JOURNAL DE LA SOCIÉTÉ FRANÇAISE DE STATISTIQUE

JSFS

Comptes rendus de lecture

Journal de la société française de statistique, tome 146, n° 4 (2005), p. 131-139

<http://www.numdam.org/item?id=JSFS_2005__146_4_131_0>

COMPTES RENDUS DE LECTURE

Histoires de Probabilités et de Statistiques

IREM – Histoire des Mathématiques
coordonné par Évelyne Barbin et Jean-Pierre Lamarche
1 vol., 296 pages, Ellipses, 2004
ISBN 2-7298-1923-1

C'est parce que chacun des articles de ce livre est intéressant que ce compte-rendu est un peu long ! Cela ne doit pas vous décourager de lire le compte-rendu ... et le livre !

L'ouvrage rassemble les interventions faites lors d'un colloque qui s'est tenu à Orléans les 31 mai et 1er juin 2002 sur « Histoire des Probabilités et des Statistiques », colloque organisé par la commission inter-IREM « Épistémologie et Histoire des Mathématiques ».

Le titre du colloque a été repris mais légèrement modifié pour mieux rendre compte du contenu : il ne s'agit évidemment pas d'une histoire exhaustive des probabilités et des statistiques mais de divers éclairages historiques sur la naissance et le développement du calcul des probabilités et de la statistique, sur la rencontre entre ces deux disciplines et leurs apports mutuels au fil des ans.

En voici le sommaire.

4. Histoires de causes : probabilités des causes et investigation des causes

9. La controverse antique sur les futurs contingents, *Michèle Villetard Tainmont et Joëlle Delattre*

10. Laplace et la « Théorie analytique des probabilités » : itinéraires de découverte, *Jean-Pierre Lubet*

11. Cournot. Statistique et raison des choses. *Thierry Martin*

5. Histoires de l'enseignement des probabilités et des statistiques

12. Sur l'histoire de l'enseignement des probabilités et des statistiques, *Norbert Meusnier*

13. Galilée ou Descartes? Étude d'un scénario d'introduction historique au calcul des probabilités, *Éric Butz*

Cet ouvrage est intéressant à plusieurs titres, surtout pour des enseignants de probabilités et de statistique, à quelque niveau que ce soit. En effet, les différents articles ne se résument pas à des anecdotes historiques permettant d'illustrer telle ou telle notion, mais présentent des résultats de recherche, bien documentés et commentés, apportant des éclairages et des réponses à des questions d'actualité concernant la conception, la pratique et l'enseignement de ces disciplines.

1) Dans le premier article, N. Meusnier reprend le « problème des partis », objet de la célèbre correspondance entre Pascal et Fermat durant l'été 1654 et souvent présenté comme à l'origine du calcul des probabilités. L'objectif est le partage équitable de la mise dans un jeu de hasard (composé de plusieurs parties) interrompu avant la fin. N. Meusnier présente le problème avec la solution de Pascal, qui repose sur l'espérance du gain décomposée en donnant « une valeur à chacune des parties », idée que l'on retrouvera dans des textes plus anciens. L'auteur rappelle les recherches déjà effectuées sur des textes parus entre 1494 (Pacioli) et 1654 avant de présenter et commenter des textes, récemment découverts, extraits de traités d'arithmétique commerciale datant probablement du début et du milieu du XV^e^ siècle. Il est intéressant de noter qu'un auteur, sans doute maître d'abaque, conseille à ses élèves de tenir secrète la solution, ce qui peut laisser penser que ces problèmes de partage sont connus des maîtres d'abaque du début du XV^e^ siècle.

2) Dans le deuxième article, D. Lanier et D. Trotoux présentent le 5ème exercice, dernier et plus difficile exercice du traité de Huygens, « *De rationiciis in ludo aleae* », paru en 1657, dont voici l'énoncé :

> « *Ayant pris chacun 12 jetons, A et B jouent avec 3 dés à cette condition qu'à chaque coup de 11 points, A doit donner un jeton à B, mais que B en doit donner 1 à A à chaque coup de 14 points, et que celui là gagnera qui sera le premier en possession de tous les jetons. On trouve dans ce cas que la chance de A est à celle de B comme 244 140 625 est à 282 429 536 481* »

Les trois principaux ouvrages de probabilité, publiés après ce traité entre 1708 et 1718, sont dus à Pierre Rémond de Montmort, Jacques Bernoulli et Abraham de Moivre. Chacun apportera une solution originale au 5ème exercice de Huygens.

D. Lanier et D. Trotoux analysent les solutions et les méthodes utilisées par ces trois auteurs ainsi que la méthode utilisée par Euler quelques années plus tard.

Il est bien intéressant de voir l'évolution des techniques de résolution. De Moivre est le premier à donner une solution générale et à raisonner sur les probabilités plutôt que sur les espérances de gain. On retrouve dans l'œuvre de Bernoulli, témoignage de 20 années de recherche, le glissement progressif du calcul des espérances vers celui des probabilités, ce qui le conduit à définir la probabilité comme le degré de certitude avec lequel un événement futur peut éventuellement se produire.

C'est dans les trois ouvrages de Montmort, Bernoulli et de Moivre qu'apparaissent les premières controverses, encore d'actualité, sur le concept de probabilité :

- probabilité « subjective » déterminée à partir d'hypothèses *a priori* (symétrie du dé, équiprobabilité des tirages, ...),
- probabilité « objective » déterminée à partir de la stabilisation des fréquences observées *a posteriori* (approche « fréquentielle » des probabilités).

La seconde approche conduit aux énoncés regroupés sous le nom de « loi des grands nombres » dont une première version, de Jacques Bernoulli, est publiée en 1713 dans *Ars Conjectandi.*

Pour les « subjectivistes », la probabilité est une opinion ou une croyance et il est illusoire de penser qu'elle livrerait une propriété objective des événements. Ils se retrouvent sur la position défendue par Bayes et étudiée dans l'article suivant.

3) Ce troisième article, de J.-P. Cléro, a pour titre : « la portée physique et sociale de la règle de Bayes ». Les textes de référence ont été publiés en 1763 et 1765, après la mort de Bayes, par son ami Price. L'intérêt de ces textes ne se résume pas à la célèbre formule de Bayes mais à son contexte d'utilisation dans un cadre chronologique, à la possibilité de réactualiser la probabilité de réalisation d'un événement lorsque l'information augmente et surtout à l'interprétation de la probabilité qui en résulte. « La réalité bayesienne », écrit J.-P. Cléro, « est complexe ; elle n'est pas celle d'une nature dont Dieu serait l'auteur ; elle est, au contraire, dès le départ, fortement « anthropologisée » ou « humanisée ». Le sujet bayesien, le *I guess* qui est mis en scène dans toutes les démonstrations, suppute des probabilités, se donne une marge de manœuvre à l'égard des événements et calcule avec quelle chance d'avoir raison il fait cette supputation ou s'accorde la marge de manœuvre. »Après avoir étudié la portée de la règle de Bayes en physique, J.-P. Cléro s'intéresse, dans le cadre des sciences sociales, à un penseur utilitariste contemporain, Harsanyi, qui interprète le *I guess* bayesien comme un *je préfère* et se demande comment le système des préférences est cohérent.

4) Le quatrième article, de B. Parzysz, a pour objet une correspondance de 1669 entre les frères Huygens à propos de la lecture de la table de mortalité publiée sept ans plus tôt par John Graunt. Leur interprétation est fort différente car l'un raisonne sur la médiane et l'autre sur l'espérance mathématique. La controverse les oblige à préciser leur pensée jusqu'à ce que Christian comprenne bien les deux points de vue :

> « *Ce sont deux choses differentes que l'esperance ou la valeur de l'aage future d'une personne, et l'âge auquel il y a egale apparence qu'il parviendra ou ne parviendra pas. Le premier est pour regler les rentes a vie, et l'autre pour les gageures.* »

5) La table de mortalité de Graunt est source de plusieurs problèmes repris par de nombreux auteurs dès sa publication. On la retrouve dans l'article de H. Plane, F. Métin et P. Guyot, qui présentent de nombreuses tables démographiques (tables de natalité, de mortalité, de répartition des sexes à la naissance, ...) ainsi que le contexte de leur émergence et les questions et les commentaires qu'elles ont suscités. C'est une source d'exercices très intéressante pour l'enseignement de la statistique et des probabilités.

6) Dans l'article suivant, M. Henry présente avec beaucoup de clarté la démonstration par Jacques Bernoulli de son théorème dont l'énoncé est :

> « *Soit le nombre de cas fertiles (...) au nombre de tous dans le rapport r/t (...) On peut concevoir des expériences en un nombre tel qu'il soit plus vraisemblable d'autant de fois que l'on veut (soit c) (...) que le nombre des observations fertiles soit au nombre de toutes les observations dans un rapport ni plus grand que $(r+1)/t$, ni plus petit que $(r-1)/t$* ».

Bernoulli obtient le premier intervalle de confiance de l'histoire pour estimer une probabilité. M. Henry ajoute le commentaire :

« Notons que Bernoulli fait intervenir deux notions de probabilité :

- la probabilité p de tirer une boule blanche, valeur déterminée *a priori*, représentant la composition de l'urne, éventuellement inconnue. L'idée géniale de Bernoulli est de l'exprimer sous la forme r/t en se laissant comme paramètre disponible la valeur de t.
- la probabilité que la fréquence F_{nt} des succès en nt épreuves reste assez voisine de p quand n est assez grand. Bernoulli préfère appeler cette probabilité 'vraisemblance'. »

7) M. Armatte présente ensuite la théorie des erreurs de 1750 à 1820, qui émerge en astronomie et géodésie et qui cherche à répondre à trois questions : « la combinaison des erreurs en un juste milieu, l'identification d'une loi de probabilité des erreurs, un critère d'optimisation pour déterminer des quantités dans un système surdéterminé. » La synthèse produite par Laplace et Gauss dans les années 1810 et 1820 conduit au triptyque : « moyenne arithmétique », « loi de Laplace-Gauss » et « méthode des moindres carrés ».

M. Armatte présente les conditions d'émergence des trois questions, montre bien les apports mutuels du développement théorique et de l'amélioration des instruments de mesure et décrit les tentatives proposées pour sortir du paradigme de la loi normale. Ce n'est qu'à la fin du XIX^e siècle qu'une autre solution émerge : «médiane», «première loi de Laplace», «méthode des moindres écarts absolus».

«L'omniprésence du modèle de Normalité», écrit M. Armatte, «se perpétue essentiellement par les usages statistiques qu'en fera Quetelet, dans une transposition de ces découvertes aux sciences morales, qui suscitera plusieurs controverses sur le concept d'homme moyen et sur l'hypothèse d'homogénéité».

Il conclut sur cette phrase que les enseignants devraient méditer : «Les tentatives faites dans l'après guerre pour fusionner ces deux approches (traitement de l'erreur dans les sciences exactes et traitement de la variabilité dans les sciences de l'homme) au sein d'un même corpus rebaptisé «statistique mathématique» montrent alors leurs limites : l'utilité de ces synthèses à des fins d'unification des sciences ou à des fins pédagogiques cache mal que les hybrides produits, cohérents d'un point de vue syntaxique, sont parfois totalement inconsistants d'un point de vue sémantique et pragmatique».

8) C'est cette période d'émergence de la statistique mathématique, entre 1920 et 1940, qu'étudie M. Zerner dans l'article suivant, avec pour objectif «d'orienter directement l'enseignement, en particulier dans la recherche de situations et d'exemples adéquats». Il est bien intéressant de retrouver dans les textes cités et leurs commentaires des questions qui font encore débat aujourd'hui autour des tentatives de définir «une expérience aléatoire», «une probabilité», «une population» et «un échantillon». Après avoir défini la tâche du statisticien comme étant «la réduction des données», Fisher écrit : «*cet objectif est accompli par la construction d'une population hypothétique infinie dont les données réelles peuvent être considérées comme un échantillon aléatoire*». Ce vocabulaire de la statistique se réfère à la notion de «population finie» dont on extrait une partie appelée «échantillon» ; lorsqu'il est transposé dans un contexte expérimental, il n'est pas sans poser problème encore aujourd'hui.

9) La quatrième partie commence par la question de savoir si le futur est nécessaire ou contingent chez Aristote et la controverse qui en a suivi chez les stoïciens et les épicuriens (M. Villetard Tainmont et J. Delattre). Il n'est pas aisé pour un non spécialiste de saisir toute la subtilité de la logique aristotélicienne mais les questions qu'il soulève et les commentaires des auteurs nous éclairent sur les liens entre cette controverse et les difficultés que peuvent rencontrer des élèves lors de l'introduction des probabilités. Sans faire intervenir explicitement le temps dans une introduction aux probabilités, il s'agit bien de quantifier la possibilité de réalisation d'un événement «futur». Le statut déterminé ou indéterminé, nécessaire ou contingent, des événements futurs constitue un réel obstacle.

10) Dans l'article suivant, J.-P. Lubet propose de présenter et commenter plusieurs passages de l'œuvre de Laplace. L'ouvrage principal, «*la Théorie analytique des probabilités*», est publié en 1812 et comprend deux parties, la première sur les fonctions génératrices, la seconde sur les probabilités (et le

théorème central limite). Cet ouvrage est réédité en 1814 avec une introduction aux probabilités bien plus développée et publiée séparément ensuite sous le titre *Essai philosophique sur les probabilités.* Laplace est déterministe. Pour lui,

> « *le hasard n'a 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* ».

Dans un mémoire de 1776, une urne contenant deux sortes de billets lui permet de présenter deux classes de problèmes : sur « la probabilité des événements » et sur « la probabilité des causes des événements » (problèmes directs et problèmes inverses).

> « *L'incertitude des connaissances humaines porte ou 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 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 ou noirs, si l'on en tire un billet blanc, déterminer la probabilité que la proportion de billets blancs aux noirs dans l'urne est celle 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.* »

Le mémoire de 1776 ne traite que de la première classe de problèmes mais, en 1774, Laplace a déjà abordé la deuxième classe dans un mémoire intitulé « *Mémoire sur la probabilité des causes par les événements* ». On y retrouve la formule de Bayes publiée peu de temps avant par Price, comme on l'a vu précédemment, et que Laplace ne semble pas connaître. Sont abordés aussi les premiers problèmes d'estimation.

11) Dans le dernier article de cette partie, intitulé « Cournot. Statistique et raison des choses », l'auteur, T. Martin, présente l'aspect philosophique et épistémologique de l'œuvre de Cournot, à propos de la statistique et de sa relation au réel, à propos du statut de la théorie et de la signification de ses résultats. Cournot cherche à promouvoir la statistique comme science apte à produire des connaissances, donc à accroître notre connaissance du réel. Il insiste également sur les conditions et donc les limites de validité des résultats fournis.

12) Dans l'avant-dernier article de l'ouvrage, N. Meusnier présente les grandes tendances de l'enseignement des Probabilités et de la Statistique en France de 1786 à 2002. Il montre le rôle positif joué par Condorcet et Laplace au début de la période, par Borel, Darmois et Fréchet avec la création de l'ISUP en 1922 ... mais il apporte aussi une contribution importante au rôle négatif joué par Bourbaki. Il cite à ce propos un extrait des mémoires de Laurent

Schwartz : « Dans l'évaluation des probabilités, Bourbaki a commis de franches erreurs. Antérieurement existaient les mesures abstraites ou de Borel sur des ensembles munis d'une tribu. Bourbaki a introduit, sous l'influence d'André Weil et de son livre remarquable « *L'intégration dans les groupes topologiques et ses applications», paru en 1940 et que j'ai beaucoup travaillé, les mesures de Radon sur les espaces localement compacts. Il y a donc deux théories de la mesure, également nobles, très étrangères l'une à l'autre : mesure abstraite et mesure de Radon. Le terme d'abstrait est un peu ridicule. Les « espaces abstraits» de Fréchet sont tous ceux qui sont « quelconques», ils généralisent les espaces usuels, la droite, le plan, l'espace à trois dimensions, « dans lequel nous vivons» ; les mesures abstraites sont des généralisations de la mesure de Lebesgue sur des « espaces abstraits». Bourbaki a totalement privilégié les mesures de Radon et rejeté les autres. Or les probabilités nécessitent les autres. ... Bref, Bourbaki s'est écarté des probabilités, les a rejetées, les a considérées comme non rigoureuses et, par son influence considérable, a dirigé la jeunesse hors du sentier des probabilités.* »

N. Meusnier retrace aussi l'enseignement de la statistique et des probabilités au collège et au lycée depuis les années 1960 et termine sur les derniers programmes du lycée mis en place entre 2000 et 2002. Les probabilités sont introduites à partir de simulations et de la « loi vulgarisée des grands nombres », ce qu'il critique vivement. Cette approche permet de justifier la théorie des probabilités mais, écrit-il, « l'éducation à l'aléatoire devrait avoir pour but fondamental la prise de conscience que toute décision s'accompagne d'un risque, mais que ce risque peut être évalué ».

13) Enfin, dans le dernier article de l'ouvrage, E. Butz propose des activités pour une progression de l'enseignement des probabilités et de la statistique au lycée faisant intervenir l'épistémologie et l'histoire des mathématiques.

En conclusion, malgré la longueur de ce compte rendu de lecture, il est bien difficile de « rendre compte » de la richesse des contributions. Elles permettent au lecteur d'approfondir sa réflexion sur les fondements des probabilités et de la statistique, sur la logique inductive que ces disciplines formalisent (avec risque contrôlé) et sur laquelle repose la démarche scientifique, sur l'enseignement de ces disciplines. Une présentation trop tardive et formelle des probabilités ne permet pas aux étudiants de sortir du raisonnement déductif et certain. Une approche très concrète, à partir de données et de simulations, de techniques de calculs et de procédures, risque de masquer les concepts importants. Ce débat, toujours d'actualité, peut être alimenté par les nombreux éclairages apportés dans cet ouvrage, dont je recommande vivement la lecture.

Jeanne Fine

Le point de vue du nombre 1936

Maurice Halbwachs et Alfred Sauvy
Édition critique
sous la direction de Marie Jaisson et Éric Brian
1 vol., 469 pages, Éditions de l'INED, 2005
ISBN 2-7332-1032-7

En préliminaire de ce compte rendu dans un journal de statistique, il convient de rappeler l'intérêt de Maurice Halbwachs pour cette discipline et la réflexion qu'il a menée sur l'apport de la statistique (et des probabilités) aux sciences sociales. Ce qui s'est manifesté, entre autres, par plusieurs articles (outre ceux dont nous parlerons plus loin citons, par exemple, *L'expérimentation statistique et les probabilités* en 1923 dans *La revue philosophique*), l'écriture d'un ouvrage avec Maurice Fréchet (*Le calcul des probabilités à la portée de tous*, 1924), la participation active aux séances de la Société de Statistique de Paris (SSP) ; Maurice Halbwachs a notamment publié dans le Journal de la SSP (voir plus bas).

« *Le Point de vue du nombre*, paru en 1936, est un ouvrage de sociologie et de démographie particulièrement méconnu. Il fut dirigé par Maurice Halbwachs. À près de soixante ans, il était alors l'un des plus actifs continuateurs d'Émile Durkheim, le fondateur en France de la Sociologie. Il s'était adjoint la collaboration d'Alfred Sauvy [qui] a connu à cette occasion sa première publication notable. Elle gouvernera son destin scientifique : après la seconde guerre mondiale, Sauvy dirigera longtemps l'Institut national d'études démographiques (INED). À leurs côtés figuraient deux autres statisticiens de l'office national de l'époque, la Statistique générale de la France, Henri Ulmer et Georges Bournier. [...] Il s'agissait de la troisième partie du septième tome d'une vaste et prestigieuse entreprise collective lancée au début des années 1930 et prolongée pendant les années 1950 et 1960 : l'*Encyclopédie française.* »

Le texte ci-dessus, extrait du début de la présentation de cette réédition, situe l'entreprise des éditions de l'INED, mais ces quelques lignes ne disent ni toute son ampleur, ni son intérêt pour le statisticien. Il ne saurait être question de faire ici une analyse détaillée de l'abondante information qu'on y trouve. Pour en donner cependant une idée, soulignons d'abord combien l'*Introduction générale*, qui précède *Le point de vue du nombre* proprement dit, représente de la part des éditeurs (Marie Jaisson et Éric Brian, avec des contributions de Walter Gierl, Jean-Christophe Marcel, Jean-Marc Rohrbasser et Jacques Véron) un travail considérable en volume (203 pages) et en qualité, par une mise en perspective très documentée de l'ensemble du projet de l'*Encyclopédie française*, avec évidemment un regard particulier sur *Le Point de vue du nombre* et l'histoire des relations entre les sciences sociales et la statistique. Cette mise en perspective est poursuivie par quantité d'annotations sur les articles ensuite reproduits. Il s'agit d'abord de l'avant propos, par Lucien Fai-