

Quels statisticiens pour l'Orstom ?

Y. Escoufier^a

Invité à intervenir dans cette table ronde, j'ai cru opportun d'organiser mon intervention autour du texte ci-joint préparé pour une des séances invitées de la 48ème session de l'Institut International de Statistiques tenue au Caire du 9 au 16 septembre 1991. J'ai tenté dans ce texte, à la lumière des expériences de formations mises en place à l'Université Montpellier II, de fournir quelques points de repères pour la définition des profils d'emploi des statisticiens. J'aurai le sentiment d'avoir été compris si le long processus susceptible de mener à un recrutement concernant la statistique s'engage par une définition précise des tâches que devra remplir le futur recruté dans le domaine de la statistique.

L'Orstom a certainement besoin de ces techniciens que j'ai appelé des "utilisateurs occasionnels de la statistique" : Agronomes, hydrologues, biologistes, ils possèdent en plus de leur spécialité spécifique des connaissances sûres sur un nombre limité de méthodes d'usage courant dans leur domaine.

Ces techniciens là n'ont pas les connaissances requises pour suivre en permanence l'évolution de la méthodologie statistique ; ils n'en ont pas non plus le temps puisqu'ils ont d'abord pour objectif des avancées dans le domaine de l'agronomie, de l'hydrologie ou de la biologie. Il me semble donc normal qu'ils puissent s'appuyer sur ces spécialistes que j'ai appelé "statisticiens praticiens". Dans une recherche finalisée en agronomie, la statistique est une technologie comme la microscopie électronique ou l'informatique. Contrairement à la plupart des technologies, la statistique ne repose pas sur des matériels mais sur des savoirs donc sur des hommes et des femmes formés. La décision de recruter un statisticien praticien relève pour moi du même niveau de décision que celle de remplacer un microscope ancien par un plus récent. En tant que "producteur" de statisticiens praticiens il me paraîtrait tout à fait anormal qu'un organisme comme l'Orstom ne se dote pas des personnels nécessaires à l'usage des statistiques les plus avancées. Ils sont déjà présents dans l'Institut. Je ne crois pas être partisan en disant qu'il me paraîtrait normal qu'ils soient plus nombreux.

Faut-il que l'Orstom ait ses propres "statisticiens théoriciens" ? Il serait étonnant qu'ils soient très nombreux. Il est cependant concevable que sur quelques chapitres statistiques particulièrement cruciaux par leurs applications, l'Orstom puisse vouloir être présent dans la communauté des experts traitant de ce sujet.

Pour terminer cette introduction, je rappellerai que mes catégorisations sont des repères. Les hommes et les femmes sont bien sûr trop divers pour se laisser enfermer dans des schémas aussi simples. Il reste qu'en statistique comme dans d'autres domaines, n'importe qui ne peut pas faire n'importe quoi sans y avoir été préparé.

^aUnité de Biométrie Ensa.M - Inra - Univ. Montpellier 2 Place Pierre Viala - 34060 Montpellier Cedex 1 (France)

Que faut-il enseigner en Statistique et à qui ?

Résumé

L'article propose de situer les niveaux de connaissances en statistique par rapport à trois repères : la culture générale, la pratique occasionnelle et la profession de statisticien qu'elle soit celle du théoricien ou celle du praticien.

L'enseignement visant la culture générale sera fondé sur l'observation et les jeux. Au non-statisticien appelé à une pratique occasionnelle de la technologie statistique, on donnera une formation fondée sur des situations de référence. Il saura les reconnaître et sera autonome dans leur traitement. Le professionnel de la statistique connaît les mathématiques et l'informatique nécessaires à la justification des méthodes, à l'étude de leurs propriétés et à la création de méthodes nouvelles. La formation doit donner au praticien une vaste panoplie de méthodes et les moyens de suivre les développements de la discipline. Elle donne au théoricien la possibilité d'aller aux limites les plus extrêmes d'un domaine de la statistique sans lui permettre d'oublier que la justification de la recherche en statistique est dans les applications.

Summary

This paper situates various levels of statistical knowledge with regards to three reference points : general education, occasional practice and professional statistics whether theoretical or practical.

Teaching for general education shall be based on observations and games. Non statistician requiring statistical technology occasionally will be given training based on referential situations. He should be able to recognise the latter and be able to do the analyses alone. The professional statistician knows sufficient mathematics and computer science to justify the methods, to study their properties and to create new ones. The education syllabus should provide the practitioner with a vast range of methods and the means to keep in touch with the subject's contemporary developments. The training should enable the theoretician to go to the extreme limits of a statistical domain without letting him forget that research in statistics is only justified by its applications.

Introduction

Il est devenu courant au moins dans les réunions de statisticiens de soutenir que tout individu est concerné par la statistique : il l'est dans sa vie quotidienne à travers les résumés numériques et graphiques que véhiculent les journaux écrits et parlés ; il l'est souvent dans sa vie professionnelle qui le conduira à prendre en compte des risques, des seuils de tolérance, des pourcentages de pénétration d'un marché, des taux de survie et bien d'autres quantités de nature statistique présentées (peut-être faut-il dire cachées) sous un vocabulaire inspiré du contexte de leur application particulière.

En réponse à ce besoin, que font les professionnels de l'enseignement de la statistique qui doivent par exemple présenter un concept aussi simple que celui de la moyenne ? Feuilletant quelques livres portant les mots "statistiques" ou "initiation à la statistique" sur leur couverture nous trouverons entre autres possibilités :

$$\frac{2 + 4 + 7}{3} \qquad \sum_{i=1}^n \frac{x_i}{n}$$

$$\sum_{i=1}^k p_i x_i \qquad < X, \underset{\sim}{1}_n >_D$$

$$\int_{-\infty}^{+\infty} x f(x) dx \qquad \int x dP$$

Toutes ces écritures représentent effectivement un calcul de moyenne mais elles font référence à des contextes de présentation différents. Le choix d'un contexte particulier de présentation ne peut pas reposer que sur les habitudes de l'enseignant : il tire sa légitimité de son adéquation aux objectifs assignés à la formation. Je vais donc essayer de définir quelques grands niveaux de formation et de préciser les objectifs qui me paraissent leur correspondre. Bien sûr on pourra objecter que les niveaux reconnus ne sont qu'artificiellement extraits d'une trajectoire continue allant de l'enseignement le plus naïf à l'abstraction et la formalisation les plus grandes. J'admets cette remarque mais je persiste à penser qu'il est nécessaire de placer quelques repères le long de cette trajectoire : s'ils ne correspondent pas exactement aux objectifs d'une formation à laquelle nous participons, ils pourront au moins nous aider à la situer en précisant ce que nous voulons faire en plus ou en moins des repères proposés.

1 Les différents niveaux

Je suggère donc de reconnaître trois niveaux, le troisième se dédoublant en deux orientations complémentaires. Pour les présenter rapidement, je leur associerai

des analogies issues du domaine de la santé qui devraient éclairer les ambitions attachées à chacun de ces repères.

- Le premier niveau est celui de la culture générale. Enseigné dans les cycles primaires et secondaires de nos écoles, il a pour but de créer un acquis culturel en statistique. Son pendant dans le domaine de la santé est l'ensemble des règles élémentaires d'hygiène alimentaire et corporelle inculquées aux enfants par la famille et l'école ;
- Le second niveau est celui de l'utilisation occasionnelle de la statistique par d'autres disciplines. Dans un processus de production ou dans une démarche de recherche qui n'ont pour but ni l'un ni l'autre le progrès de la statistique, cette dernière intervient en tant qu'outil en fournissant une procédure (représentation graphique, formalisation, test) qui contribue à la démarche. De même qu'un secouriste bénévole sait panser quelques plaies bien identifiées et saura reconnaître les blessures qui dépassent sa compétence et nécessitent un appel à plus compétent que lui, de même le spécialiste du domaine qui va devoir assurer la partie statistique est capable de mettre convenablement en oeuvre quelques méthodes d'usage courant dans le domaine et sait reconnaître les problèmes qui dépassent sa panoplie d'outils. Je ne voudrais pas qu'on donne à cette appellation d'utilisateur occasionnel une connotation négative. Au contraire, la statistique est une connaissance supplémentaire par rapport à une spécialité qu'il possède. Dans une équipe de biologistes, il apporte en plus de ses connaissances en biologie une compétence limitée mais réelle en statistique, comme d'autres le font pour d'autres technologies par exemple la microscopie électronique. Cette vision se concrétise dans les formations que dispense l'université de Montpellier II par l'ouverture d'un D.E.S.S. (Diplôme d'études supérieures spécialisées) qui accueille des biologistes et se donne pour objectif de les rendre autonomes dans l'utilisation des méthodes courantes de la statistique ;
- Le troisième niveau est constitué des professionnels de la statistique. Certains ont choisi d'appliquer leurs connaissances à l'analyse effective de problèmes de nature statistique. Ils reçoivent les problèmes comme le médecin omnipraticien reçoit les malades, cherchant à trouver la solution la mieux adaptée. D'autres se consacrent à la création et à la justification de nouveaux outils d'analyse statistique. De même que le biochimiste invente de nouvelles molécules, solutions futures pour l'omnipraticien, ils créent de nouvelles méthodes destinées à élargir les possibilités d'intervention du statisticien consultant. Les deux orientations sont nécessaires et complémentaires : seul celui qui dispose d'une large panoplie de méthodes statistiques et des connaissances nécessaires pour suivre leur évolution dans les revues spécialisées peut prétendre exercer le métier d'analyste statisticien. Le chercheur en statistique quant à lui explore des champs nouveaux, soit parce que l'analyste statisticien l'a alerté sur l'insuffisance des connaissances actuelles soit parce qu'une démarche d'élargissement systématique des domaines connus l'y a conduit.

Je n'exclus pas qu'un même individu puisse exercer successivement ou simultanément ces deux activités que sont la pratique et la recherche. Il sera alors tout à fait capable d'introduire dans la pratique les acquis récents de

la recherche et de motiver des recherches nouvelles sur des thèmes que sa pratique nécessitera. Il me paraît cependant nécessaire d'identifier les deux activités car les institutions et les organisations dans lesquelles les statisticiens exercent leur métier ne les confondent pas.

Restent à situer par rapport à ce découpage successif qu'on peut qualifier d'horizontal ce que nous appelons biométrie, psychométrie, économétrie, pharmacométrie, etc... Pour moi, ces catégorisations correspondent à un découpage transversal donc vertical (mais pas à une partition) reposant sur les problématiques et les méthodes statistiques prioritairement utilisées dans les domaines d'applications visés. Ce découpage ne concerne pas le niveau de la culture générale. Aux autres niveaux, il revient à focaliser les enseignements sur des chapitres statistiques que les domaines d'applications désignent : courbes de survie et de croissance par exemple en biométrie ; enquêtes et sondages en sociométrie, etc...

Etant membre d'un laboratoire de Biométrie, je peux préciser mon point de vue pour ce domaine. Un étudiant bien formé à la statistique pourra devenir un bon biométricien s'il accepte de s'imprégner suffisamment du vocabulaire et des problématiques des biologistes qui l'entourent. Il lui faudra de la volonté, du temps et l'acquisition de certaines connaissances biologiques. Ce long apprentissage est nécessaire pour assurer que l'étudiant sera capable d'acquérir une compréhension suffisante des problèmes posés pour définir une stratégie d'analyse statistique vraiment adaptée.

Je constate que pour ce qui concerne le D.E.A. (diplôme d'études approfondies) enseigné à l'université de Montpellier II, et du doctorat qui le suit, un certain nombre d'étudiants arrivent avec un bagage de mathématiques et de statistiques mathématiques et s'en vont quatre ou cinq ans plus tard en véritables biométriciens. D'autres étudiants nous arrivent avec une formation d'ingénieurs agronomes. C'est pour la France au moins une formation biologique qui réserve une part importante aux mathématiques. Au prix de quelques efforts pour s'adapter à une présentation plus abstraite des mathématiques que celle à laquelle ils ont été habitués, ces étudiants réussissent parfaitement dans ce D.E.A. J'avoue en contrepartie, que je n'ai jamais su transformer en biométricien un biologiste que sa formation passée ou sa propre volonté empêchent de se plonger dans des mathématiques un peu abstraites.

2 La culture générale

Si nous sommes persuadés de l'importance que revêt pour nos concitoyens l'acquisition de quelques acquis statistiques nous devons nous sentir responsables de leur en faciliter l'accès. La première chose à faire est de définir les connaissances que nous voudrions transmettre puis de choisir la méthode pédagogique adaptée à cette transmission. Nous nous adressons ici à des enfants du primaire ou du secondaire. Toute formalisation mathématique sera un obstacle à la compréhension. Il me semble pourtant possible de faire découvrir la notion de fluctuation au sein d'une population et de donner quelques outils de description de cette fluctuation. Ceci fait nous pourrions passer à la notion d'échantillon et de fluctuation d'échantillonnage et enfin à celle d'indépendance.

Il est nécessaire par exemple d'aider les élèves à reconnaître des situations dans lesquelles on parle de :

- X : la variable aléatoire ;
- x : une valeur de cette variable ;
- μ : la moyenne de cette variable dans la population ;
- M : la variable moyenne calculée sur les échantillons ;
- m : la valeur de M pour un échantillon donné.

Des expériences simples et des jeux paraissent aptes à sensibiliser à ces notions, comme ils permettront de découvrir le phénomène de dépendance. Je me range donc à ce niveau dans le camp de ceux qui souhaiteraient enseigner la statistique comme discipline d'observation et d'expérimentation.

En second lieu et ultérieurement lorsque les élèves auront acquis une certaine aisance à manipuler les mathématiques, on pourra commencer à formaliser ces notions. S'engager dans cette voie ne veut pas dire qu'on va passer son temps à résoudre des problèmes paradoxaux de combinatoire liés à la formule (à proscrire) du nombre des cas favorables sur le nombre des cas possibles dans les espaces de probabilité finis et équiprobables. Il est plus important de faire correspondre aux distributions de fréquences observées des distributions qui serviront de modèles de références, et d'étudier leurs propriétés : c'est le moment d'introduire la formalisation des probabilités. Pour les petites valeurs de n la loi binomiale peut être construite comme une utilisation exemplaire de l'indépendance. Pour les moins avertis des élèves on se contentera de vérifier que la moyenne de la proportion des succès est p et la variance $\frac{p(1-p)}{n}$. Pour eux on admettra qu'il en est ainsi pour tout n . Un commentaire est alors possible sur les sondages, leur force et leur danger. Dès que les élèves disposent de la manipulation de l'intégrale, la loi normale peut être introduite. L'expérimentation permettra encore de vérifier qu'elle est une limite pour la loi binomiale. Sans entrer dans la technique précise des tests et de l'estimation, la nécessité de rapprocher les modèles et les observations peut être commentée.

Une telle approche est-elle utopique ? Non puisque plusieurs pays proposent à des élèves du primaire et du secondaire sous la forme de concours d'affiches statistiques des démarches de collecte, d'organisation, de représentations graphiques de données qui correspondent tout à fait à l'orientation développée dans ce paragraphe. Nos collègues japonais nous avaient montré le résultat d'un tel concours lorsque l'IIS s'est réuni à Tokyo en 1987. Depuis 1989, un tel concours existe en France : d'abord limité à l'une des 22 régions, il en concerne 10 cette année. J'espère que de telles expériences ne se limitent pas à ces deux que j'ai eu la chance de connaître. Un autre moyen de rendre expérimental l'enseignement de la statistique est d'utiliser les possibilités de simulation qu'offre les micro-ordinateurs. L'ordinateur prenant à sa charge les calculs fastidieux, il est possible de soumettre l'élève à des situations particulières fournissant des éclairages variés sur un problème donné : évolution d'un comportement lorsque n augmente ; sensibilité d'une caractéristique aux variations d'une donnée, etc... Dans cet esprit, les projets de didacticiels statistiques doivent être jugés aux possibilités d'apprentissage par expérimentation qu'ils accordent à leurs utilisateurs.

3 L'utilisation occasionnelle de la statistique.

Il m'arrive assez souvent d'être invité à lire des propositions d'articles soumis à des revues traitant d'un domaine scientifique autre que la statistique. Le texte comportant des développements statistiques, les membres du comité de rédaction demandent à un expert statisticien son point de vue. Le plus souvent ces articles suivent le plan suivant :

1. La première partie est une mauvaise présentation des bases mathématiques d'une méthode bien connue des statisticiens. Ils n'apprendront donc rien à la lecture de cet article et l'utilisateur occasionnel non plus qui devra de toute façon se référer à un livre ou à un article mieux écrit.
2. La seconde partie est une application de la méthode à des données caricaturales de la réalité. Les conclusions sont donc sinon sans importance au moins insignifiantes dans le domaine considéré. En fait l'auteur voulait montrer qu'il est réellement capable de mettre la méthode en oeuvre, ce qui est un plus pour lui par rapport aux chercheurs de son domaine, mais pas un plus pour la connaissance du domaine.
3. La troisième partie décrit en détail un programme spécifiquement écrit pour ce travail dans un langage et sur un ordinateur plus ou moins exotiques ce qui le rend pratiquement intransportable.

Je ne ris pas de ces articles car je suis persuadé que les statisticiens sont collectivement responsables de leur existence. Ne valorisant pas les applications dans la carrière des statisticiens professionnels nous condamnons les non-statisticiens à se débrouiller par eux-mêmes. Certains en sont ravis qui au nom d'une pluridisciplinarité mal comprise, croient nécessaire de maîtriser par eux-mêmes l'ensemble des technologies concourant à la solution de leurs problèmes. Le risque est double : par manque de formation adéquate, il est d'abord de n'apporter au domaine scientifique concerné que des méthodes statistiques simples peut-être insuffisantes pour le problème réel et peut-être archaïques. Il est ensuite, sous prétexte de servir de pont entre le monde de la statistique et celui du domaine, de s'installer dans une position qui ne sera reconnue ni par les statisticiens qui jugeront le niveau de compétence statistique insuffisant ni par les spécialistes du domaine qui ne peuvent reconnaître pour l'un des leurs quelqu'un qui ne fait plus d'apports fondamentaux à leur domaine mais seulement des apports technologiques. L'introduction d'une technologie nouvelle dans un domaine qui ne l'utilise pas encore est un gain si elle permet d'atteindre un résultat original ou au moins d'affiner des résultats connus et mieux encore d'ouvrir une problématique qui semblait jusqu'alors inaccessible. De plus, si l'auteur est persuadé de l'intérêt de sa proposition on doit s'attendre à ce qu'il aide ses lecteurs à trouver les moyens de la mise en oeuvre de la méthode qu'il préconise.

Dans cet esprit notre enseignement devrait conduire les utilisateurs occasionnels à proposer des articles dont le plan serait le suivant :

1. D'abord une présentation des objectifs et des problématiques pour lesquelles la méthode statistique a été définie et les règles de lecture des résultats qu'elle fournit. Pour les bases mathématiques l'article renvoie à une référence bien choisie.

2. Ensuite une présentation des logiciels existants permettant la mise en oeuvre de la méthode avec leurs avantages et leurs inconvénients : matériels nécessaires ; limitations concernant les données ; mode de présentation des résultats ; insertion dans des ensembles plus grands ; précautions d'utilisation ; exemples comparés.
3. Une application réelle où l'on voit un apport significatif de la méthode dans le domaine considéré.

Nous pourrions préparer nos étudiants, futurs utilisateurs occasionnels de la statistique, à écrire de tels articles si les cours que nous leur donnons avaient des objectifs mieux définis. Si l'idée directrice est bien que pour ces étudiants la statistique est et sera une technologie, nos objectifs doivent être de les rendre aptes :

- à reconnaître s'ils possèdent ou non dans la panoplie des méthodes qu'ils connaissent une méthode utile pour résoudre le problème auquel ils sont confrontés. Dans la négative, ils trouveront naturels d'aller consulter un spécialiste ou d'apprendre plus de statistique avant de retourner à la résolution de leur problème ;
- à comprendre de manière précise les résultats des méthodes dont ils disposent ;
- à accéder aux logiciels nécessaires à la mise en oeuvre de ces méthodes.

Pour atteindre ces objectifs, l'enseignement doit être organisé autour de situations de référence qui permettront de décrire en détail les conditions requises pour l'application de la méthode : la nature des variables, les conditions de distributions à respecter, les conditions d'échantillonnage, etc... L'analyse effective de données correspondant à ces situations est nécessaire pour l'apprentissage des règles de lecture des résultats (indicateurs, graphiques) et pour sensibiliser les étudiants aux pièges inhérents à la méthode. Enfin l'enseignement comprend nécessairement l'apprentissage d'un ou des logiciels permettant la mise en oeuvre de la méthode. Bien sûr cet aspect informatique n'est possible que lorsque les logiciels existent. Leur écriture est évoquée dans le paragraphe suivant comme l'une des tâches normales du statisticien professionnel.

A enseigner à ces étudiants, futurs utilisateurs occasionnels de la statistique, j'ai peu à peu découvert le besoin d'introduire dans mon cours quelques réflexions sur la démarche scientifique elle-même. Sans prétendre à un enseignement poussé en épistémologie, la présentation des méthodes statistiques permet de poser quelques questions sur la nature des objets manipulés et sur le but de ces manipulations. Sommes-nous en train de manipuler des données ou des modèles ? Qu'avons-nous perdu de nos données pour gagner de pouvoir les décrire par tel résumé statistique ? Qu'avons-nous dû ajouter aux données ? Explorons-nous un corpus de données sans idées préconçues ou sommes-nous en train de l'utiliser pour mettre une hypothèse à l'épreuve ? Qu'avons-nous contrôlé dans notre procédure d'observation ? Plus que tout autre discipline, la statistique est sensible à ces questions et c'est pour cela qu'elles doivent être intégrées à l'enseignement.

4 La profession de statisticien.

4.1 Le praticien.

Faire métier de statisticien praticien ne se réduit pas à appliquer ponctuellement une méthode connue à un problème particulier. C'est au contraire être capable de construire une stratégie de recueil et d'analyse de données susceptible de fournir une réponse adaptée à n'importe quel problème. Le statisticien praticien doit donc être un homme à l'esprit ouvert, prêt à faire l'effort de comprendre la problématique du domaine pour lequel on vient le consulter. Dans cette démarche, il est modeste : ses interlocuteurs savent sur ce domaine plus de choses qu'il n'en sait et ce n'est pas à lui, statisticien, de juger la pertinence de leurs interrogations.

Sa formation doit l'avoir rendu maître d'une large panoplie de méthodes couvrant des champs variés de la statistique mais aussi, et peut-être plus encore, apte à trouver et à comprendre dans les livres et les revues les méthodes qu'on ne lui a pas enseignées. Cette tâche suppose qu'il ait le même langage que le statisticien théoricien. Enfin, il connaît bien sûr les logiciels de statistiques mais sa formation informatique ne se limite pas à eux : il a une bonne compétence dans les logiciels de manipulations de données, dans les logiciels graphiques, dans les systèmes experts et il possède le langage nécessaire à des écritures rapides de procédures originales. Un tel statisticien existe-t-il ? Pour ce que je connais, essentiellement en France, il existe peu dans les entreprises privées qui se satisfont, trop à mon goût, de recourir à des statisticiens occasionnels. La situation serait différente si les entreprises étaient persuadées qu'elles tireraient un bénéfice d'un emploi plus sophistiqué de la technologie statistique. Nous sommes les seuls à pouvoir les en convaincre. Nous avancerions dans ce sens si nous défendions avec plus d'ardeur qu'on ne devient pas statisticien praticien par échec à une carrière de statisticien théoricien ou par apprentissage sur le tas à partir d'une formation dans une autre discipline. La situation est plus favorable dans les établissements publics qui recrutent par des concours après annonce d'un profil de recrutement. Si le profil fait référence à la statistique, il y aura inévitablement un ou plusieurs statisticiens dans le jury qui pourront défendre le niveau requis. Reste un problème pratiquement toujours absent des recrutements : celui de la déontologie. La pratique de nombreuses professions (médecine, droit, comptabilité) n'est consentie en France qu'à des personnes dont la compétence a été reconnue par l'obtention d'un diplôme et qui se sont engagées plus ou moins explicitement à n'utiliser leurs connaissances que pour des buts et sous des conditions reconnues par un code de déontologie. L'image de la statistique souffre nettement dans le public de l'absence de telles références dès lors qu'on s'éloigne des administrations et services de l'état.

La situation est particulièrement insatisfaisante dans les domaines qui font les grands titres des journaux quotidiens : politique, économie, société. La publication partielle des résultats, les commentaires orientés et parfois contradictoires de commentateurs qu'on doit croire "autorisés" puisqu'ils s'expriment dans les journaux écrits et parlés entretiennent l'idée que la statistique n'est pas une science rigoureuse et fiable. Certes les procédures de sondage sont contrôlées par une commission officielle mais la publicité faite aux commentaires de tout ou partie des résultats occulte souvent une éventuelle remarque de la commission sur la procédure suivie. Les statisticiens praticiens ne peuvent pas rester muet devant une telle situation qui nuit à une bonne image de leur profession.

4.2 Le théoricien.

Son objectif n'est pas d'analyser des données mais de fournir de nouvelles méthodes et plus encore les outils mathématiques et informatiques nécessaires à la construction de nouvelles méthodes, à leur comparaison avec les méthodes existantes, à l'étude de leurs propriétés.

Il a une connaissance exhaustive de tout ce qui s'est écrit sur un domaine particulier de la statistique. Il est capable d'en faire un inventaire systématique mais aussi de montrer les limites des acquis actuels. Il se sent investi de la responsabilité de repousser ces limites. Pour cela, il va devoir jeter des ponts soit entre différents domaines de la statistique soit vers des domaines des mathématiques ou de l'informatique aptes à fournir les outils nécessaires à l'élargissement des acquis du moment.

Ainsi installé temporairement ou définitivement en position de médiateur entre une classe de problèmes statistiques et un potentiel mathématique et informatique, il devra garder en tête les motivations de son travail. Ce n'est pas parce qu'il a dû investir dans la géométrie différentielle ou les systèmes experts pour résoudre les problèmes statistiques qui le motivaient qu'il doit se comporter comme si toute la statistique se limitait à ces domaines. Ce n'est pas parce qu'il a dû devenir un spécialiste de topologie pour formuler correctement un problème de convergence qu'il doit se comporter comme si seuls les spécialistes de topologie pouvaient prétendre au titre de statisticien. Son travail est utile à la statistique, tant que le statisticien théoricien garde la conscience de son rôle de médiation. L'inévitable question concernant cette catégorie de statisticiens est celle de leur rapport à la pratique statistique, de leur confrontation à des données réelles. Peut-on prétendre au titre de statisticien sans s'être jamais affronté à de véritables données ? Ma réponse instinctive est non : pas de statisticiens qui n'affrontent pas au moins épisodiquement des ensembles réels de données. Dans la réalité, nous savons trop le temps et les difficultés à surmonter pour dominer certains domaines mathématiques ou informatiques pour ne pas trouver d'excuses à ceux qui s'y étant plongés y ont trouvé les solutions aux problèmes statistiques qui les motivaient mais ne trouvent plus celui de revenir vers la pratique. Ceci me conduit à atténuer ma réponse instinctive : pas de statisticiens qui oublient que la statistique est faite pour être appliquée.

Les lignes précédentes de ce paragraphe associent toujours l'informatique et les mathématiques. C'est bien sûr volontaire. Ce choix a pour but d'insister sur le fait que du point de vue de la pratique statistique, celui qui justifie mathématiquement une nouvelle méthode et celui qui écrit le logiciel qui permet sa mise en oeuvre font des apports comparables. Ni l'un ni l'autre ne traitent des données réelles ; l'un et l'autre fournissent des outils que d'autres utiliseront. Ils font donc un travail théorique, spéculatif : le programme comme le théorème sont des abstractions de situations qui se concrétiseront de différentes manières. La conséquence de ce point de vue est que la production de logiciels statistiques relève sans l'ombre d'un doute de la responsabilité des statisticiens. Seul un statisticien compétent peut définir un logiciel assez général pour faire face à des situations concrètes multiples. Mais alors la traduction de la généralité dans le programme demande des outils appropriés si bien que le spécialiste de la statistique informatique sera conduit à investir des champs de l'informatique comme son collègue spécialiste de statistique mathématique investit ceux des mathématiques.

4.3 La formation

L'organisation de ce chapitre qui présente la pratique et la théorie comme deux modes d'expression de la même profession de statisticien a pour corollaire une vision unique de la formation à cette profession. Une fois formé, l'individu exercera sa profession en tant que praticien ou en tant que théoricien. Peut-être alternera-t-il dans le temps de sa vie active, peut-être même le poste qu'il occupera le conduira-t-il à une concomitance de ces deux activités. Mais la formation est unique ; c'est celle d'un statisticien.

Ceci dit, le réalisme, c'est-à-dire le nombre d'années disponibles pour la formation, va entraîner quelques accommodations avec cette position de principe. Il reste qu'il me paraît souhaitable que la séparation entre le praticien et le théoricien, statisticien informaticien ou statisticien mathématicien, ne vienne que tardivement après un cursus commun fournissant les bases communes aux différents aspects de la profession. En France, cette différenciation me paraît relever de l'année de D.E.A. (Diplôme d'Etudes Approfondies) ; il correspond à peu près au Master Degres des universités nord-américaines.

5 Conclusion.

J'aurais été mal compris, si un lecteur pensait que les paragraphes précédents décrivent une situation que je donne comme modèle. Ils ne sont même pas complètement des objectifs que je me fixe. Ils sont des interrogations auxquelles je fais face après plusieurs années d'enseignement et d'observation sur la manière dont vit la statistique autour de moi. L'organisateur de cette séance m'a donné la possibilité d'exprimer mon point de vue et de le confronter à celui d'autres orateurs. Je l'en remercie. J'attends de cette confrontation et de nos discussions de mieux remplir mon rôle d'enseignant de statistique.

Une question surgit d'ailleurs sur ces derniers mots "enseignant de statistiques". Qui doit enseigner la statistique ? La réponse à cette question mérite une attention très particulière lorsqu'on pense aux utilisateurs occasionnels. J'ai suggéré un enseignement partant de situations de référence accompagné de traitement de données représentatives de ces situations et très peu ambitieux du point de vue de la formalisation mathématique. Ce manque d'ambition concerne l'enseignement pas l'enseignant qui ne pourra faire sentir l'essentiel sans l'écrire que s'il a beaucoup de recul par rapport à ce qu'il enseigne. Il doit de plus s'appuyer sur des situations de référence. D'où viendraient-elles sinon de sa propre expérience. Le professionnel de la statistique que j'ai baptisé "praticien" remplit ces conditions.

Paradoxalement, ce "praticien" lui aura été formé pendant de longues années en compagnie du "théoricien". Ses maîtres seront majoritairement des théoriciens et minotairement des praticiens qui l'initieront lui et les théoriciens en formation à la pratique statistique en les associant à des études réelles et des consultations. De ce compagnonnage de formation, praticiens et théoriciens garderont un langage commun qui au delà de leur spécialisation ultérieure leur permettra les échanges constants que nécessite l'exercice de leur profession.

Références bibliographiques

- Adcock Chris, Conrad Simon (1991) - Professionnal training ? The Royal Statistical Society. News and Notes. Volume 17 n° 7 p. 1-2.
- Barrabba V.P. (1990) - What Does the Future Hold for Statisticians ? The American Statistician. Vol. 44 n° 2 p. 125-128.
- Dallal E.D. (1990) - Statistical Computing Packages : Dare we abandon their teaching to others ? The American Statistician. Vol. 44, n° 4 p. 265-266.
- Finney D.J. (1989) - Is the Statistician still necessary ? Biom. Praxim 29, p. 135-146.
- Lievesley Denise (1990) - Enough Statisticians. The Royal Statistical Society. News and Notes. Volume 17 n° 1 p. 1-2.
- Morison Jim (1991) - Who should teach Statistics. The Royal Statistical Society. News and Notes. Volume 17 n° 8. p. 3.
- Noether G. (Organisateur) 1989) - Les programmes nationaux pour l'enseignement pré-universitaire de la statistique : création et évaluation. Thème 13. Actes de la 47e Session de l'I.I.S. Livraison 2. p. 183-216.
- Ruberg S.J., Mason R.L. (1988) - Increasing Public Awareness of Statistics as a Science and a Profession. Starting in the High Schools. The American Statistician. Vol 42. n° 3, p. 167-170.
- Winterbottom Alam (1990) - Teaching in Industry. The Royal Statistical Society. News and Notes. Volume 17. N° 3. p. 1-2. Zahn D.A., Boroto D.R. (1989) - Promoting Statistics : on Becoming Valued and Utilized. The American Statistician. Vol 43 n° 2, p. 71-72.

D.E.A. BIOSTATISTIQUE

Etablissement habilité :

MONTPELLIER II

Etablissements co-habités :

ENSA.M
MONTPELLIER I

Objectif : Préparation aux métiers de la recherche aussi bien en statistique que dans les applications agronomiques, médicales et pharmaceutiques de la statistique.

ETUDIANTS CONCERNES

- Titulaires des maîtrises de Mathématiques
- Titulaires des maîtrises des Sciences Médicales et Biologiques ayant suivi des unités de valeurs en Mathématiques, Informatique et Epidémiologie
- Elèves des ENSA ayant choisi une spécialisation en Mathématiques Appliquées

ORGANISATION

Le D.E.A. comporte :

- Un tronc commun :*
- Statistique inférentielle
 - Statistique exploratoire
 - Modélisation des phénomènes biologiques
 - Outils informatiques du statisticien

- trois options :*
- Statistique et santé
 - Statistique et agronomie
 - Statistique mathématique

Les étudiants suivent un séminaire commun et un séminaire lié à leur option. Au troisième trimestre, ils effectuent un stage d'initiation à la recherche.

LABORATOIRES D'ACCUEIL.

Laboratoire de Probabilités et Statistique (Montpellier II)
Chaire de Mathématiques appliquées et d'informatique (ENSA Montpellier)
Biostatistiques, Informatique et Santé Publique (Montpellier I)
Physicochimie, Formulation et Valorisation du Médicament (Montpellier I)
Laboratoire de Biométrie (Inra Montpellier)
Laboratoire de Biométrie (Inra Avignon)
Analyse Statistique et Modélisation des processus écologiques (C.E.F.E., CNRS, Montpellier)

Débouchés :

- Recherche publique et parapublique :
CNRS, INSERM, INRA, CIRAD, ORSTOM,...
- Enseignements supérieurs :
universités, grandes écoles (ENSA, ENITA...)
- Laboratoires privés :
recherche pharmaceutique, agro-alimentaire...

Responsable : Y. ESCOUFIER
Département des Sciences Mathématiques
(Case 051)
Laboratoire de Probabilités et Statistique
Université Montpellier II
Place E. Bataillon
34095 MONTPELLIER CEDEX 5

Date limite du dépôt du dossier de pré-inscription : 30 juin

UNIVERSITE MONTPELLIER II - SCIENCES ET TECHNIQUES DU LANGUEDOC

D.E.S.S.

METHODES STATISTIQUES DES INDUSTRIES AGRONOMIQUES AGROALIMENTAIRES ET PHARMACEUTIQUES

Objectif : *Rendre des spécialistes des Sciences de la Vie ou de la Terre autonomes pour la mise en oeuvre des méthodes statistiques utilisées dans les industries relevant de leur domaine.*

ETUDIANTS CONCERNES

Titulaires de maîtrises de Sciences de la Vie ou de la Terre et diplômes équivalents

ORGANISATION

Une formation générale :

- Les outils informatiques du Statisticien
- Méthodes de base en Biométrie

Une formation axée sur la pratique

- Analyse des Données psychosensorielles
- Planification d'expérience et essais thérapeutiques
- Analyse d'enregistrements biologiques
- Contrôle de qualité
- Ajustement et prévision de phénomènes évolutifs
- Echantillonnage et sondages

Un stage donnant lieu à la rédaction d'un mémoire complète la scolarité.

LABORATOIRES PARTICIPANT A LA FORMATION

Laboratoire de Probabilités et Statistique (Montpellier II)
CIRAD (Montpellier)
Institut des produits de la Vigne (INRA Montpellier)
CNFDCI (Montpellier)
ORSTOM (Montpellier)
Centre de Recherche Sanofi (Montpellier)

Débouchés :

Industries
Centres de Recherche
Instituts Coopératifs
Services de contrôle

Responsable : Y. ESCOUFIER

*Renseignements : A. DELCAMP
Département des Sciences Mathématiques (Case 051)
Laboratoire de Probabilités et Statistique
Université Montpellier II
Place E. Bataillon
34095 MONTPELLIER CEDEX 5*

*Date limite de dépôt des dossiers de candidatures 15 septembre 1991
(Lettres de motivation et curriculum vitae)*

Début des enseignements 7 octobre 1991