Journal de la société statistique de Paris

Pierre Crépel

Condorcet et l'estimation statistique

Journal de la société statistique de Paris, tome 129, n° 1-2 (1988), p. 46-67
<http://www.numdam.org/item?id=JSFS_1988__129_1-2_46_0>

CONDORCET ET L'ESTIMATION STATISTIQUE

Pierre Crépel
Université de Rennes I et REHSEIS *(CNRS)*

L'article suit, depuis les premiers manuscrits de 1772 jusqu'aux derniers textes publiés, le cheminement de la pensée de Condorcet pour traiter l'estimation statistique et en discuter les champs de validité. On montre en particulier comment, contrairement à ce qu'on dit habituellement, Condorcet y a fait preuve d'imagination, de rigueur et de lucidité.

This paper deals with the evolution of Condorcet's approach to statistical estimation, from the beginnings (1772) to his last published articles and books. Against what is generally asserted, we insist on his ingenuous, rigorous and lucid point of view.

Introduction

Traiter en détail de l'apport de Condorcet aux probabilités et statistiques est une tâche de longue haleine actuellement en chantier. Nous nous concentrerons ici sur un sujet beaucoup plus limité, mais crucial, celui de ce que nous appelons aujourd'hui « l'estimation ». L'exemple-type suivant simplifié permet déjà d'en examiner la richesse :

(*) Deux événements contradictoires A et B peuvent arriver, la probabilité x de A est inconnue. On sait qu'au cours de $m+n$ épreuves A est arrivé m fois et B n fois.

1° Quelle est la probabilité d'arrivée de A lors d'une nouvelle épreuve?

2° Dans quelle mesure peut-on avoir confiance dans une telle réponse? Y a-t-il de meilleures façons de poser la question? Quelles leçons peut-on en tirer?

Le cas où m n'est pas très grand et où $n = 0$ constitue un thème de débat particulièrement aigu.

Notre article présente un caractère assez préliminaire et technique : dans quels écrits et par quelles formules précises Condorcet répond-il à cette question? et où (pour quels problèmes?) utilise-t-il ses réponses?

Le sujet plus vaste du contexte mathématique et philosophique de l'inférence statistique à la fin du 18e siècle, des réponses ultérieures de grands statisticiens comme R.A. Fisher, J. Neyman..., exigerait un autre article [1]. Nous nous contenterons des rappels nécessaires pour permettre au lecteur non historien des sciences de suivre plus facilement l'exposé.

Une citation bien connue et spécialement tranchante de Condorcet clame ce que beaucoup de philosophes et de mathématiciens pensent plus ou moins clairement pendant le troisième quart du 18e siècle :

Dans son mémoire sur la probabilité des causes par les événements, Laplace *« traite une branche de l'analyse des hasards, bien plus importante et moins connue [...]; ici la probabilité est inconnue, c'est-à-dire que le nombre des chances pour ou contre un événement proposé, est indéterminé; on sait seulement que dans un nombre donné d'expériences, cet événement est arrivé un certain nombre de fois, et on demande comment de cette seule donnée on peut conclure la probabilité de ce qui doit arriver dans la suite. On voit que cette question renferme toutes les applications de la doctrine des hasards aux usages* *de la vie, et c'est la partie de cette science la seule utile, la seule digne d'occuper sérieusement des Philosophes; le calcul ordinaire ne sert qu'à donner les probabilités des jeux de hasards et des loteries, et il n'a pas même l'utilité de dégoûter de ces amusements également funestes à l'industrie et aux mœurs ».* [1774]

Cet aspect est donc indissolublement lié à la théorie de la connaissance, au problème de l'induction, à l'empirisme, au rationalisme. Dans le langage plus technique des probabilités, il concerne de près l'inversion du théorème de Jacques Bernoulli, les liens entre probabilités *a priori* et *a posteriori*, la probabilité des causes, les principes de raison insuffisante et de vraisemblance, et même à certains égards les arrangements réguliers...

Revenons à notre problème (*) pour distinguer deux cas :

1° Lorsque les conditions mêmes de l'expérience définissent sans ambiguïté la probabilité de A, comme par exemple au jeu de croix ou pile avec une pièce symétrique, il n'y a pas à tenir compte des résultats des épreuves précédentes pour évaluer la probabilité de A.

2° Par contre, si l'expérience résume tout ou partie de nos informations, d'autres réponses sont à envisager, et d'abord les deux suivantes :

La première consiste à juger les données insuffisantes et le problème indéterminé.

La seconde consiste à prendre pour probabilité de A dans une épreuve ultérieure la fréquence d'apparition dans les épreuves passées, à savoir $m/(m+n)$, c'est ce qu'on fait par exemple, sans le dire, en identifiant les probabilités de survie aux chiffres issus des tables de mortalité. Le théorème de Jacques Bernoulli justifie partiellement cette solution en montrant que lorsque le nombre d'épreuves devient très grand, cette fréquence tend vers la probabilité de A.

Reprécisons bien ceci : la distinction entre ces deux *« sources de probabilité »*, le cas où elles sont *« tirées de la considération de la nature même »* et celui où elles ne sont *« fondée[s] que sur l'expérience du passé »*, pour reprendre les termes de l'article « Probabilité » de l'Encyclopédie (souvent attribué à Diderot), est explicite chez J. Bernoulli, et l'objectif de son théorème concerne essentiellement ce second cas :

« on doit présumer que par la suite chaque fait peut arriver et ne pas arriver dans le même nombre de cas qu'il avait été constaté auparavant, dans un état de choses semblables, qu'il arrivait ou n'arrivait pas ».

« [...] tout être des plus stupides, par je ne sais quel instinct naturel, par lui-même et sans le guide d'aucun enseignement (chose absolument admirable) tient pour évident que, plus on aura recueilli de nombreuses observations de ce genre, moins grand sera le danger de s'écarter du but ».

J. Bernoulli ajoute qu'il va démontrer non seulement ce résultat, mais surtout celui que *« la probabilité d'atteindre le rapport réel entre les nombres de cas [...] dépasse enfin un degré quelconque donné de certitude »*, c'est-à-dire ne reste pas bornée supérieurement par un nombre strictement inférieur à 1.

Et, pour faire comprendre son propos, il indique l'exemple suivant d'une urne renfermant un nombre inconnu de pierres blanches et de pierres noires. *« [...] je suppose que pour connaître leur nombre par expérience tu tires une pierre après l'autre (en replaçant cependant chaque fois la pierre que tu as tirée avant de choisir la suivante, pour que le nombre des pierres ne diminue pas dans l'urne); tu observes combien de fois sort une pierre blanche et combien de fois une noire. On demande si tu peux le faire tant de fois qu'il devienne dix fois, cent fois, mille fois etc. plus probable (c'est-à-dire qu'il devienne moralement certain) que le nombre de fois où tu choisis une pierre noire soit dans ce même rapport [...] où se complaisent à être entre eux les nombres des pierres ou des cas, plutôt que dans tout autre rapport différent de celui-ci. »* [J. Bernoulli, 1713-1987]

On trouve donc déjà, d'une certaine façon, dans J. Bernoulli une réponse précise au problème (*) et une indication rigoureuse sur la confiance qu'on peut avoir en cette réponse, ainsi d'ailleurs que diverses discussions fines évoquant même la possibilité d'une variation de la composition de l'urne.

Ce schéma d'urnes de composition soit connue, soit inconnue, devient, après J. Bernoulli, un mode d'expression des problèmes de la théorie des chances.

Dans les années 1760-1770, l'insuffisance théorique des réponses de J. Bernoulli apparaît cependant, d'autant plus nettement que d'Alembert émet des « doutes » sur le fondement du calcul des probabilités : ces doutes serviront pour une large part de point de départ aux réflexions de Condorcet et de Laplace, tous deux disciples de d'Alembert.

Certes, T. Bayes a apporté des réponses nouvelles et argumentées à l'inversion du théorème de Bernoulli, publiées en 1764-1765, mais ni d'Alembert, ni Condorcet, ni Laplace n'en ont connaissance avant 1780 environ, alors que dès 1772 leurs travaux rejoignent par de nombreux points celui de Bayes.

Rappelons que la solution dite de Bayes-Laplace au problème (*), telle qu'elle est exprimée dans le début du mémoire de Laplace de 1774 sur la probabilité des causes par les événements, est la suivante :

$$E = [\int_0^1 x^{m+1}(1-x)^n \, dx] / [\int_0^1 x^m (1-x)^n \, dx] = (m+1)/(m+n+2)$$

(et non $m/(m+n)$, obtenu en identifiant cette probabilité à la fréquence des événements observés). La proportion inconnue x de boules blanches dans l'urne, pouvant avoir *a priori* toutes les valeurs possibles indifféremment entre 0 et 1, Laplace prend pour résultat le rapport de la probabilité moyenne d'obtenir $m+1$ blanches et n noires à la probabilité moyenne d'obtenir m blanches et n noires.

Ce mémoire contient d'ailleurs bien d'autres développements remarquables autour de ce problème, par exemple sur l'écart entre la probabilité réelle de A et la fréquence observée. Pour toutes précisions et interprétations sur cette formule de Bayes-Laplace, on se reportera particulièrement à [P.S. Laplace, 1814-1986] et [S. Stigler, 1986].

1. *Quelques lectures de Condorcet aux 19e et 20e siècles*

L'importance d'un nouvel inventaire précis des travaux probabilistes de Condorcet se justifie pour plusieurs raisons.

D'abord, de nombreux manuscrits sont encore inédits, certains viennent d'être retrouvés, d'autres sont dans un désordre inexprimable : L. Cahen, qui les a classés au début du siècle, les a recueillis dans un état lamentable et, n'ayant aucune compétence en mathématiques (comme il le reconnaît lui-même), il s'est contenté de mettre de l'ordre dans les papiers historiques et politiques [2]. Depuis lors, plusieurs chercheurs, notamment G.G. Granger, R. Rashed et K.M. Baker, ont consulté certains des manuscrits probabilistes, mais ils ne les ont pas étudiés de manière complète.

Ensuite, les articles et ouvrages de Condorcet publiés n'ont fait l'objet, du point de vue mathématique, que de deux examens systématiques : celui de I. Todhunter (vers 1865) et celui de K. Pearson (vers 1930). Chacun des deux a clarifié divers aspects, et il faut reconnaître que la tâche n'est pas aisée. Ce n'est pas sous-estimer leur travail que d'en constater les lacunes, car l'acquis de leurs études facilite grandement notre tâche aujourd'hui.

L'essentiel des publications probabilistes de Condorcet est constitué par
— ses articles de l'Encyclopédie Méthodique,
— un traité intitulé « Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix » (l'Essai),
— un « Mémoire sur le calcul des probabilités » en six parties (le Mémoire),
— un ouvrage posthume « Élémens du calcul des probabilités » (les Élémens).

I. Todhunter n'étudie que l'Essai et le Mémoire, il semble ne pas connaître les Elémens et ne consacre que quelques lignes aux articles de l'Encyclopédie Méthodique. En outre beaucoup de points lui échappent, ou lui paraissent arbitraires, obscurs ou inintelligibles : il faut reconnaître qu'il s'agit souvent de questions tout à fait étrangères à l'esprit du 19e siècle et même de la première moitié du 20e. Pour le problème de l'estimation, il analyse correctement (p. 377-385) les passages de l'Essai qui y sont relatifs; par contre, certes sans faire d'erreur mathématique, il reste vague et on ne peut plus distant en ce qui concerne le Mémoire.

K. Pearson a une vision beaucoup plus complète et mesurée des travaux publiés de Condorcet, en particulier concernant l'estimation, mais malheureusement certains passages de son manuscrit (que nous n'avons pas encore pu consulter) ont été supprimés par l'éditeur, son fils E.S. Pearson; en outre divers aspects modernes n'ont pas retenu son attention.

Enfin, et surtout, sur la foi de Todhunter, mais aussi de presque tous les auteurs du 19e siècle, y compris des biographes admiratifs de Condorcet, l'intérêt de l'œuvre probabiliste de cet Encyclopédiste est traditionnellement réduit à son apport philosophique; plus souvent encore elle est considérée au mieux comme dépassée, au pire comme extravagante, quand ce n'est pas ignorée purement et simplement. Les meilleurs livres et articles actuels d'histoire des probabilités et statistiques ne lui consacrent que quelques lignes ou pages de façon anecdotique. Même les auteurs qui ont réhabilité Condorcet depuis 1950 n'ont guère osé contredire Todhunter sur le plan des mathématiques.

2. *Présentation des articles et manuscrits*

On peut distinguer en gros trois périodes quant aux interventions probabilistes de Condorcet [P. Crépel, 1988].

La première recouvre environ les années 1772-1780 : la pensée de l'auteur est en formation, mais s'affirme rapidement, il commence plusieurs mémoires qu'il ne publie pas (en est-il mécontent? a-t-il par ailleurs un surcroît de travail? a-t-il peur de froisser d'Alembert? souffre-t-il de la concurrence avec Laplace? ... plusieurs de ces causes s'entremêlent-elles? nous ne le savons pas précisément). Une grande partie de ces textes subsiste, nous en examinerons plus loin les aspects relatifs à l'estimation. Au cours de cette période, les seuls passages imprimés de Condorcet relatifs aux probabilités sont très courts et proviennent soit d'Eloges (Huyghens, Pascal), soit de présentations d'articles de Laplace dans la partie « Histoire » des recueils de mémoires de l'Académie des sciences. Un appendice du livre « Nouvelles expériences sur la résistance des fluides » [1777] constitue le seul écrit publié d'une certaine ampleur où le calcul des probabilités joue un rôle, d'ailleurs non exclusif.

La seconde période, que nous pourrions appeler de maturité, est celle où Condorcet fait de l'arithmétique politique et de la mathématique sociale sa principale activité, soit environ 1781-1787 : alors sont publiés le Mémoire, l'Essai, les articles de l'Encyclopédie Méthodique.

La troisième est le prolongement de la précédente, sa conception n'est pas modifiée, mais l'Encyclopédiste devient de plus en plus accaparé par d'autres activités et, mise à part la nouvelle synthèse, en partie pédagogique, inachevée des Elémens, on ne dispose, comme travaux probabilistes, que de fragments, publiés ou non.

3. *Les premiers manuscrits*

On trouvera en annexe la liste de ces manuscrits, dont deux débuts d'essais sur les probabilités tentés l'un en 1772, l'autre vers 1774, et d'autres passages qui semblent s'y rattacher; aucun de ces textes n'a été publié, aucun n'est daté de la main de Condorcet. K.M. Baker a proposé des dates raisonnables pour deux d'entre eux, par comparaison de leur contenu avec la correspondance Condorcet-Turgot; les lignes suivantes nous paraissent confirmer sa position et apporter des précisions nouvelles.

Chacun de ces deux mémoires inédits, présumés de 1772 et 1774, commence par une partie historique où Condorcet prend pour base l'ouvrage de J. Bernoulli, et cherche à répondre de façon constructive aux doutes de d'Alembert.

Étudions-les par rapport au problème (*).

a) *Le manuscrit de 1772 :*

Visiblement ce premier manuscrit ignore jusqu'à l'existence du mémoire de Bayes. Il part du résultat de J. Bernoulli et ajoute à son propos :

« mais il n'en[tre] dans aucun détail sur la certitude qu'on peut en avoir se contentant d'observer que plus on a tiré [...] de boules plus la probabilité est grande. Je vais commencer par donner le calcul de cette probabilité. »

Ainsi Condorcet n'y conteste-t-il absolument pas l'identification de la probabilité cherchée à la fréquence d'observation des événements passés, il réclame simplement une évaluation de la probabilité de l'erreur faite en procédant à cette identification : en d'autres termes, il accepte la réponse classique au 1° et s'attaque à la question 2°.

Condorcet travaille dans le cas $m=n$, pour simplifier, mais note plus loin :

« on trouvera de même tout autre rapport [c.à.d. m et n quelconques] et si on n'avait tiré que des boules blanches [c.à.d. $n=0$] il en serait de même. »

Il suppose en outre le nombre total de boules fini et égal à $2N$, ce dont il « s'excusera » plus loin sous la forme suivante :

« Mais il faut observer que le nombre [sous-entendu total] des boules entre dans l'évaluation exacte de la probabilité, il faut même du moins qu'on en connaisse à peu près les limites. »

Voici alors son raisonnement :

« si j'ai N blanches et N noires [dans l'urne] le nombre des combinaisons pour amener m [blanches] et m [noires] est

$$C_{2m}^{m}\, N^{2m}.$$

Le nombre des combinaisons qui donnent le même résultat si l'on a $N+1$ noires et $N-1$ blanches et réciproquement [c'est-à-dire : $N-1$ noires et $N+1$ blanches] est

$$C_{2m}^{m}\, (N+1)^m (N-1)^m.$$

La probabilité qu'il y a dans le sac autant de blanches que de noires est donc à celle qu'il y en a une de plus comme N^{2m} à $2\,(N+1)^m(N-1)^m$ [...] »

Plus généralement la probabilité qu'il y a dans le sac autant de blanches que de noires est à celle qu'il y en a un nombre différent « comme N^{2m} au double de la somme [...] $(N^2-1)^m+(N^2-4)^m+(N^2-9)^m$ [+...] ».

De la même façon, la probabilité que le nombre de boules blanches soit compris entre $N-p$ et $N+p$, où p est donné est à celle du contraire comme

$$N^{2m}+2\,[(N^2-1)^m+(N^2-4)^m+\ldots+(N^2-p^2)^m]$$

au reste de la somme.

Il s'agit donc, au moins implicitement, d'une application du principe de « vraisemblance » suivant, que Laplace énoncera de façon précise :

On envisage les différentes hypothèses possibles qui ont pu produire l'événement et, sous chacune de ces hypothèses, on calcule la probabilité d'arrivée de cet événement; on estime alors que les probabilités de ces différentes hypothèses sachant que l'événement s'est produit sont proportionnelles à ces nombres.

Supposant alors N et p donnés, se fixant une probabilité r, nous dirions « de confiance », il conclut :

« ainsi l'on saura combien il faut d'expériences [m] pour conclure à tant d'erreur près [p] avec une probabilité donnée [r] le rapport d'égalité entre le nombre des boules ».

Il commente ensuite ce résultat :

« La manière de MM. Bernoulli et Moivre est différente de la mienne et conduit au même résultat général mais les résultats particuliers sont différents.

Ces principes posés on veut que pour pouvoir avoir une grande probabilité et regarder en conséquence comme presque exact le rapport trouvé entre un certain nombre d'effets pour le vrai rapport des causes il faut un très grand nombre d'expériences et que par conséquent pour dresser les tables de mortalité etc. il faut établir les raports sur un très grand nombre d'homes.

2° Le degré de probabilité et l'exactitude du résultat dépendent du nombre des expériences, des rapports et de telle probabilité ainsi dans chaque résultat il faudrait soigneusement marquer l'un et l'autre.

3° Lorsque les calculs sont pour un nombre de choses supérieur à celui des expériences, les résultats sont exacts, mais lorsqu'il est de les appliquer à une seule on rentre dans les difficultés du jeu inégal ainsi par exemple s'il est très probable qu'il ne meurt qu'un homme sur mille de l'inoculation l'état où il a mille inoculés est presque sûr d'en conserver 999, mais chaque inoculé peut ou vivre ou mourir [...]. »

Le contenu de ce passage montre aussi que Condorcet ignore à ce moment le mémoire de Laplace sur la probabilité des causes par les événements.

b) *Le manuscrit dit de 1774 :*

Ce manuscrit comprend deux parties : l'une, historique, intitulée « Histoire abrégée de ce calcul », l'autre de caractère plus philosophique, appelée « De la nature du calcul des probabilités ». Il est clairement rédigé en plusieurs fois.

Le premier jet, légèrement postérieur au mémoire de Laplace sur la probabilité des causes par les événements, ne fait qu'évoquer le problème précis que nous traitons ici; nous le transcrivons tel quel avec les ratures, d'après le manuscrit autographe [m.a] [3] :

« [m.a. : ff 85/89] En considérant les probabilités des <> événemens de la vie, J. Bernoulli apperçut une différence esentielle entr eux et les coups des jeux de hazards. Dans ceux-ci le nombre total des événemens et celui des événemens favorables sont connus <*> rigoureusement. Il n'en est pas de même dans les événemens de la vie : <*> on ne peut juger du futur que par le passé de la cause que par les effets; sur dix mille personnes, par exemple, il en est <*> mort dix d'une maladie, < de la petite vérole> dans un espace de tems doné. »*

Le m.a. comprend alors la partie barrée suivante :

« <Combien est-il probable que sur dix mille autres persones placées dans des circonstances semblables il en mourra aussi dix dans le même tens tel est la 1re question qui se présente dans cette recherche elle paraît d'abord beaucoup plus difficile que les questions que l'analyse des jeux de hazard avait fait naître [rajouté (semble-t-il, avec la même encre que « place du folio A ») :] mais on peut la résoudre avec les mêmes principes.

Plusieurs savans Géomètres de ce siècle se sont occupés de ces questions qui exigent un usage adroit de la science des combinaisons, du calcul intégral ordinaire, de celui des différences finies <souvent> quelquefois même des deux réunis. M. de la Place <a doné> vient de doner [je pense contrairement à Baker que Condorcet a d'abord écrit « a doné » puis corrigé en « vient de doner » (?)] sur cet objet un <*> Mémoire très profond inséré dans le volume 6e [on pourrait lire 5e, mais je ne pense pas] de la Collection des Savans étrangers.> »

Le m.a. est alors corrigé ultérieurement, la partie transcrite ci-dessus y est remplacée par le « folio A » suivant :

« [m.a. : (place du folio A), f 86 ...]

A. Quelle probabilité y a-t-il qu'un homme attaqué de cette maladie n'y succombera point. <*> Dans toutes les questions <*> où on n'examine que la probabilité sans faire entrer dans le calcul le plus ou moins d'importance des événemens <*> la probabilité se peut réduire à celle de tirer une boule blanche dans un nombre de boules blanches et noires <*>. Il suffit de supposer le nombre total des boules égal au nombre des événemens ou des combinaisons et le nombre des boules blanches égal au nombre des cas où arrive l'événement dont on cherche la probabilité.

Pour réduire <*> la probabilité à cette expression dans <*> la question qu'on vient de donner pour exemple <*>, on observera qu'elle se réduit à celle-ci ayant un nombre <infini> inconnu de boules blanches et noires et ayant tiré 10 000 [?] mille dont 10 seulement sont noires trouver la probabilité d'en tirer encore une blanche. <<*> Si le nombre <connu [?]> est in [?] fini, on peut <ici> supposer ici ce nombre de boules [?] ou fini ou infini [?] come [?] [très difficile à lire car barré d'un trait plein]>

Dans la question précédente on doit regarder le nombre des boules comme <*> infini puisqu'il n'y a pas de nombre si grand qu'il soit au-delà duquel celui des événemens puisse s'étendre encore ainsi on peut sans inconvénient regarder le nombre total comme l'unité, et alors le nombre aussi inconnu des boules blanches pourra être exprimé par <*> un nombre <inconu mais> fractionnaire quelconque entre zéro et l'unité. L'on voit que ce nombre fractionnaire n'est alors que le rapport des boules blanches au nombre total.

On prendra donc une expression algébrique de la probabilité <*> d'amener <*> le nombre donné de boules blanches et noires en supposant <*> <*> le nombre <*> fractionnaire inconu.

[m.a. : ff 86/87, feuille notée A2 par Condorcet] Ce nombre pouvant être depuis 0 jusqu'à l'unité, il est clair que si l'on a la formule général, et qu'on la divise par le nombre des valeurs possibles représenté ici par l'unité on aura la probabilité d'amener en général le nombre donné. On cherchera ensuite par la même méthode la probabilité d'amener en dix mille un coup 10 boules noires seulement divisant ces deux quantités, l'une par l'autre on aura la probabilité d'amener une boule blanche au dix mille et unième coup, lorsque dans les dix mille premiers on en a déjà amené 9990. Si dans la même hypothèse on cherchait combien il y a à parier que sur 1 000 malades il n'en mourra <*> qu'un ou que sur cent malades il n'en périra aucun, on chercherait dans le 1er cas la probabilité pour amener 11 boules noires seulement sur onze mille et on le diviserait par celle d'amener dix sur dix mille et dans le second on prendrait la probabilité d'amener dix boules noires sur 10 100, [rajouté par Condorcet sur la copie; sur le m.a. il y a écrit 10 000] et on la diviserait par celle de les amener sur cent mille [?]

<[écrit avec l'encre que du folio A] On trouvera le détail de tous ces calculs dans le Mémoire de la Place imprimé dans le tome VI>

Il existe également une copie de ce mémoire effectuée par un copiste à l'intention duquel Condorcet a ajouté, comme correction, avec une encre fine sur le m.a. :

« <*> Les principes de ces calculs se trouvent dans les Transactions philosophiques année 1764 No LIII <par> Dans différens morceaux de Mrs Bayes et Price. M. Hébert laissera ici cinq à six lignes en blanc. » [ce que le copiste a fait, et ce qui confirme au passage que Condorcet ne connaissait pas auparavant le mémoire de Bayes]

Plus loin, dans la partie philosophique, une nouvelle allusion au mémoire de Laplace est ajoutée dans un folio B (f 96), mais l'objectif de Condorcet à cet endroit concerne le lien entre le calcul des probabilités et la théorie de la connaissance et ces considérations ne comportent pratiquement pas de calculs, sinon le suivant, appartenant déjà au premier jet :

« Par exemple, je suppose que j'aie <> seulement 100 expériences bien constatées d'un fait phisique et que je sache que sur <*> un million de fois <*> que l'on a ainsi 100 expériences uniformes la centième la été aussi. Je prendrai la probabilité pour que ayant tiré cent boules blanches j'en tire une cent unième <*>. La probabilité <*> qu'ayant tiré dix millions de boules noires, j'en tire une blanche. Enfin la probabilité qu'ayant tiré dix millions de boules blanches il en vienne une noire, je multiplierai les deux probabilités <*> l'une par l'autre, je les diviserai par la dernière et j'aurai la probabilité que ma cent unième expérience réussira. »*

La lecture de ce texte avec ses ratures, son orthographe et ses hésitations paraîtra sans doute un peu pénible, mais elle nous permet de constater deux aspects frappants. D'abord, dans le premier jet, Condorcet a bien noté la profondeur du mémoire de Laplace, mais il ne l'utilise pas explicitement dans ses propres considérations où la notion de probabilité moyenne, en particulier, semble absente. Ensuite, dans les corrections ultérieures, il semble se rallier à la méthode de Bayes-Laplace sans discussion ni critique : il l'utilise à des fins philosophiques, mais ne cherche pas à la relativiser ni à l'approfondir comme ce sera le cas plus tard au milieu des années 80.

Pour l'étude du lien avec la théorie de la connaissance, nous renvoyons à [B. Bru, 1988 b].

c) *Le manuscrit MS 875 ff 116-125, 115*

Il subsiste dans les volumes de l'Institut un autre manuscrit inédit important dans lequel sont développés des calculs destinés à répondre au problème (*). Voici le contenu de ces calculs :

Ce manuscrit prend visiblement appui sur celui de 1772 :

« Quant aux événements de la vie, s'il n'y a que deux manières possibles, il faut trouver la possibilité supposée (cela est détaillé dans l'ancien mémoire) mais d'où vient cette probabilité d'amener une boule blanche au premier coup? [...] ».

Après quelques réflexions, Condorcet passe aux calculs en étudiant d'abord *« la situation de celui qui après avoir vu [un] événement arriver constamment un nombre m de fois attend que cet événement arrive ou le contraire » :* il s'agit donc du problème (*) avec $n=0$.

« Je suppose que le nombre de fois que l'événement peut arriver ou ne pas arriver soit infini. Alors appelant m le nombre de fois qu'il est arrivé il est clair de voir que les probabilités pour que l'événement arrive toujours, ou ne manque qu'une fois, ou manque deux etc. suivent entre elles les rapports de ces nombres ∞^m, $(\infty-1)^m$, $(\infty-2)^m$... *Maintenant si on tire un coup de plus, il est clair que le nombre de toutes les combinaisons possibles auxquelles il faut rapporter la probabilité que l'événement arrivera ou non sera* $\infty\,(\infty^m+(\infty-1)^m...) = \infty^{m+2}/(m+1)$; *maintenant l'espérance d'amener le même événement est*

$\infty^{m+1}+(\infty-1)^{m+1}... = \infty^{(m+2}/(m+2)$

donc la probabilité que l'événement arrivera est $(m+1)/(m+2)$ *c'est-à-dire la même que j'aurais pour amener une boule noire. » [?]*

Donnons quelques explications pour montrer le lien avec le calcul de 1772 : notons N au lieu de ∞ le nombre de boules dans l'urne. Condorcet examine les différentes hypothèses possibles concernant la composition de l'urne :

H_N : N blanches et 0 noire,
H_{N-1} : $N-1$ blanches et 1 noire,
...
H_k : k blanches et $N-k$ noires
...
H_0 : 0 blanche et N noires;

et il applique le même principe de vraisemblance que plus haut, en affirmant que les probabilités de ces diverses hypothèses, sachant que A est arrivé m fois et B 0 fois (c'est-à-dire qu'on a tiré m fois une boule blanche, avec replacement, et 0 fois une boule noire), sont proportionnelles aux différentes probabilités de cet ensemble de tirages sous chacune des hypothèses, c'est-à-dire aux nombres :

N^m,
$(N-1)^m$,
...
k^m,
...
0.

On remarquera en outre qu'il raisonne en termes de combinaisons, et non de probabilités.

La probabilité de tirer une boule blanche au $(m+1)^e$ coup y apparaît alors comme le quotient du nombre des combinaisons qui amènent cet événement, soit

$$N^{m+1}+(N-1)^{m+1}+\dots$$

au nombre total des combinaisons, soit

$$N\,[N^m+(N-1)^m+\dots]$$

et comme $N = \infty$ cela revient au quotient

$$\int x^{m+1}dx \,/\, \int x^m dx.$$

L'auteur ne dit pas explicitement qu'il considère toutes les hypothèses possibles comme *a priori* équiprobables, mais son dénombrement revient finalement à le supposer.

Condorcet passe alors au cas $n \neq 0$, mais son brouillon en partie raturé semble inachevé. Si nous avons bien compris, on pourrait compléter son calcul de la manière suivante :

Puisqu'on sait qu'on a tiré m blanches et n noires, les probabilités respectives des différentes hypothèses H_N, H_{N-1}, ..., H_k, ..., H_0 sont proportionnelles aux probabilités de réalisation de l'événement arrivé sous chacune de ces diverses hypothèses, soit

$N^m.0^n$,
$(N-1)^m.1^n$,
...
$k^m.(N-k)^n$,
...
$0^m.N^n$;

par suite la probabilité d'arrivée de A, lors du coup qui suit les $m+n$ épreuves déjà réalisées, sera le quotient de

$$N^{m+1}0^n+(N-1)^{m+1}1^n+\dots+k^{m+1}(N-k)^n+\dots+0^{m+1}N^n$$

par

$$N\,[N^m0^n+(N-1)^{m+1}1^n+\dots+k^m(N-k)^n+\dots+0^mN^n].$$

Il est alors clair que si l'on fait $N = \infty$, cela revient à la formule de Bayes-Laplace

$$[\int x^{m+1}(1-x)^n dx] \,/\, [\int x^m(1-x)^n dx] \;=\; (m+1)\,/\,(m+n+2).$$

Ce calcul n'étant pas plus difficile que celui effectué dans le cas $n = 0$, il est étonnant que Condorcet ne parvienne pas explicitement à cette formule.

Revenant apparemment au cas $n = 0$, Condorcet répond ensuite d'avance à quelques objections qu'on pourrait lui faire :

« Si ces calculs où j'ai supposé le nombre des cas infini choquaient quelques géomètres, il serait aisé de l'établir en supposant le nombre des coups fini et qu'à chaque fois qu'on tire une boule on la rejette

*[?] dans le < * > vase où elles sont, et si cette supposition paraît [?] encore précaire [?] supposons le fini et qu'on ne remette pas nous aurons si le nombre [total de boules] est N*

$$N\,(N-1)(N-2)\,\ldots\,(N-m+1),$$
$$(N-1)(N-2)(N-3)\,\ldots\,(N-m),$$
$$\ldots$$

pour les cas de probabilité [des hypothèses respectives H_N, H_{N-1}, ..., H_0] [...] ». Renvoyant à certains passages d'Euler, il donne alors quelques indications pour poursuivre ce calcul inachevé qui évoque plus qu'implicitement ce que nous appelons la répartition hypergéométrique, puis il note :

« On voit ici que si on suppose m fini et N infini on a le même résultat que ci-dessus.

J'ai pris ces deux cas parce qu'il me paraît que celui de N fini ne peut représenter le cas de la question qu'autant que les événemens dont il est question ne pouvait [?] s'étendre au-delà d'un certain nombre de cas. Or c'est ce qu'on ne doit pas supposer en général, je crois même qu'il serait nécessaire de supposer le contraire dans une infinité de cas. »

Et il ajoute une illustration de la nécessité de supposer N infini :

« Prenons un de ces cas, celui par exemple que, dans tous les instants toutes les aiguilles placées dans un même lieu affectent la même direction. On demande combien il y a à parier que le même phénomène s'observera l'instant suivant. Il est clair que je dois supposer le nombre de ces observations infini puisqu'il n'y a pas de nombre si grand qu'il puisse être que je ne puisse supposer que celui des cas est plus grand. Cela posé, il me semble qu'étant à l'expérience à déduire [?]. »

d) *Quelques remarques sur ces manuscrits*

Dans l'état actuel des recherches, il subsiste quelques interrogations sur la datation de ces manuscrits : en particulier, le MS 875 ff 116-125, 115 a-t-il été écrit par Condorcet avant qu'il ne prenne connaissance du mémoire de Laplace sur la probabilité des causes par les événements? Ont-ils discuté entre eux de ce problème? Ni Condorcet, ni Laplace n'ont laissé d'indice à ce sujet dans les papiers disponibles. Au-delà de la question de priorité éventuelle, la réponse à ces questions permettrait de mieux suivre le cheminement de la pensée de Condorcet.

Nous pouvons cependant observer ceci : dans ce dernier manuscrit, la position de Condorcet s'est affirmée, il a compris la simplification analytique qu'on pouvait obtenir par la supposition d'un modèle continu pour la composition de l'urne (à savoir $N = \infty$), il discute un peu son opportunité, mais sans pour autant envisager ce que nous appellerions d'autres répartitions *a priori* dans ce cas continu.

e) *A propos des textes publiés de cette période*

Les présentations par Condorcet des mémoires de Laplace ne contiennent pas de calculs d'estimation ni de critiques sur les calculs mêmes de Laplace. Pour ce qui nous concerne ici, le commentaire de [1781] reprend une idée présente dans les manuscrits précédents :

« On peut supposer le nombre des boules fini ou infini [...]. Le cas du nombre infini est celui qui a lieu lorsqu'on applique les questions aux événemens naturels : en effet, il est aisé de voir qu'alors elles embrassent l'immensité des temps, et que le nombre des combinaisons est infini. »

Enfin, dans l'Appendice du livre [1777], la formule de Bayes-Laplace est utilisée comme une formule classique sans qu'aucun doute ne soit émis sur sa validité.

f) *Le manuscrit MS 873 f 289*

Nous disposons encore du fragment manuscrit suivant :

« Soit un nombre inconnu de boules tel que 1 *représente le nombre total de boules, x celui des boules blanches,* 1 − *x celui des boules noires, que l'on ait tiré m boules blanches, n boules noires et qu'on demande la probabilité de tirer une boule blanche, ou p boules blanches et q boules noires.*

On peut envisager le problème de différentes manières.

1° On peut prendre la probabilité moyenne depuis x = 0 ... *jusqu'à x* = 1 *d'amener p + m blanches et q + n noires et la diviser par la probabilité d'amener m blanches et n noires. C'est la méthode de M. de Laplace, cette méthode n'est pas rigoureuse. Nous aimerions mieux dans ce cas prendre la probabilité moyenne d'amener p, q et la diviser par la probabilité d'amener p + m, q + n. Ceci n'est pas non plus rigoureux.*

Voici la méthode qui me paraît préférable.

Je suppose qu'on ait tiré m et n. Si on cherche la probabilité que le rapport de x à 1 − *x soit a étant quelconque il sera infiniment petit mais on peut chercher la probabilité s'il sera entre a − b et a + b, et il résultera que si <m et n sont très grands> l'on prend a = m/n, la quantité b <et* 1*> [?] qui marquera la probabilité serait très grande à ce point et qu'ainsi b [?] pourra être assez petit.*

L'objection contre cette manière de considérer l'objet vient de ce qu'elle est indépendante de la grandeur m, ce qui est faux, mais l'objection est la même dans les autres cas si p et q sont indéfinis, il faudra donc si on veut avoir égard à cette observation supposer que p + m, q + n soi[en]t finis, alors le nombre de boules sera fini et le problème dépendra des quadratures aux différences finies. »

Le texte s'arrête là; dans le recueil MS 873 des manuscrits, plusieurs autres feuilles éparses, contiguës à celle-ci, de formats divers, contiennent des brouillons de calculs touchant aux témoignages et aux jugements; elles pourraient être reliées à ce fragment, mais cela n'est pas absolument certain. Alors le contenu de ces brouillons évoquerait naturellement divers passages de l'Essai. Puisque les textes tant publiés que manuscrits de Condorcet jusqu'à l'année 1781 n'émettent aucune réserve sur la pertinence de la formule de (Bayes-)Laplace, on pourrait alors penser que c'est la réflexion naissante approfondie de Condorcet sur les témoignages et jugements, lors de la préparation de l'Essai, qui l'a conduit à retravailler de façon critique les méthodes possibles d'estimation. Mais cela ne constitue qu'une hypothèse. Ce manuscrit daterait alors des années 80. En tout état de cause cette note embryonnaire évoque une remise en chantier du problème (*) sous une forme qui préfigure la Troisième Partie de l'Essai et certains passages des Élémens, nous le verrons partiellement plus loin.

4. *Les 3 premiers problèmes de l'Essai et les 3 hypothèses du Mémoire*

Nous passons maintenant aux écrits de maturité : limitons-nous au 1°) de (*) et indiquons les réponses de Condorcet par un tableau d'ensemble, en précisant les différentes conditions de validité, les interprétations en termes de schémas d'urnes et les justifications qu'on peut en donner.

Pour exprimer les schémas d'urnes correspondants, nous supposerons une suite d'urnes U_k dont la composition inconnue (c'est-à-dire le rapport du nombre de boules blanches au nombre total de boules) sera notée x_k.

Les différentes solutions générales proposées :

a) *Problème 1, p. 176-177, dans l'Essai, et § I-II, p. 539-541, dans le 4e article du Mémoire :*

Conditions de validité : 1°) La probabilité inconnue de l'événement *A* reste constante, soit *x*, dans toute la série d'épreuves, ou bien 2°) cette probabilité est variable, mais *« quoique pouvant être* *différente pour chaque événement, est cependant prise au hasard pour chacun d'après une certaine probabilité générale x pour A ».*

Solution :

$$\int x^{m+1}(1-x)^n dx \,/\, \int x^m(1-x)^n dx \;=\; (m+1)/(m+n+2).$$

Interprétation (explicite) en termes d'urnes : 1°) à chaque épreuve la boule est tirée de la même urne, ou de diverses urnes ayant toutes la même composition; 2°) à la k^e épreuve la boule est tirée de l'urne U_k, les x_k sont inconnus et éventuellement différents, mais on sait que toutes les urnes ont été remplies en tirant des boules au hasard d'une seule urne (−mère) de composition inconnue x, qui représente la valeur moyenne des x_k.

Commentaire : c'est la solution de Bayes-Laplace. Si m et n sont grands, ce résultat est évidemment proche de la fréquence observée $m/(m+n)$.

b) *Problème II, p. 177-178, dans l'Essai :*

Conditions de validité : La probabilité de A *« n'est pas la même dans tous les événements, mais [...] elle peut avoir pour chacun une valeur quelconque depuis zéro jusqu'à l'unité ».*

Interprétation (implicite) en termes d'urnes : à la k^e épreuve, la boule est tirée de l'urne U_k, les x_k sont indépendants les uns des autres et équirépartis sur [0,1].

Solution :

$$\{(\int x dx)^{m+1}\,[\int(1-x)dx]^n\} \,/\, \{(\int x dx)^m\,[\int(1-x)dx]^n\} \;=\; 1/2.$$

Commentaire : cela revient à dire que cette probabilité varie d'une épreuve à l'autre de façon indépendante de sorte que les expériences ne peuvent rien nous apprendre à son sujet (voir aussi le commentaire de Todhunter).

c) *Problème III, p. 178-180, dans l'Essai :*

Conditions de validité : on ignore si à chaque fois la probabilité de A reste la même comme au problème I ou si elle est variable comme au problème II.

Solution :

$$\{[n!\,(m+1)!\,/\,(m+n+2)!]+1/2^{m+n+1}\} \,/\, \{[n!\,m!\,/\,(m+n+1)!]+1/2^{m+n}\}.$$

Justification de Condorcet : on sait que A est arrivé m fois parmi $m+n$, il examine alors séparément sous chacune des deux hypothèses précédentes quelle était la probabilité de ce fait, soit

$\rightarrow\ p_1 \;=\; C^n_{m+n}\; n!m!/(m+n+1)! \;=\; 1/(m+n+1)$, dans le premier cas,

$\rightarrow\ p_2 \;=\; C^n_{m+n}\; 1/2^{m+n}$, dans le second cas,

et il applique alors ce que nous avons appelé ci-dessus le principe de vraisemblance, c'est-à-dire qu'il pondère chacune des solutions des problèmes I et II proportionnellement à ces probabilités, soit :

$$\{1/[p_1+p_2]\}\;\{p_1.[(m+1)/(m+n+2)]+p_2.[1/2]\}$$

Commentaire : Condorcet ne pondère pas en aveugle de façon équiprobable entre les deux hypothèses, parce que, au vu des observations passées, ces deux hypothèses ne sont pas également plausibles. D'autre part la première hypothèse revient à dire que les $m+n+1$ répartitions possibles des nombres de boules blanches et noires sont équiprobables (probabilités égales à $1/(m+n+1)$), bien qu'il y ait évidemment beaucoup plus de combinaisons donnant m blanches et m noires que de combinaisons donnant $2m$ blanches et 0 noire, par exemple. Par contre, la seconde hypothèse correspond à une équiprobabilité sur les combinaisons.

d) *§ V du Mémoire, p. 543-545 :*

Conditions de validité : La probabilité est *« variable, mais indépendante du temps où les événements sont arrivés et de l'ordre dans lequel ils ont été observés ».*

Interprétation (implicite) en termes d'urnes : on dispose de $m+n$ urnes, leurs compositions inconnues x_k sont indépendantes, mais à chaque épreuve on tire au sort équitablement (c'est-à-dire avec probabilité $1/(m+n)$ pour chaque urne) le numéro de l'urne qui servira pour tirer la boule : c'est un procédé de double tirage : ceci revient à prendre pour probabilité à chaque tirage la quantité $(x_1+x_2+...+x_{m+n})/(m+n)$, chacun des x_k prenant indépendamment et uniformément toutes les valeurs possibles entre 0 et 1.

Solution : C'est le quotient de l'intégrale multiple :

$\int...\int [(x_1+...+x_{m+n})\,/\,(m+n)]^{m+1}\,[1-(x_1+...+x_{m+n})\,/\,(m+n)]^n\,dx_1...dx_{m+n}$

par

$\int...\int [(x_1+...+x_{m+n})\,/\,(m+n)]^{m}\,[1-(x_1+...+x_{m+n})\,/\,(m+n)]^n\,dx_1...dx_{m+n}$.

Justification de Condorcet : *« il est clair [que la probabilité] ne peut être assujettie à aucune autre loi que celle qui naît de la probabilité qu'elle sera plutôt la même que différente pour les divers événements ».*

e) *§ VI du Mémoire, p. 545-548 :*

Conditions de validité : « On les suppose dépendants ou plutôt pouvant dépendre de cet ordre ».

Interprétation (implicite) en termes d'urnes : les x_k sont toujours inconnues, indépendantes entre elles et uniformément réparties sur [0,1], mais cette fois, on procède comme suit : à la 1e épreuve on tire la boule de l'urne U_1; à la seconde on tire au sort équitablement entre l'urne U_1 et l'urne U_2, et on effectue le tirage de la boule dans l'urne obtenue...; pour la ke épreuve, on tire au sort équitablement entre les urnes U_1, ..., U_k, et on effectue le tirage de la boule dans l'urne obtenue. Les probabilités successives à utiliser pour le calcul sont donc x_1 pour la 1e épreuve, $(x_1+x_2)/2$ pour la 2e, ..., $(x_1+...+x_k)/k$ pour la ke, les x_k étant indépendants et uniformément répartis sur [0,1].

Solution : C'est le quotient de

$\int...\int f_1(x_1) f_2[(x_1+x_2)/2]...f_{m+n}[(x_1+...+x_{m+n})\,/(m+n)]\,[(x_1+...+x_{m+n+1})\,/\,(m+n+1)]\,dx_1...dx_{m+n+1}$

par

$\int...\int f_1(x_1)\,f_2[(x_1+x_2)/2]...f_{m+n}[(x_1+...+x_{m+n})/(m+n)]dx_1...dx_{m+n}$;

où $f_i(u) \;=\; u$, si A est arrivé au ie coup,

$\quad\quad\;\; = \; 1-u$, si A n'est pas arrivé au ie coup.

Justification de Condorcet : *« comme l'on ne connaît pas la loi de l'ordre des événements, mais qu'on sait seulement qu'il peut en exister une, la méthode consiste, de même que dans l'article précédent, à prendre seulement la probabilité que celle des événements successifs sera ou ne sera pas la même, avec cette seule diffiérence qu'ici l'on a égard à l'ordre que les événements se sont suivis ».*

Comparaison entre ces méthodes d'estimation :

Nous allons d'abord indiquer l'application numérique aux cas $m = 2$ ou 3 et $n = 0$.

Cas $m = 2$:	Cas $m = 3$:
a) 3/4 = 0,75	4/5 = 0,80
b) 1/2 = 0,50	1/2 = 0,50
c) 9/14 = 0,64	21/30 = 0,70
d) 3/5 = 0,60	40824/67200 = 0,61
e) 25/42 = 0,59	1799/3000 = 0,60
f) 1 = 1 (fréquence observée)	1 = 1

Les méthodes extrêmes b) et f) représentent le cas où les observations n'apprennent rien, et celui où elles disent tout. L'ordre b) < c) < a) est évident par construction, puisque l'estimation c) est une moyenne pondérée des deux autres. L'ordre e) < d) < a), noté par Condorcet, s'explique également avec facilité, puisque les hypothèses correspondantes sont de plus en plus restrictives.

Il est donc clair que Condorcet propose diverses réponses au problème (*). Alors oscille-t-il, comme le lui reproche Todhunter, entre une application systématique de la règle de Bayes et une panoplie incohérente de méthodes d'estimation arbitraires? Nous ne le pensons pas.

Bien au contraire, la cohérence de Condorcet ne doit pas être mise en doute : plus que tout autre mathématicien de son époque (et de quelques autres...), il a parfaitement conscience qu'il subsiste une part d'arbitraire dans les différentes méthodes d'estimation, et il cherche justement à localiser cet arbitraire (y compris dans les équiprobabilités supposées), à le traquer, à intégrer le maximum de connaissances positives, pour reculer le plus longtemps possible le moment de son introduction, sans jamais baisser les bras : il l'exprime à plusieurs reprises, en particulier dans le Mémoire :

« dans toute autre hypothèse [que celles formulées au a)], la formule [...de Bayes-Laplace] ne peut être regardée comme donnant les résultats rigoureux; et il faut examiner s'il n'y a pas entre ces hypothèses et celle qui suppose tous les événemens indépendans [c'est-à-dire b)], quelqu'autre supposition qui soit propre à représenter la probabilité, d'une manière plus vraie dans une partie des questions qu'on peut avoir à résoudre; autrement il suffiroit, au lieu d'employer sans restriction la méthode [...de Bayes-Laplace], de suivre celle que j'ai indiquée dans l'Essai ...[à savoir le c)] ». L'objectif de l'article IV du Mémoire est justement, pour une large part, de répondre à cette interrogation. Ici, la méthode du problème III de l'Essai nous semble représenter à la fois une étape intermédiaire de sa réflexion et un pis-aller.

Au § III de l'article IV du Mémoire, Condorcet note qu'une étude plus fine des résultats des épreuves successives, notamment en les considérant par paquets de 100 ou de 1 000, permet d'étudier l'évolution du phénomène, de « tester la stabilité des quotients statistiques » (selon l'expression fort juste de K. Pearson, contre Todhunter). Non seulement cette remarque pertinente est neuve, mais Condorcet l'exploite de manière constructive en cherchant des formules pour traiter l'aléatoire non indépendant, pour se dégager du cas stationnaire [4]. On aura une idée de l'audace théorique en constatant que, soixante ans plus tard, le perspicace Cournot, pourtant confronté au problème, propose en fait de baisser les bras :

« Si les conditions du triage changent d'une épreuve à l'autre suivant une loi connue, il n'y a rien à conclure des événements observés aux événements futurs. Condorcet a donné, pour ce cas, des formules tout à fait illusoires ». [Cournot, 1843-1984]

K. Pearson estime qu'une bonne réponse au problème doit faire intervenir des hypothèses sur la corrélation entre les épreuves successives, et il ajoute :

« This correlation can, as far as I see, only be determined by some hypothesis, or from the observations themselves. Condorcet's hypothesis is only one of many and not a very probable one in itself. If the chances do change with the time then it seems to me that Condorcet by his limits is supposing them always to increase, and his limits of integration are very open to criticism ». (p. 459)

Il n'empêche que la statistique mathématique moderne, suit (sans le savoir), certes par d'autres formules, l'état d'esprit de Condorcet. On se rendra compte de la difficulté éprouvée par les scientifiques du 19e siècle à apporter des réponses satisfaisantes à de telles questions, en lisant [Stigler, 1986].

Nous verrons en outre plus en détail que, dans les applications des méthodes d'estimation opérées par Condorcet, la discussion de ce qu'il est d'usage d'appeler les modèles et des procédés d'estimation est un souci réel : c'est particulièrement frappant pour l'évaluation des droits éventuels.

5. *Les Élémens*

Nouvelle synthèse, et à la fois traité pédagogique inachevé reprenant les cours faits au Lycée, les Élémens abordent à deux reprises le problème qui nous intéresse ici :

— Dans l'article III, Condorcet considère une urne renfermant des boules blanches et noires dans une proportion inconnue, successivement dans deux cas,

a) celui où le nombre total N de boules dans l'urne est connu (p. 65-68),

b) celui où l'on ignore ce nombre (p. 68-79).

Dans le premier cas, se limitant à l'exemple particulier $m = 3$, $n = 1$, $N = 4$, il expose simplement les idées devenues classiques : examen des 5 hypothèses possibles, détermination de leurs probabilités respectives par le principe de vraisemblance, calcul de la probabilité d'amener une boule blanche lors d'une nouvelle épreuve, dans le même esprit qu'au problème III de l'Essai.

Dans le second, il suppose implicitement que N est infini : *« le rapport du nombre des boules blanches au nombre total [peut] être également exprimé par tous les nombres, depuis zéro jusqu'à l'unité ».* Toutes les hypothèses sur la valeur de x étant également plausibles, il donne donc la solution de Bayes-Laplace, à ceci près que le traité étant « élémentaire », il effectue un intéressant calcul « à la main » sans utiliser le signe $\int$ (p. 68-74, pour le cas particulier $m = 3$, $n = 1$; p. 76-79, pour le cas général).

Surtout, il suggère d'autres manières de poser le problème : remarquant que *« la probabilité que x ait une certaine valeur déterminée [est] nécessairement nulle »,* il *« cherche la probabilité qu'il soit plutôt [tel nombre] que [tel] autre »,* puis demande *« s'il est plus probable que x est au-dessus de 1/2 qu'au dessous »* (p. 74-76).

— Dans l'article IV, il envisage encore d'autres utilisations de ces idées débouchant notamment sur ce que nous appelons les tests de signification (p. 94-100).

Ces aspects, liés à la seconde question de (*) mériteraient un tout autre développement que nous ne ferons pas ici.

6. *A quoi Condorcet applique sa théorie et à quoi il ne l'applique pas*

Nous distinguerons d'un côté les thèmes d'illustrations et exemples, de l'autre les problèmes travaillés effectivement par Condorcet; et pour cela nous procéderons approximativement de façon chronologique.

a) *Exemples cités (mais non approfondis) dans les premiers travaux :*

— Le manuscrit de 1772 parle d'appliquer ces principes aux événements de la vie et évoque brièvement les tables de mortalité et l'inoculation.
— Le manuscrit commencé en 1774, dans sa partie corrigée vers 1780, n'évoque que la probabilité de mourir d'une maladie sans précision, le mot « inoculation » ayant été rayé. Le reste concerne des applications à la théorie de la connaissance.
— Le MS 875 ff 116-125, 115 fait allusion aux sciences des faits et aux événements de la vie; un peu plus loin, pour justifier l'hypothèse d'un nombre de boules infini, cite le cas où « dans tous les instants toutes les aiguilles placées dans un même lieu affectent la même direction ».
— Le manuscrit (sans doute plus tardif), MS 873 ff 289 sqq, où la méthode de Laplace est jugée non rigoureuse, est illustré par un calcul sur le nombre des hommes « qui ont le sens commun » [?] et un autre sur « la probabilité du jugement d'un homme ».

— La présentation déjà citée de l'article de Laplace de 1774 parle des usages de la vie, et la fin du commentaire fait allusion à l'art de déduire et à l'art de se conduire.
— Une lettre à Turgot de 1775, concernant partiellement Laplace, note, sans autre précision, que la partie la plus utile du calcul des probabilités est son application à l'économie politique.
— La présentation du mémoire de Laplace prévue dans le volume pour l'année 1778 (paru en 1781) cite l'exemple des naissances (mais c'est celui de Laplace dans l'article), il indique aussi les erreurs d'observations en astronomie et l'inégalité entre joueurs dans les jeux, enfin, pour justifier l'hypothèse du nombre de boules infini, il invoque les événements naturels, et le commentaire final s'occupe de la constance des lois dans les sciences de la nature et dans la conduite de la vie.
— Dans toute cette époque 1772-1781, le seul cas où Condorcet utilise vraiment l'estimation pour une recherche dans un domaine d'application se trouve dans l'appendice au livre [1777] : il s'agit, comme le titre l'indique, de dégager une méthode mathématique générale pratique et *« d'un usage très-simple » « pour trouver les loix des phénomènes d'après les observations »*. En d'autres termes, au cours de toute cette période, sauf peut-être dans son « Mémoire sur le canal de Picardie » (dont nous n'avons pas retrouvé les calculs explicites, à supposer que de tels manuscrits existent encore), la motivation « appliquée » de Condorcet semble toute théorique.

b) *Dans les écrits de maturité :*

La situation est tout à fait différente, il subsiste certes deci delà des exemples d'illustrations, mais on note surtout que la théorie de l'estimation joue un rôle fondamental dans des problèmes d'arithmétique politique et de mathématique sociale étudiés en tant que tels :
— Dans cet ouvrage capital qu'est l'Essai, cela concerne les votes et les jugements. La question, abordée dans [G.G. Granger, 1956] et [R. Rashed, 1974], mériterait une étude mathématique à part, qui dépasse le cadre de cet article.
— Dans la 3e partie du Mémoire, il s'agit d'évaluer le montant équitable du rachat de certains types de droits féodaux appelés droits éventuels ou droits casuels, critiqués à la fin de l'Ancien Régime (*cf.* [P. Crépel, 1988]).
— Dans l'article « Probabilité » de l'Encyclopédie Méthodique, au III.5, il est question d'économie publique en général et les § 6-8 traitent le cas des rentes viagères (*cf.* aussi l'article « Remboursement »), alors que le § 9 reprend une partie du calcul des droits éventuels.
— La 4e partie du Mémoire, spécialement consacrée au sujet qui nous intéresse ici, a pour objet une discussion méthodologique générale et vise en particulier des questions telles que la constance des lois dans les faits naturels ou leur évolution dans le temps. Un passage avait paru sous une forme un peu différente dans [1777].
— Les Élémens, dans l'article IV, dont le contexte concerne les voies de la connaissance, cite deux cas originaux, mais sans entrer dans le détail des calculs : décider, d'après les données connues, d'éventuelles différences entre les êtres humains selon les nations et selon les sexes : en particulier, il s'agit de critiquer certains préjugés.
— Un article prévu pour les Élémens, mais resté inédit (MS Z 39 et MS 873 ff 300-301), étend (comme nous le verrons plus loin) la théorie de Bayes-Laplace au cas où plusieurs événements, au lieu de 2, sont possibles à chaque épreuve, et l'applique aux tables de mortalité et aux rentes viagères.
— Un manuscrit inédit intitulé « Application du calcul à la probabilité des faits » (MS Z 37), et datant vraisemblablement de la même époque, s'attache à appliquer certains aspects de la théorie de l'estimation aux probabilités de témoignages.
— Un manuscrit inédit intitulé « Arithmétique politique ou application des mathématiques aux sciences économiques » (MS 855 ff 177-179) évoque des applications diverses y compris l'évaluation

de la population; il contient des remarques sur les causes d'erreurs. Malheureusement ce fragment annonce divers calculs d'estimation qui ne semblent pas se trouver dans les papiers de l'auteur.
— Condorcet, dans un mémoire signé conjointement avec Du Séjour et Laplace, est confronté à l'estimation de la population de la France; mais, en fait, l'essentiel du Mémoire est dû à La Michodière et il est difficile de savoir dans quelle mesure Condorcet y a vraiment contribué [B. Bru, 1988 a, dans ce même volume].
— Signalons enfin que, dans le IVe Mémoire sur les Monnaies [1790], à propos du *« jugement qui doit constater que les monnaies fabriquées sont au titre et ont le poids déterminé par la loi »*, Condorcet remarque : *« On n'essaye pas toutes les pièces de monnaie fabriquées, mais seulement un certain nombre prises au hasard, dans chaque fabrication »*. Il discute qualitativement les méthodes employées, précisant par exemple *« qu'il faut constater d'abord que ces monnaies, destinées à subir l'examen, ont été réellement choisies au hasard »*. C'est très intéressant, mais l'article ne renferme pas de calculs.

c) *Une première remarque :*

La liste sèche qui précède fait immédiatement apparaître que Condorcet, bien qu'il soit évidemment parfaitement au courant des trois domaines d'application (et de débats) les plus répandus du calcul des probabilités, à savoir les erreurs d'observations, l'inoculation et le problème des naissances, ne les aborde pas personnellement autrement que par allusion ou à titre de brève illustration.

Ses domaines d'intervention concernent exclusivement l'économie publique, la mathématique sociale et la théorie de la connaissance.

7. *Quelques aspects remarquables et méconnus :*

Nous en avons noté trois importants traités plus en détail ailleurs [P. Crépel, 1988].

a) *Cas de plusieurs événements*

Le premier concerne la nécessité d'étendre l'estimation bayésienne au cas où plusieurs événements, et non pas seulement 2, sont possibles, c'est-à-dire, en termes modernes, au cas des variables aléatoires, notion dont la lente émergence serait intéressante à étudier en détail [B. V. Gnedenko, 1986]. Le problème n'est pas qu'évoqué, il est résolu : seul le résultat est indiqué dans l'article « Probabilité » de l'Encyclopédie Méthodique, mais la démonstration se trouve dans le manuscrit inédit destiné à devenir un article des Élémens. La difficulté analytique n'est pas énorme même si la démonstration fait intervenir plusieurs intégrations successives. Ce pas conceptuel, en partie explicite, est intéressant et rejoint le souci manifesté par Condorcet à propos de la théorie de l'espérance : ne pas se contenter d'événements, mais étudier des grandeurs dépendant du hasard. Le résultat s'énonce simplement : si les événements possibles $A_1, A_2, \ldots, A_n$ se sont produits respectivement $m_1, m_2, \ldots, m_n$ fois au cours de $q = m_1 + m_2 + \ldots + m_n$ épreuves passées, les probabilités respectives de ces événements au coup suivant sont

$$(m_1+1)/(q+n),\ (m_2+1)/(q+n),\ \ldots,\ (m_n+1)/(q+n)$$

au lieu de

$$m_1/q,\ m_2/q,\ \ldots,\ m_n/q$$

dans la méthode ordinaire.

Il applique cela aux tables de mortalité et note que la méthode ordinaire, moins exacte, suffit *« lorsque* m_1, m_2, ..., m_n *et q sont tous de très grands nombres : ce qui par exemple n'a pas lieu pour les âges très avancés, même si le nombre des morts contenu dans les tables est fort grand ».*

b) *Estimations relatives à d'autres lois*

Le second point que nous voulons souligner ici a trait à l'estimation bayésienne de la loi géométrique. Dans l'évaluation des droits éventuels, Condorcet examine la possibilité pour représenter les époques aléatoires futures, correspondant aux paiements de certains droits, ce que nous appelons une loi géométrique de paramètre x; et comme on ne peut connaître que les époques antérieures où ce droit a été dû, il formule la valeur du droit par un quotient d'intégrales qui revient à appliquer le procédé de Bayes-Laplace pour estimer le paramètre x de sa loi géométrique en fonction des observations.

c) *La dépendance*

Le dernier point qui nous a semblé particulièrement original concerne toujours l'évaluation des droits éventuels dans le cas de deux droits de types différents dépendant l'un de l'autre. Condorcet, n'esquivant pas la difficulté, construit très explicitement ce que nous appelons aujourd'hui un modèle markovien général dans lequel là encore il estime les paramètres par un procédé bayésien dans le même esprit. Ainsi donc se trouvent réunis dans ce passage, complètement ignoré jusqu'à nos jours, une tentative délicate d'estimation dans un cadre compliqué et le premier assaut victorieux face à la dépendance en probabilités, anticipant de 120 ans ce dont Markov fera la base de la théorie des processus qui a révolutionné les probabilités au 20e siècle.

Voici ce qu'en dit Todhunter :

« We may add that on pages 687-690 Condorcet gives an investigation of the total value arising from two different rights. It is difficult to see any use whatever in this investigation, as the natural method would be to calculate each separately. Some idea of the unpractical character of the result may be gathered from the fact that we have to calculate a fraction the numerator and the denominator of which involve $n+n'+n''+n'''-2$ *integrations. This complexity arises from an extravagant extension and abuse of Bayes' Theorem. »* (p. 398)

Ce commentaire doit moins faire condamner son malheureux auteur qu'insister sur la modernité du propos de Condorcet, et sur l'imperméabilité du 19e siècle à la dépendance en probabilités.

8. *Premières conclusions :*

Malgré son caractère limité, notre étude permet déjà de dégager quelques conclusions sur les soucis de Condorcet relativement à ce que nous appelons l'estimation statistique :

Dès ses premiers brouillons, il insiste sur l'importance d'une réponse élaborée pour l'évaluation des probabilités des événements futurs en fonction des observations des événements passés : dès 1772, il cherche à approfondir la question de l'écart entre cette probabilité et la fréquence d'apparition dans les épreuves passées; un peu plus tard, il a clairement conscience de la nécessité d'une réponse, différente de la fréquence, lorsque le nombre des expériences passées n'est pas très grand par rapport à celui des épreuves futures.

Puis, dans sa période de maturité probabiliste, il examine avec beaucoup de soin les conditions de validité des méthodes maintenant connues, en les reliant à un projet plus vaste. Ceci est noté en particulier dans [G.G. Granger, 1956], p. 75-80 et [R. Rashed, 1974], p. 52-58.

Mais nous pensons que, même sur la simple question de l'estimation ponctuelle, l'apport de Condorcet ne se limite pas à approfondir la portée de l'application du théorème de Bayes : il procède à l'introduction de formulations mathématiques nouvelles, diversifiées et pertinentes traitant
— de l'évolution dans le temps des phénomènes aléatoires (et cela sous plusieurs formes),
— de l'estimation des probabilités dans le cas de plusieurs événements,
— de l'estimation des paramètres de lois diverses.

Son souci, même embryonnaire, évoque plus qu'implicitement un calcul prenant en compte ce qu'on appellera les variables aléatoires et les processus, et non pas seulement les probabilités d'événements isolés.

Bien sûr, nous n'avons pas abordé l'aspect capital de savoir quelle « confiance » on peut accorder aux estimations ponctuelles, quelles manières plus pertinentes on peut avoir de poser le problème et comment cela s'insère dans la théorie de la connaissance de Condorcet : c'est l'objet de [B. Bru, 1988 b], comme nous l'avons déjà dit en introduction. On disposera alors d'un éclairage d'une autre valeur.

On peut toutefois, auparavant, s'interroger sur les raisons qui ont poussé Condorcet à ces préoccupations bien en avance sur son époque. Par une étude des principaux textes publiés, R. Rashed suggère que *« cette compréhension à la fois meilleure et nouvelle provient précisément des préoccupations moins techniques de Condorcet, du sujet quelque peu différent de sa réflexion »* (p. 54) et ajoute qu'elle répond *« apparemment au désir de Condorcet de mieux manier le calcul des prabablités a posteriori lors de son application au domaine complexe du social »* (p. 57).

Ces remarques nous semblent corroborées par le cheminement de la pensée de Condorcet telle qu'elle s'exprime dans les manuscrits, et par l'examen détaillé des formules compliquées contenues dans certains articles jusqu'ici jugés obscurs, comme celui consacré aux droits éventuels.

On pourrait même dire que le fait de ne se sentir pratiquement jamais astreint à des résultats numériques autorise Condorcet à des audaces théoriques qui auraient paru d'autant plus impossibles dans le cas contraire que les moyens analytiques et les données statistiques de l'époque n'auraient jamais permis de surmonter les difficultés.

Il est vrai d'autre part que la multiplicité des paramètres et leur variabilité dans les phénomènes sociaux pousse à un approfondissement théorique différent de celui qu'appelle le problème des erreurs d'observations en physique; néanmoins il faut remarquer que dans la question des droits éventuels, les paramètres ne sont ni nombreux, ni complexes, ni variables, et la capacité de construction de modèles originaux apparaît là, pour Condorcet, avec la puissance théorique méconnue qui s'exprime dans ses travaux de mathématiques pures (*cf.* [C. Gilain, p. 1988]).

Notes

1. Outre les ouvrages classiques d'histoire des statistiques de I. Todhunter et K. Pearson, on pourra notamment consulter [S. Stigler, 1986] qui constitue une référence essentielle (mais où Condorcet est malheureusement presque absent), la traduction annotée et commentée par J.P. Cléro de [T. Bayes, 1764-1987] et la réédition annotée et commentée par B. Bru de [Laplace, 1814-1986]. Parmi l'abondante littérature traitant de l'histoire et des principes de l'estimation statistique signalons trois articles peu connus : [J.E. Estienne, 1903-04], [M. Fréchet, 1947], [M. Dumas, 1969].

En ce qui concerne Condorcet, les principales études assez récentes sur ses conceptions philosophiques de l'inférence statistique se trouvent dans [G.G. Granger, 1956], [R. Rashed, 1974], [K.M. Baker, 1975]. Nous avons indiqué une bibliographie des articles ultérieurs dans [Collectif, 1986]; il faut signaler aussi la conférence orale (non publiée) de R. Rashed au Collège de France, en 1985 : « L'introduction du concept de probabilité conditionnelle et l'invention du théorème de Bayes ». Enfin, l'article [B. Bru, 1988 b] est, à bien des égards, complémentaire de notre étude.

2. Le fonds de manuscrits du Bureau des Longitudes a été retrouvé en 1983, dans des circonstances étonnantes; malheureusement une partie en a disparu et fait actuellement l'objet de spéculations qui la rendront peut-être inaccessible aux chercheurs (voir le dossier Condorcet aux Archives de l'Académie des Sciences). La partie sauvée a été inventoriée par M. Chapront et déposée à la Bibliothèque de l'Observatoire.

En ce qui concerne les manuscrits scientifiques déposés à la Bibliothèque de l'Institut, un nouveau catalogue est en cours, mais celui de Bouteron et Tremblot donne déjà quelques indications incomplètes. Voir aussi [L. Cahen, 1904].

3. Pour la transcription faite ici des manuscrits, nous avons en général conservé les fautes, l'orthographe, les ratures et signalé les mots incertains; nous avons par contre modifié certaines notations et indices afin de les harmoniser dans tout le cours de cet article. Les passages entre [] indiquent nos ajouts; les passages entre < > indiquent les parties rayées par Condorcet lui-même, et lorsque ceux-ci sont sans importance ou illisibles, nous l'avons noté par <*>.

4. La Bibliothèque de l'Institut dispose de plusieurs brouillons ou fragments de brouillons de cette IVe partie du « Mémoire sur le calcul des probabilités » : MS 873 f 179; MS 873 f 183; MS 873 ff 176-177, 174-175, 178; MS 875 f 191; MS 875 ff 192-192 bis. Pour tenir compte de ces variantes, il faudrait entrer dans des précisions encore plus minutieuses, qui n'infirmeraient pas la trame générale de notre article.

BIBLIOGRAPHIE

1) Écrits de Condorcet traitant de l'estimation statistique :

a) Textes publiés :

L'édition en 12 volumes des œuvres de Condorcet (1847-1849), par A. Condorcet-O'Connor et F. Arago, chez Firmin-Didot, Paris, ne contient pas les textes scientifiques, mais inclut toutefois certaines réflexions liées aux sciences; nous en mentionnons quelques extraits ci-dessous.

Dans l'ordre probable de publication, les textes de Condorcet traitant directement du problème de l'estimation sont les suivants :

(1774) : « Sur les suites récurrorécurrentes et le calcul des probabilités », Préface des Mémoires des savants étrangers, tome VI, p. 17-19 : présentation des articles de Laplace.

(1777) : « Essai d'une méthode pour trouver les loix des phénomènes d'après les observations », dans d'Alembert, Bossut, Condorcet : « Nouvelles expériences sur la résistance des fluides », Paris, p. 195-230.

(1781) : « Sur les probabilités », Hist. Acad. Roy. Sci. pour l'année 1778, p. 43-46 : présentation de l'article de Laplace inséré dans ce volume.

(1785) : « Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix », Paris, réimpression Chelsea 1972. Essentiellement dans la troisième partie : Discours Préliminaire, p. lxxxij-xciij et Essai, p. 176-213.

(1785) : Article « Probabilité » dans la section « Mathématiques » de l'Encyclopédie Méthodique, Panckoucke, tome II. Plus précisément § III, p. 656-663.

(1786) : « Mémoire sur le calcul des probabilités », « Quatrième partie. Réflexions sur la méthode de déterminer la probabilité des événemens futurs, d'après l'observation des évenemens passés », Hist. Acad. Roy. Sci. pour l'année 1783, p. 539-553.

(1805) : « Élémens du calcul des probabilités et son application aux jeux de hasard, à la loterie, et aux jugements des hommes », édité par Fayolle, Paris. En particulier dans l'article III, p. 65-79 et dans l'article IV, p. 94-100. Ce texte posthume a été écrit vers la fin des années 80.

Principaux autres textes abordant la question de façon partielle et indirecte.

(1780) : « Mémoire sur le canal de Picardie », O.C., t. XI, p. 315-350.

(1784) : « Mémoire sur le calcul des probabilités », « Seconde partie. Application de l'analyse à cette question : déterminer la probabilité qu'un arrangement régulier est l'effet d'une intention de le produire », Hist. Acad. Roy. Sci. pour l'année 1781, p. 720-728.

(1785) : Idem, « Troisième partie. Sur l'évaluation des droits éventuels », Hist. Acad. Roy. Sci. pour l'année 1782, p. 674-691.

(1786) : Idem, « Cinquième partie. Sur la probabilité des faits extraordinaires », Hist. Acad. Roy. Sci. pour l'année 1783, p. 553-559.

(1786) : « Essai pour connaître la population du royaume... » (avec Duséjour et Laplace), Hist. Acad. Roy. Sci. pour l'année 1783, p. 703-718.

(1787) : « Mémoire sur le calcul des probabilités », « Article VI. Application des principes de l'article précédent à quelques questions de critique », Hist. Acad. Roy. Sci. pour l'année 1784, p. 454-468.

(1787) : « Discours sur l'astronomie et le calcul des probabilités », O.C., t. I, p. 482-502.

(1790) : « Quatrième mémoire sur les monnaies », O.C., t. XI, p. 624-641.

b) Manuscrits :

Ils proviennent de deux fonds : Bibliothèque de l'Institut (cotes MS 848 à 885) et Bureau des Longitudes (cotes Z 19 à 45). Ce sont :

Le manuscrit commencé en 1772 : MS 883 ff 216-221, Z 30 ff 1-6, MS 875 ff 132-133, en particulier ff 3-6.

Le manuscrit commencé en 1774 : MS 875 ff 84-99 et 100-109, en particulier ff 101-103.

Un autre manuscrit antérieur aux articles publiés : MS 875 ff 116-125, 115, en particulier à partir du f 122.

Une lettre inédite de Condorcet à Turgot (écrite pendant l'été 1775) : MS 855 ff 157-158.

Un brouillon contenant une note critique sur Laplace et quelques fragments : MS 873 ff 289-292 bis.

Plusieurs brouillons successifs de ce qui donnera la quatrième partie du Mémoire sur le calcul des probabilités : MS 873 ff 174-178, f 179, f 183; MS 875 ff 191-192 bis.

La note sur le chapitre II de la Thèse de Nicolas Bernoulli : MS 875 f 190.

Un texte prévu pour constituer un article des Elémens de calcul des probabilités : Z 39 ff 1-4, MS 873 ff 300-301, Z 39 ff 5-6.

2) Articles et ouvrages sur Condorcet (cités ici) :

K.M. BAKER (1975) : *Condorcet — From natural philosophy to social mathematics,* The University of Chicago Press.

B. BRU (1988 a) : « Estimations laplaciennes. Un exemple : « la recherche de la population d'un grand empire » (1785-1812) », dans ce même volume.

B. BRU (1988 b) : « Condorcet, statistique mathématique et théorie de la connaissance », *Revue de Synthèse,* CXIX n° 1, à paraître.

L. CAHEN (1904) : *Condorcet et la Révolution Française,* Paris, Alcan.

Collectif (1986) : Compte rendu de conférences et bibliographie : « Condorcet et les statistiques », Séminaire Sciences, histoire, société, Publications de l'Université de Rennes.

P. CRÉPEL (1987) : « Le premier manuscrit de Condorcet sur le calcul des probabilités — 1772 », *Historia Mathematica,* 14, p. 282-284.

P. CRÉPEL (1988) : « Condorcet, la théorie des probabilités et les calculs financiers », dans *Sciences à l'époque de la Révolution Française — Recherches historiques,* sous la direction de R. Rashed, A. Blanchard éd., à paraître.

C. GILAIN (1988) : « Condorcet et le calcul intégral », ibid.

G.G. GRANGER (1956) : *La mathématique sociale du marquis de Condorcet,* PUF.

G.Th. GUILBAUD (1952-1968) : *Elements de la théorie mathématique des jeux,* Dunod, Paris.

K. PEARSON (1978) : *The history of statistics in the 17th and 18th centuries,* édité par E.S. Pearson, London, Charles Griffin, en particulier p. 423-505. Ces travaux historiques furent exposés oralement entre 1921 et 1933, mais non publiés avant 1978.

R. RASHED (1974) : *Condorcet, mathématique et société,* avec choix de textes, Hermann.

I. TODHUNTER (1865) : *A history of the mathematical theory of probability,* réimpression Chelsea (1949), en particulier, p. 351-410.

3) Autres références :

La liste suivante est très sélective, on se reportera aussi aux bibliographies des ouvrages cités.

Lorsque deux dates sont mentionnées entièrement, la première indique l'édition originale, la seconde celle utilisée dans le texte.

T. BAYES (1764-1987) : « An essay towards solving a problem in the doctrine of chances », *Phil. Trans 53* (1764), p. 370-418. Ce texte est aussi publié en édition bilingue par Jean-Pierre Cléro, sous le titre « Essai en vue de résoudre un problème de la doctrine des chances », avec de nombreuses notes, des commentaires et références bibliographiques : *Cahiers d'Histoire et de Philosophie des Sciences,* n° 18, Paris (1987).

J. BERNOULLI (1713-1987) : *Ars conjectandi.* Nous citons ici la traduction française de N. Meusnier : « Quatrième partie de l'art de conjecturer », IREM de Rouen.

A. COURNOT (1843-1984) : *Exposition de la théorie des chances et des probabilités,* J. Vrin (1984). Notes de B. Bru.

M. DUMAS (1969) : « Conditions de réception et statistique. Exposé historique, critique et conjectural », *Mémorial de l'Artillerie Française,* p. 631-694.

J.E. Estienne (1903-04) : « Essai sur l'art de conjecturer », *Revue d'Artillerie,* t. 61, p. 405-449; t. 62, p. 73-117; t. 64, p. 5-39 et 65-97.

M. Fréchet (1947) : « Enquête sur l'estimation statistique des paramètres », Institut International de Statistique, Congrès de Washington.

B.V. Gnedenko (1986) : « Sur l'histoire des concepts de base de la théorie des probabilités » (en russe), *Istor. Metodol. Estestv. Nauk,* 32, p. 81-88.

P.S. Laplace (1774-1891) : *cf* notamment « Mémoire sur la probabilité des causes par les événements », *Mémoires des Savants étrangers,* tome VI, p. 621-656 (1774), repris dans *Œuvres complètes,* Gauthier-Villars (1891), tome 8, p. 27-65.

P.S. Laplace (1814-1986) : *Essai philosophique sur les probabilités,* Christian Bourgois, Paris (1986). Voir notes et postface de B. Bru.

M. Paty (1988) : « D'Alembert et les probabilités », dans *Sciences à l'époque de la Révolution Française — Recherches historiques,* sous la direction de R. Rashed, A. Blanchard éd.

S. Stigler (june 1978) : « Laplace's early work : chronology and citations », *ISIS, 69* N° 247, p. 234-254.

S. Stigler (1986) : *The history of statistics,* Belknap Harvard.