

JSFS

Comptes rendus de lecture

Journal de la société française de statistique, tome 140, n° 4 (1999),
p. 67-77

http://www.numdam.org/item?id=JSFS_1999__140_4_67_0

© Société française de statistique, 1999, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

COMPTES RENDUS DE LECTURE

Statistique théorique et appliquée

Pierre DAGNELIE

Tome 1, Les Presses Agronomiques de Gembloux, 1992, 492 p.

Tome 2, De Broeck Université, 1998, 660 p.

Tous les statisticiens, mais aussi de nombreux utilisateurs de la statistique, connaissent les ouvrages de Pierre Dagnelie. En 1969 et 1970 sont sortis des presses de l'Université de Gembloux deux volumes de *Théorie et méthodes statistiques : applications agronomiques*. En trente ans bien des choses ont changé : de nouvelles tendances ont vu le jour, l'utilisation de l'informatique a modifié bien des pratiques. Pierre Dagnelie a entrepris une refonte complète de ses ouvrages qu'il nous soumet aujourd'hui : le lecteur trouvera dans ces nouvelles éditions un texte entièrement remanié et quelquefois complètement réécrit.

Le premier tome (STAT1) est divisé en quatre parties :

- Introduction générale et collecte des données (2 chapitres).
- La statistique descriptive (2 chapitres : une et deux dimensions).
- La probabilité mathématique et les distributions théoriques (3 chapitres : généralités, une et deux dimensions).
- Les principes de l'inférence statistique (3 chapitres : les distributions d'échantillonnage, les problèmes d'estimation, les tests d'hypothèse).

Le second tome (STAT2) est divisé en quatre parties :

- Notions préliminaires (4 chapitres : le choix d'une méthode statistique, les conditions d'application et l'examen initial des données, les tests d'ajustement et les observations aberrantes, les transformations de variables).
- L'étude des données qualitatives (2 chapitres : proportions et pourcentages, tableaux de contingence).
- Les méthodes relatives aux moyennes et à la dispersion (6 chapitres : dispersion, une ou deux moyennes, analyse de variance à un, deux ou davantage de critères de classification, les comparaisons particulières et multiples de moyennes).
- L'inférence statistique à deux dimensions (5 chapitres : corrélation simple, régression linéaire simple, régression non linéaire et modélisation, régression multiple et modèle linéaire général, analyse de covariance).

Un index bibliographique (plus de 400 références dans STAT1, plus de 780 dans STAT2) et un index des traductions anglaises complètent chaque ouvrage. Les principaux mots clés sont indiqués dans STAT1 et une solution des exercices proposés est fournie dans STAT2.

Au-delà du contenu des deux tomes qui couvre la majorité des situations qu'un statisticien, professionnel ou simple utilisateur, rencontre ou rencontrera dans ses activités professionnelles, l'auteur a une ambition plus grande : certes il veut donner au lecteur la possibilité d'utiliser correctement les méthodes statistiques, mais aussi il souhaite qu'il en comprenne les fondements et les justifications. L'auteur insiste sur le choix judicieux d'une méthode d'analyse ; mais que serait ce choix si l'utilisateur ne savait pas interpréter correctement les résultats ? A l'heure où nombre d'entre eux effectuent les calculs nécessaires sur ordinateur cette précaution est devenue encore plus indispensable qu'à l'époque où les calculs étaient manuels : le temps gagné par l'utilisation d'un logiciel statistique doit être consacré à cette analyse fine.

Comme manuel de base, les deux tomes peuvent être lus à différents niveaux : élémentaire, intermédiaire ou avancé. La typographie utilisée le permet très facilement. Un enseignant de statistique y trouvera des modèles de cours (même pour des enseignements théoriques) ; un utilisateur y trouvera des réponses à ses questions plus pratiques.

Nous n'entrerons pas dans plus de détails techniques sur ces deux ouvrages. Nous voulons simplement dire qu'ils n'ont sûrement pas d'équivalent dans toute la littérature statistique de langue française. Quitte à heurter la modestie de Pierre Dagnelie, nous sommes convaincus qu'il peut remplacer la majorité des autres ouvrages sur le sujet. Ce n'est certes pas son objectif, puisqu'il nous convie bien souvent à aller plus loin, à consulter des ouvrages ou des articles plus spécialisés. C'est un ouvrage de base qui devrait très rapidement se trouver dans toutes les bonnes bibliothèques, et devenir tout simplement le « Dagnelie » !

Richard Tomassone

Introduction aux modélisations mathématiques pour les sciences du vivant

Jacques ISTAS, avec la collaboration de François Rodolphe

Collection « Mathématiques et applications de la SMAI »,
n° 34, 2000, Springer-Verlag.

1 vol., 160 pages, 29 figures. ISBN : 3-540-67254-0

Cet ouvrage devrait occuper une place originale dans la littérature biomathématique de langue française, par le champ couvert et par sa technicité. Il complète et prolonge, sous l'angle méthodologique, des ouvrages plus anciens comme *l'Introduction aux modèles mathématiques en biologie* (E. Jolivet, 1983) et *Modèles dynamiques déterministes en biologie* (Ed. : J.-D. Lebreton et C. Millier, 1982). Par rapport à d'autres ouvrages, destinés à un public comparable, mais plus spécialisés autour d'un outil de modélisation particulier comme *La régression non linéaire, méthodes et applications en biologie* (S. Huet, E. Jolivet, A. Messéan, 1992) ou d'une problématique particulière comme *Identification de modèles paramétriques* (E. Walter et L. Pronzato, 1994), ce texte est nettement plus éloigné de la mise en œuvre pratique de l'opération de modélisation proprement dite (déterministe en tout cas) et des problèmes numériques et statistiques qu'elle peut poser. Son originalité principale par rapport à cette littérature réside dans la confrontation entre les approches déterministe et stochastique de la phase de caractérisation d'un modèle et dans la place réservée à la seconde approche. Si l'on se tourne maintenant vers la littérature anglo-saxonne et les grands classiques que constituent notamment les différents ouvrages de May (*e.g. Stability and complexity in model ecosystems*, 1975) et de Murray (*Mathematical Biology*, 1989) ce texte apparaît sans doute beaucoup plus compact, à la fois moins exhaustif pour ce qui concerne l'approche déterministe mais aussi plus général par son approche stochastique, et d'un niveau mathématique plus élevé. Rapporté aux thèmes applicatifs proposés, notamment génétique, son niveau mathématique est à rapprocher de ce que l'on peut trouver dans certaines monographies des *Lecture Notes in Mathematical Biology* de Springer (*cf.* par exemple l'ouvrage de Maruyama), mais son intérêt est d'être plus général qu'une monographie.

En plus de l'introduction, l'ouvrage comprend six chapitres, qui sont autant d'ouvertures mathématiques fondamentales vers la modélisation des systèmes biologiques : Systèmes dynamiques à temps continu, Systèmes dynamiques à temps discret, Théorie des jeux, Chaînes de Markov et diffusions, Processus arborescents, Estimation et tests d'hypothèses. Trois annexes en forme de résumés techniques terminent ce texte : Equations différentielles ordinaires (EDO), Equations d'évolution, Probabilités. La matière ainsi proposée est riche, tant au niveau méthodologique qu'au niveau applicatif. Les auteurs

ont voulu regrouper dans les trois premiers chapitres un ensemble de modélisations déterministes et dans les deux suivants deux types de modélisations stochastiques. Le dernier chapitre apporte des outils statistiques fondamentaux nécessaires aux deux types de modélisations précédemment abordés : estimation de maximum de vraisemblance, propriétés des estimateurs et tests asymptotiques. Chacun de ces chapitres mêle et entrelace concepts de modélisation particuliers et applications qui leur servent d'illustration immédiate, tirées de problèmes réels d'écologie et de dynamique de populations (chenille de l'épicéa, dynamique de *Tribolium* ou du charançon), de biologie moléculaire (amplification PCR), ou encore de génétique quantitative (localisation d'un gène majeur). Cette alternance systématique, qui peut surprendre par moment, est une des originalités de l'ouvrage et contribue à sa valeur pédagogique et démonstrative, y compris pour un public de mathématiciens. Les auteurs fournissent par ailleurs des annexes de bon niveau technique sur l'étude qualitative des EDO, sur les équations d'évolution, sur les inégalités fondamentales en théorie des probabilités et sur les convergences de processus stochastiques (peut-être auraient-elle pu être complétées par des éléments analogues sur l'identification numérique des modèles et sur l'intégration des EDO, démarches opératoires nécessaires à toute modélisation quantitative). Chacun des chapitres est suivi d'exercices qui complètent bien l'exposé, et de références spécialisées.

Ce texte correspond clairement au second type d'ouvrages publiés dans la collection *Mathématiques et Applications* de la SMAI, « devant conduire le lecteur aux modèles mathématiques bien posés et aux problèmes pratiques liés à leur résolution ». Il peut s'adresser à un public de mathématiciens intéressés par la modélisation en biologie. Il doit s'adresser tout autant à un public de biologistes de bonne culture mathématique, désireux de mieux maîtriser des outils d'analyse et de modélisation de systèmes dynamiques et spatio-temporels. A l'intention des premiers, tous les exemples biologiques objets des modélisations étudiées sont replacés dans leur contexte d'application et souvent enrichis de rappels des notions écologiques ou biologiques concernées, en particulier génétiques. Pour les seconds, l'effort de lecture sera sans doute plus important, mais les annexes devraient leur faciliter la tâche.

Un tel ouvrage devrait donc contribuer à éclairer un jeune public d'étudiants mathématiciens sur d'importants champs d'application de leur discipline favorite et susciter parmi eux quelques vocations de biométriciens. Il devrait permettre également à des chercheurs et étudiants dans les sciences du vivant une meilleure maîtrise de certains outils d'analyse qualitative et quantitative de leurs systèmes biologiques ou écologiques et faciliter ainsi le dialogue avec leurs collègues mathématiciens ou statisticiens.

Jean-Pierre Vila

Réseaux et files d'attente : méthodes probabilistes

Philippe ROBERT

Collection « Mathématiques et Applications » de la SMAI,
n° 35, 2000, Springer-Verlag.
368 pages, 32 illustrations. ISBN : 3-540-67872-7

La modélisation des réseaux de communication et des réseaux informatiques a depuis une trentaine d'années fortement dynamisé la recherche sur les modèles d'attente. Au-delà des outils markoviens de calcul explicite, classiques et faciles d'emploi, mais d'applicabilité limitée, l'effort moderne a surtout porté sur le développement de techniques asymptotiques permettant d'étudier les modèles à différentes échelles de temps ou d'espace. Depuis plus de vingt ans, aucun auteur francophone n'avait entrepris de faire le point sur le large éventail de méthodes probabilistes qui ont été employées pour l'étude des modèles d'attente. Spécialiste reconnu du domaine, Philippe Robert était sans doute un des mieux placés pour combler cette lacune, et on doit le remercier de l'avoir fait de manière exhaustive et claire.

Il était difficile d'unifier dans un même exposé les techniques de processus ponctuels, de martingale, de renormalisation, les limites fluides, les critères de stabilité à base de fonctions de Lyapunov, tout en rappelant les outils classiques du domaine. Le risque était de livrer au lecteur un fourre-tout où tout aurait été survolé sans que rien ne soit approfondi ni surtout que l'intérêt pour l'étude des phénomènes d'attente n'apparaisse clairement. Philippe Robert évite cet écueil avec beaucoup d'adresse. Il s'appuie pour cela sur les modèles les plus simples et les plus classiques, comme les files $M/M/1$ et $M/M/\infty$, pour lesquels l'intuition permet une compréhension aisée des phénomènes. Apporter des résultats nouveaux et revisiter de manière élégante et originale les théorèmes classiques pour des processus aussi connus, voilà qui démontre amplement la puissance des outils introduits et qui emporte l'adhésion du lecteur plus que ne l'auraient fait des développements techniques compliqués sur des modèles plus ésotériques.

Ce livre constitue à mon sens un ouvrage de référence sur un ensemble de méthodes probabilistes dont l'applicabilité ne se limite pas aux phénomènes d'attente. Pour longtemps sans doute, il servira de base aux étudiants et aux chercheurs dans de nombreux domaines des probabilités appliquées.

Bernard Ycart

COMPTES RENDUS DE LECTURE

Le calcul des probabilités compréhensible pour tous

Exercices avec corrigés

Marie BERRONDO-AGRELL et Jacqueline FOURASTIÉ

Gaëtan Morin éditeur, 1998, 198 pages.

ISBN : 2-910749-35-5

Statistiques et probabilités

Rappels de cours, questions de réflexion
Exercices d'entraînement, annales corrigées

Jean-Pierre LECOUTRE

Dunod, 2000, 209 pages

ISBN : 2-10-003998-9

C'est un constat très courant, devenu en quelque sorte officiel depuis le rapport de l'Académie des Sciences de Paris sur la Statistique (voir le compte rendu par Gilbert Saporta, Président de la SFdS, dans ce même volume) : d'une part la France manque de statisticiens professionnels, d'autre part la formation de tout un chacun à cette discipline est insuffisante, deux aspects d'une même carence. Il existe à cela de nombreuses raisons, bien analysées par ailleurs. L'une d'entre elles est sans doute la difficulté à enseigner le calcul des probabilités, à faire admettre que les mathématiques fournissent des modèles utiles pour la vie courante, même en présence d'incertitude quand nous sommes confrontés à des phénomènes aléatoires. Il y a donc un défi pédagogique à relever. C'est ce que les auteurs des deux ouvrages considérés ont cherché à faire (à deux niveaux différents) et c'est la raison pour laquelle il nous est apparu pertinent de rendre compte de ces deux ouvrages sous le même chapeau. Pour ces auteurs, comme pour le présent lecteur, l'initiation au calcul des probabilités ne peut correctement se faire que par l'exercice et, au delà de l'exercice traditionnel « de mathématiques », par l'entraînement à la modélisation de situations réelles simples (parfois même, pour commencer, il faut savoir être simpliste). En outre, à côté des techniques mathématiques de base, il convient d'insister encore davantage sur les concepts à appréhender : la réflexion « généraliste » (par opposition à la seule réflexion technique) est

de la plus haute importance. Mais elle est bien plus difficile à stimuler dans un ouvrage écrit.

L'ouvrage de M. Berrondo-Agrell et J. Fourastié s'adresse en priorité à des élèves d'IUT ; mais, comme le soulignent les auteurs à juste titre, il peut convenir à tous ceux qui souhaitent faire du calcul des probabilités et éprouvent des difficultés à assimiler les notions de base. Ils y trouveront un guide en huit chapitres. Le premier, fort utile à ce niveau, est sur « l'art de résoudre un problème » et, essentiellement, comment passer d'un problème concret à une formalisation mathématique. Dans les trois chapitres suivants, on trouve les bases de logique et de théorie élémentaire des ensembles, en insistant sur l'utilisation des diagrammes de Venn. Trois chapitres introduisent ensuite au calcul des probabilités ; il s'agit presque exclusivement de probabilités discrètes. Le dernier chapitre introduit à la statistique à un niveau très élémentaire : distributions, histogrammes, moyenne, médiane, variance empiriques. Chaque chapitre contient des rappels de cours, mais insiste surtout sur les exercices qui cherchent à être attractifs et de difficulté graduée ; des solutions détaillées sont données. Le souci pédagogique est constant comme indiqué plus haut. L'ouvrage sera donc fort utile pour ceux qui cherchent une initiation aux probabilités en ayant une bonne motivation mais un bagage mathématique et des capacités d'abstraction limitées.

L'ouvrage de J.-P. Lecoutre s'adresse à un public d'un niveau plus avancé ; l'auteur pense d'abord aux étudiants d'Economie et Gestion de premier cycle et de licence, mais des ingénieurs, par exemple, y trouveront aussi une bonne introduction au maniement des probabilités et de la statistique, sur des bases mathématiques solides, sans exiger un degré d'abstraction trop élevé. En matière de probabilités, l'ouvrage considère les bases générales, puis la description des lois de probabilités discrètes et des lois continues réelles et vectorielles, l'espérance mathématique et les moments (mais ne parle pas de fonction caractéristique), les convergences en probabilité et en loi. En matière de statistique, sont considérés les problèmes usuels d'estimation ponctuelle (maximum de vraisemblance, information de Fisher et inégalité de Cramer-Rao) et par intervalles, ainsi que les problèmes de tests ; à côté des tests usuels « du débutant », on trouvera aussi une comparaison méthodologique (et pratique) des approches bayésienne et de Neyman-Pearson. Sans négliger les aspects mathématiques et techniques, les aspects de modélisation sont très présents avec de nombreux exercices qui envisagent des questions très concrètes. Et l'effort pédagogique est important, dans l'organisation même de l'ouvrage comme dans la rédaction. Je recommande ce livre à tous les étudiants qui souhaitent s'initier au calcul des probabilités et à la statistique en ayant un niveau mathématique raisonnable, disons juste inférieur au niveau (de principe !) des étudiants de mathématiques. Mais un grand nombre de ces derniers pourront y trouver aussi une aide précieuse

Henri Caussinus

Un événement : le rapport de l'Académie des Sciences sur la statistique.

« La Statistique »

Académie des Sciences,
Rapport sur la Science et la Technologie n°8
Juillet 2000, Technique & Documentation,
ISBN : 2-7430-0411-8

L'Académie des Sciences vient de consacrer un rapport de plus de 180 pages à la statistique, dans le cadre du rapport biennal sur l'état de la science et de la technologie demandé par le conseil interministériel du 15 juillet 1998. Pour notre discipline, il s'agit d'un événement important pour maintes raisons : la nouveauté, le prestige de l'institution, la qualité et la diversité des contributeurs, le diagnostic porté sur l'état de la discipline.

Ce rapport est l'œuvre d'un groupe de travail animé par Paul Malliavin et réunissant les membres suivants : Paul Deheuvels, Yves Escoufier, Edmond Malinvaud, Bernard Picinbono, Daniel Schwartz, Alain-Jacques Valleron et Marc Yor. Y ont également collaboré à divers titres : Henri Caussinus, Paul Embrechts, Nelly Hanoune, Nicole El Karoui, François Marcotorchino. Parmi les commentaires d'un groupe de lecture critique, on notera les contributions de nos collègues Philippe Besse (au nom de la SFdS), Jean-Luc Besson, Dominique Collombier, Alain Monfort et Elisabeth de Turkheim.

De façon quelque peu inhabituelle, le rapport débute par neuf constats et recommandations, ce qui lui donne un aspect immédiatement offensif. En effet l'Académie des Sciences formule d'emblée le constat suivant : « *La France apparaît en retard dans cette discipline où historiquement elle a joué un rôle majeur... Aujourd'hui, il n'est pas exagéré de considérer la statistique comme une discipline émergeant difficilement...* ». L'Académie dénonce « *l'absence de formation en statistique dans les collèges, les lycées et de vastes secteurs de l'enseignement supérieur qui entrave le développement économique et l'exercice des droits des citoyens* » et déclare « *un effort majeur est nécessaire. C'est une nécessité nationale urgente* ».

A l'appui de ce constat le rapport s'articule en quatre chapitres : présentation de la discipline, recherche méthodologique, interactions avec les autres disciplines, formations et métiers.

Le premier chapitre commence par le rappel indispensable de la distinction entre *la* statistique et *les* statistiques, c'est-à-dire entre d'une part la statistique méthodologique, et d'autre part les données et résultats. Tout statisticien appréciera la nuance, et on sait bien que ceux qui ne la font pas ne connaissent pas vraiment la statistique : à ce propos, on ne peut être que consterné de voir le nombre d'intitulés de cours, d'ouvrages et mêmes de diplômes parlant de statistiques avec un *s*, ce qui prouve que la statistique est trop souvent en-

seignée par des non-statisticiens! Le chapitre 1 brosse ensuite un tableau des liens entre statistique, mathématiques et probabilités, et insiste sur les développements récents liés à l'usage de l'informatique, avec ce que l'on appelle « data mining ». L'emploi de titres comme « statistique sans probabilité » et « statistique et informatique : la nouvelle convergence » donne bien l'esprit de certains développements et insiste sur le fait qu'il ne s'agit pas d'une mode qui va passer mais d'une tendance lourde. Les conclusions de la dernière partie de ce chapitre insistent sur la nécessité de recherches coordonnées en statistique et informatique, avec l'adaptation des cursus universitaires et la création d'un Institut National de Sciences et Techniques Statistiques; et elles évoquent, si rien n'est fait, la menace d'une domination encore plus forte des États-Unis. Si l'on peut juger que certaines parties sur la gestion de la relation-client sont trop détaillées, et d'autres sur les sondages un peu expéditives, ce chapitre pose clairement les problèmes de l'avenir de la statistique au sein de la société de l'information. J'ajouterais pour ma part, en suivant J. Friedman (1997), que la statistique a laissé se développer en dehors d'elle, et a largement ignoré, des pans entiers de la recherche appliquée actuelle, ainsi développés en dehors de la statistique conçue en une acception trop stricte. Voici ce point de vue quelque peu ironique de Friedman : « *One view recognizes that while the amount of data (and related applications) continue to grow exponentially, the number of statisticians is not growing that fast. Therefore our field should concentrate that small part of information science that we do best, namely probabilistic inference based on mathematics. This is a highly defensible point of view that may well turn out to be the best strategy for our field. However, if adopted, we should become resigned to the fact that the roll of Statistics as a player in the information revolution will steadily diminish over time. This strategy has the strong advantage that it requires relatively little change to our current practice and academic programs* ». L'autre stratégie consisterait à : « *First (and foremost) we would have to make peace with computing. It is here to stay; that's where the data is. Computing has been one of the most glaring omissions in the set of tools that have so far defined Statistics. Had we incorporated computing methodology from its inception as a fundamental statistical tool (as opposed to simply a convenient way to apply our existing tools) many of the other data related fields would not have needed to exist. They would have been part of our field* » et de citer : la reconnaissance des formes, les SGBD, les réseaux de neurones, l'apprentissage automatique, les modèles graphiques, la chimiométrie, le traitement du signal...

Le chapitre 2 du rapport est consacré à un état de la recherche méthodologique en statistique en France avec deux contributions. Celle de Paul Deheuvels s'attache à montrer tout d'abord, à partir de comparaisons chiffrées, une exception française en ce qui concerne la population des statisticiens se déclarant tels : si les statisticiens français membres de l'IIS forment le deuxième groupe national après les USA avec 221 membres, on n'y trouve qu'un tiers d'enseignants et chercheurs et deux tiers de « statisticiens officiels » au contraire des autres « grands pays scientifiques », et le réservoir formé par les membres des sections de l'IIS (Société Bernoulli, IASC, etc.) est quasi nul. On peut l'expliquer par la faible appétence des collègues français à adhérer à

des associations, et considérer que cette situation est déplorable au pays de la loi de 1901 ! La faiblesse de cette participation a des conséquences nuisibles sur la représentativité des associations et l'audience de leurs membres. Paul Deheuvels analyse ensuite la production scientifique française en statistique, à partir des travaux de Christian Genest, et tire un constat inquiétant : si la France est au deuxième rang mondial en probabilités, selon le critère du nombre de pages publiées, et au troisième en nombre de pages par habitant, elle n'est qu'au sixième rang et au vingtième rang respectivement pour les mêmes critères, en statistique. De plus le nombre de chercheurs productifs est faible. Le rapport avance que « *le nombre de chercheurs en statistique en France devrait être multiplié par un facteur minimal de 5 pour que le niveau d'activité soit comparable à celui des États-Unis* ». Le rapport propose quelques explications au faible nombre de chercheurs actifs en France : tout d'abord la relative pénurie des formations de statisticiens professionnels et l'appel du marché du travail ont pour résultat que les meilleurs étudiants se dirigent vers des emplois dans les entreprises et les administrations plutôt que vers la recherche et l'enseignement. Dominique Collombier (p. 161) y ajoute la faible attractivité des carrières universitaires. Je serais tenté de rajouter que le manque d'enseignants de statistique entraîne que ceux-ci sont très sollicités pour intervenir dans de nombreuses formations, au détriment de leur activité de recherche. Yves Escoufier insiste d'ailleurs dans la deuxième partie du chapitre 2 sur la dispersion des statisticiens et le manque de lisibilité de la discipline (13 DEA seulement la concernent, en comptant large, souvent intitulés sans référence explicite à la statistique).

Le chapitre 3 est un panorama des principaux domaines d'application de la statistique montrant bien les enjeux et les apports de la statistique à l'agronomie, la médecine, l'économie, la finance, l'assurance et le traitement du signal. Chaque partie est écrite par un connaisseur et acteur du domaine et constitue une introduction aux problématiques. On note au cours de la lecture le manque de spécialistes dans certains domaines (génomique, actuariat).

Le chapitre 4 est consacré à la formation sous ses deux aspects : formation de spécialistes et formation à la statistique. Les deux premières contributions (Yves Escoufier et Daniel Schwartz) sont majoritairement consacrées à la formation du citoyen et des non-statisticiens en se référant à des travaux et expériences français et étrangers, que connaissent bien les membres de la SFdS (voir les diverses tables rondes pendant les journées annuelles). Ici l'enjeu est plus de former à la pensée statistique et au bon usage des méthodes que de s'appesantir sur les fondements mathématiques. Les deux auteurs insistent sur le nécessaire côté expérimental et l'utilisation de situations concrètes. Une dernière partie rédigée par Nelly Hanoune brosse un tableau, forcément rapide, des divers métiers de la statistique pour présenter les formations professionnelles : celles-ci sont diverses et existent à tous niveaux de Bac+2 à Bac+5. Le texte insiste sur le couplage nécessaire avec une bonne formation en informatique, pour conclure qu'actuellement les diplômés sont embauchés très rapidement, et que le développement de doubles compétences, très recherchées, se fait tant en formation initiale qu'en formation continue. On peut regretter

que ce chapitre ne traite pas de la formation des ingénieurs dans les grandes écoles, où les situations sont très diverses et pas toujours satisfaisantes.

Quelles sont donc les neuf recommandations de l'Académie des Sciences ?

Elles sont d'ordre général : « faire un effort pour rattraper le retard de développement de la France », « mettre en œuvre une réflexion concertée entre entreprises et établissements d'enseignement supérieur », « faire un effort de formation des professeurs des lycées et collèges », mais portent également sur des points précis : « constituer des structures de recherche et de formation doctorale », « recruter et former des enseignants et des chercheurs », « afficher clairement des recrutements de statisticiens dans les universités », « tenir compte des collaborations dans l'évaluation des chercheurs et enseignants-chercheurs », « mener des recherches coordonnées en statistique et informatique », « développer la recherche en actuariat avec les organismes professionnels pour aboutir à des thèses de doctorat ».

Au delà de ces recommandations, que certains pourraient juger trop vagues, il est clair que le constat (souvent provocant) de l'Académie des Sciences ne peut qu'être bénéfique pour le développement de la statistique en France. Il est clairement montré que la statistique est une discipline scientifique autonome d'intérêt fondamental, qui utilise, mais sans s'y réduire, mathématiques, probabilités, informatique. On doit donc se réjouir de cette publication qui devrait inciter autorités de tutelle, collègues d'autres disciplines, entreprises, à aider les statisticiens afin de mettre en place les réformes et de donner à la statistique toute la place qui convient. L'Académie des Sciences apporte en effet une caution et un soutien à des arguments trop souvent considérés comme *pro domo*.

Souhaitons donc que ce rapport soit lu par le plus grand nombre, statisticiens et non statisticiens.

Gilbert Saporta

RÉFÉRENCES BIBLIOGRAPHIQUES

- FRIEDMAN, J. (1997). *Data Mining and statistics, what's the connection ?*
<http://www-stat.stanford.edu/~jhf/ftp/dm-stat.ps>
- GENEST, C. (1997). *Statistics on statistics : Worldwide performance based on journal publication over the period 1985 1995*. The Canadian Journal of Statistics, 25, 427-443.
- GENEST, C. (1999). *Probability and statistics : A tale of two worlds ?* The Canadian Journal of Statistics, 27, 421-444.