard selon de Moivre .....                                                        | 132        |

|                                                  |     |
|--------------------------------------------------|-----|
| 6.7. Les <i>Annuities upon lives</i> .....       | 135 |
| 6.8. Biographie de De Moivre (1667-1754).....    | 137 |
| 6.9. Principales publications de De Moivre ..... | 140 |
| 6.10. Références.....                            | 141 |

## Chapitre 7

---

|                                                    |            |
|----------------------------------------------------|------------|
| <b>Le principe de probabilité inverse de Bayes</b> | <b>143</b> |
| 7.1. La publication du mémoire de Bayes .....      | 143        |
| 7.2. Extrait du mémoire de 1763 .....              | 146        |
| 7.3. Le commentaire de Laplace.....                | 153        |
| 7.4. Biographie de Bayes (1701-1761).....          | 155        |
| 7.5. Ouvrages de Thomas Bayes .....                | 158        |
| 7.6. Références.....                               | 158        |

## Chapitre 8

---

|                                                                         |            |
|-------------------------------------------------------------------------|------------|
| <b>Buffon</b>                                                           | <b>159</b> |
| 8.1 Le jeu du franc-carreau .....                                       | 159        |
| 8.2. <i>L'Histoire Naturelle</i> .....                                  | 162        |
| 8.3. <i>Des probabilités de la durée de la vie</i> .....                | 163        |
| 8.4. <i>L'Essai d'arithmétique morale</i> .....                         | 163        |
| 8.5. Biographie Georges-Louis Le Clerc, Comte de Buffon (1707-1788).... | 166        |
| 8.6. Principaux ouvrages de Buffon (1707-1788) .....                    | 167        |
| 8.7. Références.....                                                    | 168        |

## Chapitre 9

---

|                                                               |            |
|---------------------------------------------------------------|------------|
| <b>D'Alembert</b>                                             | <b>169</b> |
| 9.1. Doutes et questions sur le calcul des probabilités ..... | 169        |
| 9.2. L'article « Croix-Pile » de l'Encyclopédie .....         | 172        |
| 9.3. L'inoculation .....                                      | 174        |
| 9.4. Le mémoire sur la durée de la vie .....                  | 177        |
| 9.5. Biographie de d'Alembert (1717-1783) .....               | 177        |
| 9.6. Principaux ouvrages scientifiques de d'Alembert .....    | 181        |
| 9.7. Références.....                                          | 181        |

## Chapitre 10

|                                                                       |            |
|-----------------------------------------------------------------------|------------|
| <b>Les contributions de Cotes et Simpson à la théorie des erreurs</b> | <b>183</b> |
| 10.1. Cotes .....                                                     | 183        |
| 10.2. Simpson .....                                                   | 185        |
| 10.3. Biographie de Cotes (1682-1716) .....                           | 188        |
| 10.4. Ouvrages de Cotes et Références .....                           | 188        |
| 10.5. Biographie de Simpson (1710-1761) .....                         | 189        |
| 10.6. Ouvrages de Simpson et références .....                         | 190        |

## Chapitre 11

|                                                                     |            |
|---------------------------------------------------------------------|------------|
| <b>Lambert</b>                                                      | <b>191</b> |
| 11.1. La méthode du maximum de vraisemblance .....                  | 191        |
| 11.2. Les autres travaux de Lambert relatifs aux probabilités ..... | 194        |
| 11.3. L'équipossibilité selon Laplace et Poincaré .....             | 195        |
| 11.4. Biographie de Lambert (1728-1777) .....                       | 197        |
| 11.5. Principaux ouvrages de Lambert .....                          | 199        |
| 11.6. Références .....                                              | 199        |

## Chapitre 12

|                                                                                                                                                                                      |            |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>Laplace</b>                                                                                                                                                                       | <b>201</b> |
| 12.1. Les premiers mémoires 1774-1776 .....                                                                                                                                          | 201        |
| 12.2. Le mémoire <i>Sur les Probabilités</i> .....                                                                                                                                   | 206        |
| 12.4. <i>Mémoire sur les approximations des formules qui sont fonctions des très grands nombres et leur application aux probabilités</i> .....                                       | 213        |
| 12.5. <i>Sur les intégrales définies et leur application aux probabilités et spécialement à la recherche du milieu qu'il faut choisir entre les résultats des observations</i> ..... | 217        |
| 12.6. La <i>Théorie analytique des Probabilités</i> .....                                                                                                                            | 221        |
| 12.7. Biographie de Laplace (1749-1827) .....                                                                                                                                        | 243        |
| 12.8. Les publications de Laplace .....                                                                                                                                              | 247        |
| 12.9. Références .....                                                                                                                                                               | 248        |

## Chapitre 13

|                                                                                                                          |            |
|--------------------------------------------------------------------------------------------------------------------------|------------|
| <b>Condorcet</b>                                                                                                         | <b>249</b> |
| 13.1. Les travaux publiés dans les <i>Mémoires de l'Académie royale des sciences de Paris</i> , entre 1781 et 1784.....  | 249        |
| 13.2. <i>L'Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix</i> ..... | 254        |
| 13.3. Les considérations relatives aux élections .....                                                                   | 256        |
| 13.4. Biographie de Condorcet (1743-1794) .....                                                                          | 260        |
| 13.5. Ouvrages scientifiques de Condorcet.....                                                                           | 265        |
| 13.6. Références.....                                                                                                    | 265        |

## Chapitre 14

|                                                     |            |
|-----------------------------------------------------|------------|
| <b>Lagrange</b>                                     | <b>267</b> |
| 14.1. Le mémoire de Turin sur les probabilités..... | 267        |
| 14.2. Biographie de Lagrange (1736-1813) .....      | 270        |
| 14.3. Principales publications de Lagrange.....     | 274        |
| 14.4. Références.....                               | 275        |

## Chapitre 15

|                                                                                     |            |
|-------------------------------------------------------------------------------------|------------|
| <b>Legendre</b>                                                                     | <b>277</b> |
| 15.1. <i>Nouvelles méthodes pour la détermination des orbites des comètes</i> ..... | 277        |
| 15.2. Application à la mesure des degrés du méridien.....                           | 282        |
| 15.3. Biographie de Legendre (1752-1833) .....                                      | 283        |
| 15.4. Principales publications de Legendre.....                                     | 287        |
| 15.5. Références.....                                                               | 287        |

## Chapitre 16

|                                                                      |            |
|----------------------------------------------------------------------|------------|
| <b>Gauss</b>                                                         | <b>289</b> |
| 16.1. La loi normale et la méthode des moindres carrés de 1809 ..... | 289        |
| 16.2. Les remarques de Joseph Bertrand .....                         | 300        |
| 16.3. La fonction d'erreur .....                                     | 302        |
| 16.4. <i>La théorie de la combinaison des observations</i> .....     | 302        |
| 16.5. Biographie de Gauss (1777-1855) .....                          | 306        |
| 16.6. Principaux ouvrages de Gauss.....                              | 309        |
| 16.7. Références.....                                                | 309        |

## Chapitre 17

|                                                                                          |            |
|------------------------------------------------------------------------------------------|------------|
| <b>Poisson</b>                                                                           | <b>311</b> |
| 17.1. Le <i>Mémoire sur la proportion des naissances des filles et des garçons</i> ..... | 311        |
| 17.2. La publication de 1837.....                                                        | 317        |
| 17.3. Biographie de Poisson (1781-1840).....                                             | 321        |
| 17.4. Ouvrages de Poisson.....                                                           | 325        |
| 17.5. Références.....                                                                    | 326        |

## Chapitre 18

|                                                                      |            |
|----------------------------------------------------------------------|------------|
| <b>Les statistiques d'Ulpie à Quetelet</b>                           | <b>327</b> |
| 18.1. Ulpie.....                                                     | 327        |
| 18.2. Graunt (1620-1674).....                                        | 328        |
| 18.3. La correspondance des frères Huygens.....                      | 331        |
| 18.4. Le calcul politique de Leibniz (1646-1716).....                | 332        |
| 18.5. Arbuthnot (1667-1735).....                                     | 334        |
| 18.6. Süßmilch, le père de la démographie allemande (1707-1767)..... | 338        |
| 18.7. Euler (1707-1783).....                                         | 338        |
| 18.8. Deparcieux (1703-1768).....                                    | 341        |
| 18.9. Lambert (1728-1777).....                                       | 342        |
| 18.10. Quetelet (1796-1874).....                                     | 343        |
| 18.11. Biographie de Graunt (1620-1674).....                         | 348        |
| 18.12. Biographie d'Arbuthnot (1667-1735).....                       | 348        |
| 18.13. Principaux ouvrages d'Arbuthnot et références.....            | 349        |
| 18.14. Biographie de Deparcieux (1703-1768).....                     | 350        |
| 18.15. Principaux ouvrages de Deparcieux et références.....          | 352        |
| 18.16. Biographie de Quetelet (1796-1874).....                       | 352        |
| 18.17. Principaux ouvrages de Quetelet et références.....            | 355        |

## Chapitre 19

|                                                                                       |            |
|---------------------------------------------------------------------------------------|------------|
| <b>Cournot</b>                                                                        | <b>357</b> |
| 19.1. Les écrits de Cournot concernant les fondements du calcul des probabilités..... | 357        |
| 19.2. Le principe de Cournot.....                                                     | 359        |
| 19.3. Biographie de Cournot (1801-1877).....                                          | 361        |
| 19.4. Principales publications de Cournot et références.....                          | 363        |

## Chapitre 20

|                                                                                                                                           |            |
|-------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>Bienaymé</b>                                                                                                                           | <b>365</b> |
| 20.1. <i>Mémoire sur la probabilité des erreurs<br/>d'après la méthode des moindres carrés</i> .....                                      | 365        |
| 20.2. Les travaux antérieurs de Bravais .....                                                                                             | 374        |
| 20.3. <i>Considérations à l'appui de la découverte de Laplace<br/>sur la loi de probabilité dans la méthode des moindres carrés</i> ..... | 375        |
| 20.4. La controverse Bienaymé-Cauchy .....                                                                                                | 377        |
| 20.5. L'Inégalité de Bienaymé-Chebyshev .....                                                                                             | 378        |
| 20.6. Biographie de Bienaymé (1796-1878) .....                                                                                            | 379        |
| 20.7. Principaux mémoires de Bienaymé .....                                                                                               | 380        |
| 20.8. Références.....                                                                                                                     | 381        |

## Chapitre 21

|                                                                                |            |
|--------------------------------------------------------------------------------|------------|
| <b>Boole</b>                                                                   | <b>383</b> |
| 21.1. Les diverses publications de Boole .....                                 | 383        |
| 21.2. Le <i>challenge Problem</i> de Boole et la controverse avec Cayley ..... | 387        |
| 21.3. Biographie.....                                                          | 387        |
| 21.4. Publications de Boole relatives aux probabilités .....                   | 389        |
| 21.5. Références.....                                                          | 389        |

## Chapitre 22

|                                                                         |            |
|-------------------------------------------------------------------------|------------|
| <b>Chebyshev</b>                                                        | <b>391</b> |
| 22.1. <i>Mémoire sur Des valeurs moyennes</i> .....                     | 391        |
| 22.2. <i>Mémoire Sur deux théorèmes relatifs aux probabilités</i> ..... | 393        |
| 22.3. Biographie de Chebyshev (1821-1894) .....                         | 394        |
| 22.4. Publications de Chebyshev .....                                   | 396        |
| 22.5. Références.....                                                   | 397        |

## Chapitre 23

|                                                        |            |
|--------------------------------------------------------|------------|
| <b>Bertrand</b>                                        | <b>399</b> |
| 23.1. Le cours de <i>Calcul des probabilités</i> ..... | 399        |
| 23.2. Biographie de Bertrand (1822-1900) .....         | 404        |
| 23.3. Principaux ouvrages de Bertrand .....            | 405        |
| 23.4. Références.....                                  | 405        |

*Chapitre 24*

---

|                                                       |            |
|-------------------------------------------------------|------------|
| <b>Poincaré</b>                                       | <b>407</b> |
| 24.1. Contenu du <i>Calcul des Probabilités</i> ..... | 407        |
| 24.2. Biographie de Poincaré (1854-1912).....         | 416        |
| 24.3. Références.....                                 | 422        |

*Chapitre 25*

---

|                                                                                     |            |
|-------------------------------------------------------------------------------------|------------|
| <b>Épilogue : Kolmogorov et l'axiomatisation<br/>de la théorie des probabilités</b> | <b>423</b> |
|-------------------------------------------------------------------------------------|------------|

*Annexe*

---

|                                                                                                                                                     |            |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>Chronologie des événements constituant<br/>le contexte historique, politique et scientifique<br/>de l'élaboration du calcul des probabilités</b> | <b>427</b> |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|------------|

*Références*

---

|                                                                                                                 |            |
|-----------------------------------------------------------------------------------------------------------------|------------|
| <b>Quelques publications ou livres relatifs à l'Histoire<br/>de la période pré-axiomatique des Probabilités</b> | <b>449</b> |
|-----------------------------------------------------------------------------------------------------------------|------------|

