

La statistique : une discipline mathématique à part entière ?

Jean Pierre Raoult est un « professeur des universités émérite », membre du laboratoire CNRS d'Analyse et de Mathématiques Appliquées de l'Université de Marne-la-Vallée, laboratoire particulièrement impliqué en calcul des probabilités théorique et appliqué et en statistiques. Il a enseigné notamment à l'Université de Rouen et au département « informatique » de l'IUT de Paris V.

Il s'est depuis fort longtemps préoccupé de l'enseignement des mathématiques dans les lycées et collèges. Il s'y est fortement investi depuis 2004, en prenant la présidence du Comité Scientifique des IREM. Il y joue un rôle déterminant, fort apprécié, pour impulser des débats de fond pertinents et des analyses critiques, aussi bien que pour défendre efficacement la survie des IREM, très menacés ces derniers temps.

Il est également l'auteur de nombreux articles sur la statistique publiés dans le Bulletin Vert de l'APMEP.

Cet article reprend, en partie, le texte d'une conférence présentée le 22 septembre 2006 à Nantes dans le cadre du Colloque « IREMs du Grand-Ouest ». Je remercie ses organisateurs de m'avoir autorisé à proposer ma contribution pour publication dans PLOT.

La statistique, ma spécialité, fut le lieu en France, récemment, de ruptures institutionnelles, sous forme de changements de programmes dont on sait combien les enseignants de mathématiques les ont souvent durement ressentis. Je citerai à cet égard le nouveau programme de seconde, où l'initiation à l'aléatoire se fait, ce qui était une nouveauté, par l'usage de la simulation et avant l'introduction du formalisme et du vocabulaire probabilistes, qui interviennent en classe de première.

Mais on ne peut pas parler de programmes à des enseignants sans analyser l'évolution de la science qui les sous-tend. Et il y a là, intrinsèquement liées, conti-

nuité et ruptures, sous forme d'un processus de maturation continue, qui débouche sur des ruptures de deux types :

- ruptures conceptuelles (en général assez bien datées) : des savants font émerger (souvent dans des délais rapprochés) un concept ou une théorie innovants ;
- ruptures sociales (plus difficiles à dater) : un corps d'idées ou de pratiques jusque là éparses est désormais vu comme une discipline à part entière.

Nous sommes d'abord amenés à nous préoccuper de ce qui caractérise une discipline scientifique. Nous y distinguerons plusieurs éléments :

- un **centre** reconnaissable (ainsi nous savons dire : cela, c'est de la géométrie ; cela, c'est du calcul différentiel) ;
- des **interactions** avec d'autres disciplines reconnues (d'où parfois des frontières mouvantes) ;
- un catalogue d'**applications** (éventuellement) ;

- des **acteurs** qui s'en réclament (chercheurs, professeurs, éventuellement techniciens) ;
- des **institutions** (sociétés savantes, revues, colloques ...) ;
- une place dans une **classification** (par exemple le calcul différentiel est inclus dans l'analyse, qui est incluse dans les mathématiques, qui sont incluses dans les sciences exactes).

Ma propre pratique de la statistique me conduit à y reconnaître ces différentes caractéristiques. Mais on ne peut faire admettre par les enseignants de collège ou lycée **une matière d'enseignement** si on ne les a pas au préalable eux-mêmes convaincus de son statut de discipline scientifique et, éventuellement, de sa place comme « sous-discipline » par rapport à celle qui, à leurs yeux, fonde leur légitimité disciplinaire (en ce qui nous concerne ici, les mathématiques). Or cette légitimité est largement fondée sur la conscience d'une histoire.

S'impose donc à nous une première interrogation : **comment s'inscrit la statistique dans la culture ?** Hélas, en ce qui concerne la France, la réponse pourrait encore, de nos jours, tenir en un mot : faiblement.

En effet, le mot statistique évoque d'abord pour nos concitoyens (et je pense que les professeurs de mathématiques ne se distinguent pas beaucoup en cela) une **technique** :

- **technique qui paraît rudimentaire** : elle fait appel à des calculs

qu'on imagine assez simples dans leurs principes et à des représentations visuelles (graphiques, camemberts...) et non pas à des théories fines d'exploration de la nature comme le font les applications des sciences physiques ou de la biologie ;

- **technique qui paraît mystérieuse** : on ne sait trop quels programmes d'ordinateurs président aux sorties publiées ;

- **technique qui paraît suspecte**, et ce "par les deux bouts" : absence de garanties sur l'origine des données et risque de manipulation des résultats (par l'état, les puissances économiques...).

L'aspect scientifique de la statistique se trouve ainsi occulté au profit :

- du calcul des probabilités (qui, lui, chez les mathématiciens français, a ses lettres de noblesse) ;
- de l'informatique ;
- des disciplines usagères de statistiques publiées (économie, médecine...).

Mon lecteur ne s'étonnera pas que je m'élève contre ce déficit de reconnaissance. Ce déficit est à la fois analysé et combattu dans le livre d'Alain Desrosières : *La politique des grands nombres : histoire de la raison statistique* (1993, La découverte, Paris). Dans un chapitre où la France est comparée à la Grande-Bretagne, l'Allemagne et les USA :

« En France, l'état est centralisé et la statistique l'est aussi, de deux points de vue,

administratif et territorial. De façon générale, la compétence d'expertise est plutôt interne à l'administration, à travers les corps d'ingénieurs et de fonctionnaires et l'université a une influence moindre que dans les trois autres pays. »

La nécessité d'une émergence de la statistique en tant que discipline reconnue en France a fait aussi l'objet d'un rapport assez récent (2000) de l'Académie des Sciences : *La statistique* (Rapports sur la Science et la Technologie, N° 8, Editions Tec et Doc). Le groupe de travail chargé de la rédaction de ce rapport était présidé par un mathématicien, Paul Malliavin (probabiliste) et comprenait en tout 4 mathématiciens (2 que je classerais plutôt « probabilistes », 2 que je classerais plutôt « statisticiens »), 1 économètre, 2 biostatisticiens, 1 théoricien du signal. Voici quelques extraits de sa préface :

« L'Académie des sciences, après avoir entendu les responsables des organismes de recherche publique, a constaté que la statistique imprégnait tous les programmes de recherche et de développement qu'elle avait examinés. La France - au terme de l'analyse - apparaît en retard dans cette discipline...

Tant en recherche que pour les nombreuses applications industrielles, la part de la France est nettement plus faible qu'elle ne l'est dans de nombreuses autres activités. Ces constats conduisent à proposer des voies et des moyens de développement de cette discipline, essentielle dans les champs scientifiques, technologiques et économiques. Aujourd'hui, il n'est pas exagéré de considérer la statistique en France comme une discipline

émergeant difficilement.

La recherche de base est conduite au sein d'unités de mathématiques appliquées. La spécificité des statistiques n'y est pas clairement reconnue et la discipline est ainsi mise dans l'impossibilité (ou, du moins, dans la difficulté) de se développer. Ceci est le lot classique des disciplines nouvelles. Des dispositions administratives s'imposent pour remédier à cet état de fait. »

C'est Jean-Pierre Kahane, non membre de ce groupe de travail, qui a fait une présentation critique du rapport devant l'académie le 3 juillet 2000. Voici un extrait de ce qu'il dit de ses recommandations :

*« Le rapport distingue bien **les** statistiques, issues d'une grande variété d'activités, et **la** statistique, qui conceptualise et développe les méthodes de recueil et de traitement **des** statistiques. Il dégage l'importance théorique et pratique des statistiques et de la statistique dans un grand nombre d'activités humaines. Il donne une vision d'ensemble convaincante des interactions de la statistique avec plusieurs branches des mathématiques et avec diverses disciplines...*

*Ses recommandations touchent principalement à l'enseignement de **la** statistique, à tous les niveaux, et à la place de la statistique dans les enseignements supérieurs. Les enfants et les jeunes gens doivent se familiariser avec la statistique de manière à ne pas être "écrasés" par les statistiques... À l'université s'impose une statistique en prise avec les métiers de la statistique et avec la recherche internationale, et c'est l'occasion d'une réflexion de*

fond sur la place de la statistique comme discipline scientifique...

En France, à la différence d'autres pays européens, les citoyens n'ont pas une formation suffisante à la prise en compte du mode de pensée statistique. Pour améliorer cette situation, des initiatives récentes ont été prises dans le cadre d'une réforme de l'apprentissage des mathématiques dans l'enseignement primaire et secondaire.

Le groupe souligne l'opportunité de ces réformes et encourage les responsables de l'enseignement à en assurer la mise en œuvre, en particulier par un effort de formation initiale et continue des professeurs des lycées et collèges. »

C'est un geste véritablement militant auquel s'est livrée là l'Académie des Sciences. Bien d'autres pourraient être cités. Évoquons, par exemple le livre de Didier Dacunha-Castelle, *Chemins de l'aléatoire. Le hasard et le risque dans la société moderne*, (Flammarion, 1996). Il y plaide pour une « culture du hasard et du risque » chez nos concitoyens ; en voici un extrait qui, à mon sens éclaire à la fois le besoin et les difficultés d'un enseignement de la statistique imbriqué dans celui des mathématiques.

« *Le manque de rigueur des sondeurs-politologues en appelle à une déontologie du statisticien et le manque de pédagogie des politiques en appelle à une amélioration de la manière de présenter l'information chiffrée. Le recours aux mathématiques du hasard donne un instrument pédagogique important qu'il faut utiliser pour porter en avant la réflexion sur les*

sondages, sur l'acte médical ou sur les risques du nucléaire [...] Peut-être ne médite-t-on pas assez sur l'heuristique statistique qui va bien au-delà de la simple technologie. »

Et Didier Dacunha-Castelle titre l'épilogue de son livre : *Provocation à une pédagogie de l'aléatoire*.

C'est donc bien à des **ruptures** que nous sommes conviés de différents côtés, en ce début du XXI^{ème} siècle dans notre pays, en ce qui concerne la place de la statistique, à la fois en tant que discipline scientifique et en tant que matière d'enseignement. Mais si nous acceptons de donner à la statistique le statut de discipline scientifique, sa spécificité interdit de ne la considérer que comme une branche des mathématiques. Citons ici un extrait d'un article de Jean Dhombres : *La modélisation doit-elle être la partie vive de l'enseignement des mathématiques ?* (Actes du colloque *Quelles mathématiques au lycée ?*, Limoges, 2004, 26-65). Il s'interroge sur différents aspects de la "modélisation" dans l'enseignement des mathématiques ; il considère que, en statistique, *la modélisation est une part majeure de l'acte mathématique* et précise :

L'exemple des probabilités et des statistiques est bien plus profond, alors même que la familiarisation des élèves est encore plus grande dans la mesure où la société vit quotidiennement de statistiques (celles du chômage, celles des morts des suites de la canicule, etc.)... En statistiques, le professionnel doit choisir ses modèles de calcul en fonction des données qu'il a et, s'il n'y a pas de modèle

qui s'impose impérativement, c'est bien que la modélisation y est cette fois une part majeure de l'acte mathématique, jointe à une habitude expérimentale à acquérir... Les données numériques apportent un monde expérimental, et c'est aussi cela peut-être la grande nouveauté pour l'enseignant de mathématiques.

Personnellement, si je souhaite que la statistique et le calcul des probabilités fassent partie du bagage des bacheliers, selon des modalités diverses selon les filières, ce n'est pas essentiellement en tant que composantes de la culture mathématique, mais d'abord en tant qu'éléments de la culture du citoyen (ou plus généralement de l'être humain en tant que sujet doté de sens critique). À ce titre, Condorcet reste pour moi d'actualité, qui écrivait dans son texte sur l'enseignement des « mathématiques sociales » :

« Cette exposition montrera toute l'utilité de cette science : on verra qu'aucun de nos intérêts individuels ou publics ne lui est étranger, qu'il n'en est aucun sur lequel elle ne nous donne des idées plus précises, des connaissances plus certaines ; on verra combien, si cette science était plus répandue, plus cultivée, elle contribuerait au bonheur et au perfectionnement de l'espèce humaine. »

En d'autres termes, je trouve la statistique digne d'être incluse dans la formulation de Paul Gagnon dans son article *What Should Children Learn* (in *The Atlantic Monthly* (décembre 1995, p. 71-72) : « *the essential core of learning that all stu-*

dents in a modern democracy have the right not to be allowed to avoid. »

Mais subsiste la question : **dans quelle mesure la statistique peut-elle être vue comme une branche des mathématiques ?**

En ce qui concerne **la recherche**, cette place de la statistique, en tant qu'usagère de certains outils mathématiques (calcul des probabilités, graphes, algorithmes, algèbre linéaire, ...), ainsi que du mode de pensée mathématique, ne fait pas de doute pour les mathématiciens qui y réfléchissent ; mais elle est facilement oubliée par tous les autres ! J'ai cherché la statistique dans le volume *Mathématiques à venir* : quels mathématiciens pour l'an 2000 ? (Actes du colloque organisé par la SMF (Société Mathématique de France) et la SMAI (Société de Mathématiques Appliquées et Industrielles), École Polytechnique, 9 et 10 décembre 1987.

Je ne l'y ai trouvée que deux fois :

- Mathématiques et sciences de la vie (par Richard Tomassone, professeur de statistique à l'Institut National Agronomique),
- Les mathématiques vues par un grand groupe industriel (par Claude Jablon, directeur scientifique d'ELF Aquitaine).

La situation a progressé sensiblement depuis, mais sans qu'on puisse parler de rupture !

Qu'en est-il pour l'enseignement ? De nombreuses raisons sont avancées pour situer l'enseignement de la statistique au

sein de celui des mathématiques. Certaines ne sont, à mon sens, que partiellement valables ; passons-les en revue :

a. *Le bon usage de la statistique (ou des statistiques) ne va pas sans raisonnement; or le cours de mathématiques est le lieu privilégié d'apprentissage du raisonnement.* Il est vrai, mais force est de reconnaître aussi que, en matière de regard raisonné sur la statistique, il ne s'agit pas uniquement du raisonnement hypothético-déductif qui fonde les mathématiques.

b. *On fait grand usage en statistique, comme en mathématiques, d'outils de calcul numérique et de représentations géométriques.* Il est vrai, mais force est de reconnaître aussi que d'autres disciplines scolaires (physique, biologie, économie...) font aussi usage de tels outils et le font parfois de manière plus accessible aux élèves, car plus proche des besoins immédiats.

c. *L'usage des calculatrices, des ordinateurs, est de plus en plus partie intégrante des enseignements de mathématiques ; or la statistique en constitue un domaine d'emploi privilégié.* Il est vrai, mais force est de reconnaître aussi qu'employer ces outils dans des secteurs plus traditionnels de l'enseignement des mathématiques (tels les calculs approchés de fonctions, d'intégrales, de solutions d'équations numériques ou différentielles...) fournit déjà un vaste champ de mise en œuvre et qu'on pourrait alors envisager de laisser un usage « routinier » des fonctions statistiques aux autres disciplines.

Comme je l'ai par ailleurs développé dans mon article : *Simulation statistique et enseignement* (dans *La place des mathématiques vivantes dans l'éducation secondaire*, Actes de l'université d'Animath, août 2004 à Saint-Flour, Brochure APMEP 168, 264-275), la raison principale à mon avis, pour faire figurer la statistique au sein du cours de mathématiques est la suivante : **le caractère « universel » des mathématiques est indispensable pour donner à l'élève les clefs pour un usage lui aussi universel de la statistique.** Pour mettre ici en avant le danger de la spécialisation dans tel ou tel type d'emploi, que seul le cadre du cours de mathématiques permet d'éviter, je m'appuie sur l'avis de « grands praticiens » de la statistique dans des domaines déterminés, tels Daniel Schwartz (en biologie et médecine) ou Edmond Malinvaud (en économie). Il faut en effet pouvoir à la fois :

- donner vie aux concepts statistiques, indépendamment d'un domaine d'application particulier (et cela, seul l'enseignant de mathématiques peut le faire) ;
- motiver et concrétiser ces concepts par des applications diversifiées, ce qui suppose de fournir au professeur de mathématiques de la documentation appropriée et des occasions de contact avec les collègues d'autres disciplines ;
- mettre en évidence les « vraies structures » compatibles avec la nature de données et les traitements qu'on leur apporte (et ceci encore est l'apanage du professeur de mathématiques).

Cette attention portée aux **structures** est l'un des exemples de la manière dont le professeur peut situer son cours de statistique dans la **continuité** du reste de son cours de mathématiques. Pour prendre un exemple très rudimentaire (et de ce fait souvent traité avec un peu de condescendance par les enseignants), l'étude des quartiles, bien menée par le professeur de mathématiques, sera une occasion de montrer que cette notion est liée uniquement à la structure d'ordre total sur \mathbb{R} et s'étend donc naturellement à tous les types de données régis par cette structure (par exemple les intervalles, même d'amplitudes différentes). Je renvoie ici à deux articles : celui de Valérie Henry, *Une étude sur les quartiles d'une série statistique univariée* (2006, Bulletin de

l'APMEP, 464, 359-372) et celui de Jean-Claude Girard, *Quartiles, déciles et tutti quantiles* (2005, Statistique au lycée, Tome 1 : les outils de la statistique, Commission Inter-IREM Statistique et Probabilités, Brochure APMEP 156, 39-52).

À travers ces citations et ces exemples, j'espère avoir prouvé que, derrière la rupture apparente (au plan thématique et au plan temporel) entre l'enseignement actuel de la statistique en France et celui des mathématiques, il y a en fait une profonde continuité conceptuelle. Et donc je souhaite avoir un peu contribué à l'intérêt, voire à l'éclosion du goût, de mes collègues pour cette discipline proprement scientifique.

