



IREM de Lyon
Institut de recherche sur l'enseignement des mathématiques
Université Lyon 1 – Académie de Lyon

Compte rendu du stage ATSM “Probabilités et statistique”

Lyon, 25 août – 3 septembre 2010

Faisant appel à des formateurs dans tout le réseau des IREM, l'IREM de Lyon a organisé, du 25 août au 3 septembre 2010, une formation intensive au profit de dix professeurs tunisiens membres de l'Association tunisienne des sciences mathématiques (ATSM), sur les thèmes des probabilités et de la statistique. Ce stage s'inscrit dans une tradition de coopération entre l'ATSM, l'APMEP et les IREM.

1 Compte rendu général

Organisation

Les dates du stage ont été choisies sans tenir compte du Ramadan, ce qui a conduit à des conditions de travail pas très faciles. Par rapport aux habitudes françaises, les journées ont été concentrées et décalées vers le matin mais cette adaptation n'a pas été tout à fait suffisante. Pour une prochaine session, la question serait réglée en évitant simplement la période de jeûne.

L'organisation générale (transport, logement, bourses de stage) était parfaitement prise en charge par l'association Égide. Les quelques problèmes pratiques (passage de la douane, perception des bourses, voyage du retour à l'aéroport) qui sont survenus pendant le stage ne leur sont en rien imputables et ont été tous résolus, notamment grâce à la présence précieuse de Régis Goiffon (IREM).

Compte rendu synthétique

Ce paragraphe synthétise les réactions des collègues stagiaires et des formateurs, telles qu'elles ont été recueillies par différentes conversations et un questionnaire. Tous les impétrants se sont déclarés satisfaits : les stagiaires ont jugé le stage intéressant, certains étant vraiment enthousiastes ; du côté des formateurs, ils ont unanimement apprécié l'écoute, l'attention portée par les stagiaires, leur attitude active, leurs nombreuses questions et les échanges qui en ont découlé.

Les deux idées qui ressortent comme ayant été les plus intéressantes sont celles de *modélisation* et de *simulation*. Ces notions sont depuis quelques années inscrites aux programmes de collège et lycée français. Comme elles sont nouvelles pour la plupart des enseignants, elles forment le cœur des efforts de formation en probabilités en France, à travers les documents pédagogiques, les stages et les consignes des inspecteurs. Il n'est donc pas étonnant que les intervenants aient insisté sur ces points et que les stagiaires y aient été particulièrement sensibles.

Vu les questionnaires remplis par les stagiaires, l'équilibre général du programme, son niveau et l'ordre de traitement ont été bien adaptés aux demandes. Par exemple, le fait que les activités “clés en main” soient arrivées en toute fin de stage a permis d'une part d'éviter les questions théoriques sur les probabilités, déjà réglées, et d'apporter un regard didactique distancié sur les activités.

Quelques points du programme ont attiré des appréciations plus contrastées des stagiaires : la quantité de théorie, jugée inutile par certains et très pertinente par d'autres,

et la quantité de séances sur ordinateur, jugée convenable par certains et insuffisante par d'autres. Sans doute le temps passé à étudier les programmes français a été un peu trop long : c'est un sujet pour lequel il y a eu des répétitions. Pour autant, les redites ont été minimales ; de plus, quand le même thème (par exemple, l'échantillonnage) était abordé par plusieurs intervenants, leurs points de vue et leurs traitements (théorique, par exercices, sur machine) étaient assez différents pour justifier une nouvelle occurrence.

Parmi les ajustements à envisager pour une prochaine session, on peut retenir :

- allonger un peu la partie théorique pour la rendre moins dense, par exemple en ajoutant des exercices,
- proposer un peu plus de séances sur ordinateur : peut-être une séance sur tableur, en amont des activités spécifiques aux probabilités ;
- deux idées intéressantes ont été mentionnées : proposer un projet informatique en autonomie (autour du week-end) ; proposer une analyse de séquence filmée ; pour ce dernier point, le groupe "Collège" de l'IREM de Lyon disposera peut-être d'une telle séquence d'introduction aux probabilités l'an prochain ;
- la question des séries statistiques à deux variables, omises dans ce stage, a été posée.

Le bilan du stage est, pour résumer, très satisfaisant pour les stagiaires, les intervenants et donc pour les organisateurs.

Intervenants

Ce stage est d'évidence une manifestation du *réseau* des IREM. L'IREM de Lyon a assuré la logistique du stage proprement dit (programme, salles, ordinateurs, photocopies) mais c'est tout le réseau qui a fourni le contenu scientifique : enseignants et enseignants-chercheurs de sept IREM (Aix-Marseille, Clermont-Ferrand, Franche-Comté, Lyon, Montpellier, Paris 7, Toulouse) ont animé les différentes journées, la cohérence de l'ensemble étant assurée par le comité scientifique du stage et par la concertation. Les intervenants sont des formateurs aguerris habitués à travailler avec le public des enseignants de leur académie. Enfin, le stage a été clôturé par le président du Comité scientifique des IREM, un statisticien particulièrement investi dans les questions d'enseignement des probabilités et de la statistique.

Voici la liste des intervenants (la mention CII désigne les membres de la commission inter-IREM Statistique et probabilités) :

- Brigitte Chaput, professeur à l'ENFA d'Auzeville Tolosane (École nationale de formation agronomique, pour les professeurs du secteur agricole), responsable de la CII,
- Annette Corpart, professeur au lycée de Thiers, CII,
- Jean-Claude Girard, formateur à l'IUFM de Lyon retraité, CII, responsable du programme scientifique du stage,
- Michel Henry, professeur agrégé retraité, CII,
- Nelly Lassalle, professeur au lycée de Clermont-Ferrand, CII,
- Jacqueline Mac Aleese, maître de conférences à l'université Paris 7,
- Anne Perrut, maître de conférences à l'université Lyon 1,
- Jean-Pierre Raoult, professeur à l'université de Marne-la-Vallée retraité, statisticien, président du Comité scientifique des IREM, CII,
- Hubert Raymondau, professeur au lycée agricole de Carpentras, CII,
- Brigitte Sotura, professeur au lycée de Vanves,
- Claudine Vergne, professeur au lycée retraité, CII.

Plus de la moitié des interventions ont été effectuées en binômes. Ce fait a été très apprécié des stagiaires, notamment pour les binômes universitaire-secondaire. Le nombre élevé des intervenants a aussi été apprécié pour la variété des points de vue qu'il permet.

Programme

25/8	AP	Bases théoriques des probabilités (6 h)
26/8	BS & JMA	Les probabilités au collège et au lycée (6 h)
27/8	BS & JMA	Les probabilités au collège et au lycée - suite (6 h)
28/8		Travail libre des stagiaires
29/8	RG & AP	Visite de Lyon (après-midi)
30/8	AP	Bases théoriques des probabilités - suite (3 h)
30/8	BC & CV	Échantillonnage, tests statistiques (3 h)
31/8	BC & CV	Échantillonnage, tests statistiques - suite (6 h)
1/9	HR	Outils de description en statistiques, versant informatique (6 h)
2/9	JCG	Statistique et formation du citoyen (3 h)
	MH	Regard historique et didactique sur les probabilités (3 h)
3/9	AC & NL	Activités “clés en main” sur les probabilités (5 h)
3/9	JPR	Clôture du stage (1 h)

Documents

Les documents distribués aux stagiaires se trouvent sur le site de l’IREM de Lyon : <http://math.univ-lyon1.fr/irem/spip.php?rubrique77>.

2 Compte rendu détaillé

a) Intervention d’Anne Perrut

Elle a essentiellement consisté en un mini-cours de probabilités (voir le texte de support –65 pages– en annexe, dont la partie sur les statistiques n’a pas été abordée).

Dans ce domaine, les connaissances des stagiaires remontaient à leur formation initiale à l’université et présentaient un caractère très théorique. Il est significatif que pour tous, les applications numériques, même pour des problèmes simples, ne leur semblent pas nécessaires. La plupart des stagiaires étant professeurs de lycée, ils avaient en arrivant l’habitude des exercices standards sur la loi binomiale mais ils ne connaissaient pas d’outils statistiques, et peu de modélisation proba.-stat.

L’intervention a été principalement axée sur la modélisation. Au début, ce sujet a semblé inutile aux stagiaires qui objectaient : “en classe, on met toutes les hypothèses dans l’énoncé, donc on ne les discute pas.” Au fur et à mesure, ils ont mesuré que la modélisation était intéressante (ce qui témoigne de l’esprit ouvert avec lequel ils sont arrivés) et justifiée (préciser le modèle probabiliste permet d’expliquer les calculs proprement), si bien qu’ils se sont montrés de plus en plus demandeurs.

Ainsi, les questions ont porté progressivement sur la modélisation : l’utilisation pratique des lois, la signification pratique des paramètres... L’intervenante espère avoir répondu de manière satisfaisante. Malheureusement, le temps a manqué pour fouiller la modélisation statistique autant qu’il aurait été souhaité mais le stage n’était pas fini après cette partie ! Clairement, si la notion de modélisation probabiliste était loin de leurs préoccupations, la problématique statistique était aux antipodes ! Les programmes tunisiens sont en décalage par rapport aux programmes français car la statistique n’a pas encore fait son apparition. L’intervenante, pour sa part, ne peut pas s’empêcher de lier les deux...

Quelques simulations sur Excel ont été proposées. Les stagiaires ne les ont pas beaucoup aimées. On l’imputera à leur relatif manque d’habileté avec l’outil informatique : c’est toujours douloureux au début. Apparemment, leur sensibilité, leur formation et leurs habitudes ne les rapprochent pas de la bureautique ou de la programmation.

b) Intervention de Jacqueline Mac Aleese et Brigitte Sotura

Jeudi 26 août au matin (2 h 30 + 1 h 30)

1. Réflexions sur l'évolution des programmes d'enseignement de la statistique et des probabilités au collège et au lycée d'enseignement général et comparaison avec les programmes tunisiens : différentes approches de la notion de probabilité (laplacienne, fréquentiste...), place de l'expérimentation, de la simulation, de la modélisation, initiation à la démarche de la statistique inférentielle.

Documents 1a et 1b (voir annexe).

2. À partir d'une activité réalisée en classe de seconde (d'après une activité de Guy Brousseau, quelle est la composition de la bouteille opaque contenant 5 boules de deux couleurs différentes, noires et blanches ?), description effective avec la réalisation matérielle, illustration sur tableur des travaux faits par les élèves et développements possibles :
 - exploitation des données expérimentales obtenues avec les élèves dans la classe avec les quatre bouteilles transparentes dont on connaît la composition (notée p) (1N4B, 2N3B, 3N2B, 4N1B correspondant à 4 valeurs du paramètre p soit 0.2, 0.4, 0.6, 0.8),
 - puis de données simulées,
 - notions d'échantillon, tailles d'échantillon, comparaisons d'échantillons,
 - observation des fluctuations des fréquences, de la stabilisation vers la proportion de boules noires,
 - représentations graphiques des distribution de fréquences (diagrammes en bâtons et en boîte), etc.

Jeudi 26 août, après-midi (salle informatique) (2 h 30)

3. Initiation au tableur (Open office 4) pour simuler des expériences aléatoires et réaliser des échantillons de taille différentes. Travail individualisé.

Vendredi 27 août, matin (3 h)

4. Retour sur la situation (la bouteille) présentée la veille : formalisation théorique et modélisation dans le cadre probabiliste, en insistant sur la différenciation entre ce qui est aléatoire, ce qui est certain, ce qui est connu et ce qui ne l'est pas, avant l'expérience et après l'expérience (en s'appuyant sur bouteilles transparentes de composition p connue et la bouteille opaque de composition p inconnue).
5. Intervalles de fluctuation d'une variable aléatoire (fréquence), intervalles au pluriel pour un même choix de niveau explicités à partir d'exemples sur tableur par BS : lois binomiales de paramètres $n = 10$ et $p = 0.2$ ou 0.6 (lois exactes et pas expérimentales!).
6. Intervalles de confiance pour le paramètre p : toujours à partir du travail sur les quatre diagrammes en boîtes simultanés de l'expérience de la bouteille, début de construction d'abaques pour des tailles d'échantillons différentes, puis passage aux abaques non expérimentales mais exactes en utilisant les lois exactes.
7. Passage des intervalles de fluctuations aux intervalles de confiance à l'aide d'abaques en insistant sur d'une part sur le lien quasi-fonctionnel entre taille de l'échantillon, largeur de l'intervalle et niveau de confiance (après avoir énoncé sans preuve l'inégalité de Bienaymé-Chebyshev) et d'autre part sur le caractère aléatoire connu ou non, certain connu ou non, des objets manipulés.

8. Estimation d'une proportion et prise de décision à partir d'un échantillon : toujours sur l'exemple de la bouteille opaque, et élargissement au cas d'un lancer de pièce.

Documents : Abaques.

Vendredi 27 août, après-midi (salle informatique) (2 h 30)

9. Suite du travail individuel sur tableur (Open office 4) : travail moins guidé sur le papier que la veille, en partant de l'exemple suivant :

Simuler un tirage dans une urne :

- (a) Simuler le tirage d'une boule dans une urne contenant des boules noires et des boules blanches dans les proportions respectives p et $1 - p$ (p compris entre 0 et 1). (Définir p et utiliser la formule : `=si(alea()<p; 1; 0.)`)
- (b) Simuler un tirage avec remise de 100 boules (échantillon de taille 100) et observer les fluctuations de la fréquence de boules noires.
- (c) Simuler 1000 échantillons de taille 100. Représenter la distribution des 1000 fréquences observées. Faire varier p . Observer.

Certains stagiaires sont allés beaucoup plus loin dans cet exercice que ce qui était demandé, d'autres ont cherché d'autres exercices de la feuille (il y en a 13).

c) Intervention de Brigitte Chaput et Claudine Vergne

Thèmes abordés : échantillonnage, intervalle de fluctuation, intervalle de confiance, tests d'hypothèse. Elles n'ont pas pu approfondir les tests autant qu'elles l'auraient souhaité, faute de temps. Le public s'est montré très volontaire, très demandeur : beaucoup de participation et de questions. Il a manqué un peu de temps pour mettre en œuvre ce qui a été présenté, en particulier avec les élèves (mais cette question a été largement abordée par Annette Corpart et Nelly Lassalle vendredi), et pour illustrer nos propos avec une séance sur ordinateur ou calculatrice (mais Hubert Raymondaut a pris en compte cette dimension le lendemain).

Le travail sur l'échantillonnage a donné lieu à une réflexion sur l'expression "tirage au hasard" et au problème de l'équiprobabilité ou non des tirages selon la modélisation adoptée. Le lien a été fait avec les actuels programmes de seconde générale français. Enfin le traitement d'un exemple a permis de définir le vocabulaire et de décrire une distribution d'échantillonnage simple. L'influence de la taille des échantillons sur les distributions d'échantillonnage a été mise en évidence, cela a permis d'introduire la loi des grands nombre et le théorème limite central qui conduit dans certaines conditions à approcher une loi binomiale par une loi normale.

La journée du mardi a commencé par l'introduction des intervalles de fluctuation : définition conformément aux nouveaux programmes de seconde générale français, application à un exemple et explication de la formule $[p - 1/\sqrt{n}, p + 1/\sqrt{n}]$ du programme. Les intervalles de confiance ont ensuite été présentés ainsi que le contexte de leur utilisation pour l'estimation.

Les tests d'hypothèse ont été abordé par le traitement d'un exemple qui a permis d'introduire le vocabulaire et la démarche des tests.

d) Intervention d'Hubert Raymondaut

L'objectif était de mettre en œuvre des outils de la description statistique pour exploiter les résultats d'une expérience aléatoire simulant un contrôle de qualité sur un lot de semences devant contenir au plus 30 % de graines d'*Agropyrum repens*.

La matinée a été occupée par la simulation avec des bouteilles en plastique opaques, la saisie des données dans un tableau, la création de fichiers contenant les données récapitulatives pour l'ensemble des stagiaires, effectifs et fréquences, et la création des fichiers au format texte csv pour l'importation dans le logiciel R.

L'après-midi est consacrée à quelques points théoriques sur branches et feuilles, boîtes à pattes et histogrammes et à la difficulté de les réaliser avec un tableur. Quelques éléments sur les caractéristiques des distributions observées : symétrie, dispersion, position. Les limitations du tableur ont justifié de passer sur le logiciel R : importation des données avec la bonne structure et mise en œuvre de quelques méthodes de description des séries à une variable. Avec tout cela, les stagiaires ont pu estimer que les lots contenaient une proportion de grains d'*Agropyrum repens* autour de 30 %, conformément aux spécifications.

L'attrait principal a été exprimé pour l'utilisation de R, ses possibilités graphiques et la perspective de pouvoir faire des analyses statistiques plus poussées (inférence). Les stagiaires ont aussi découvert les branches et feuilles et boîtes à pattes avec curiosité. Beaucoup de questions sur :

- Le fond : à quoi cela sert-il de décrire une série et comment décrire une distribution observée ?
- Pourquoi passer d'une structure de données à une autre ? (Réponse : il peut arriver qu'une structure de données soit adaptée au terrain pour la saisie mais devienne inexploitable dans un autre logiciel (exemple : trois colonnes de données correspondant à des variétés de semences différentes : OK pour Excel, pas OK pour R).)
- Que faire avec R ? Mise en évidence des limites des tableurs et des possibilités de R.

e) Intervention de Jean-Claude Girard

À partir de quelques exemples de (mauvais) traitements statistiques pris dans la presse, l'objectif était d'illustrer ce qu'on entend par "développer chez les élèves la capacité à comprendre, analyser et critiquer l'information chiffrée" et "développer une culture numérique" au niveau secondaire. Les notions suivantes ont été abordées :

- Les différents types de variables statistiques, la difficulté pour les choisir et les mesurer. Exemple d'analyse de résultats obtenus pour une variable quantitative (parce que ce que c'est plus facile à mesurer) conduisant à des conclusions abusives à propos d'une autre (qualitative).
- Importance de l'ensemble de référence dans les calculs de pourcentages ou dans leur comparaison. Exemple de manipulation dans le choix du référentiel (ou dans la définition de la variable) pour obtenir le pourcentage souhaité.
- Passage d'une réalité à un modèle mathématique : pertinence du modèle, domaine de validité. Exemple de modélisation sur un nombre très restreint de données et de prévisions très éloignées de l'intervalle d'observation.
- Opposition : observation - expérimentation.
- "Défense et illustration" des sondages malgré, ou à cause, de la marge d'erreur et des fluctuations d'échantillonnage.

f) Intervention de Michel Henry

Le thème de l'après-midi était un regard épistémologique et didactique sur l'enseignement des probabilités au lycée. L'objectif inavoué était de montrer que c'était une mission difficile et d'expliquer pourquoi. (On peut s'en tenir au quotient des cas favorables par les cas possibles mais ça ne va pas bien loin et ça n'est pas intéressant.) Il faut réussir à faire saisir la dualité du concept, entre définition a priori liée à un modèle et compréhension a posteriori via l'observation de fréquences.

La problématique était donc la recherche du (d'un ?) sens de la notion de probabilité. Le plan était en trois parties.

1. Historique (cf. l'article récent de *Repères-IREM* joint) : pourquoi la notion classique de probabilité a émergé. La partie s'est finie sur une critique didactique de la notion classique (quels problèmes didactiques elle pose).
2. Approche fréquentiste des programmes de 1990, qui n'allait pas jusqu'à une définition des probabilités à partir des fréquences qu'ont formulée certains auteurs. Cela a donné l'occasion de parler de la loi des grands nombres, notamment la première version formulée par Bernoulli : elle a été énoncée avec ses mots puis traduite en langage moderne. Une critique didactique a suivi ces considérations d'ordre historique, à l'appui de la conviction de Michel Henry que cette approche aboutit à des impasses.
3. Approche fondée sur la modélisation : la probabilité est un concept mathématique (donc un objet idéalisé qui doit représenter les attentes que l'on a des réalisations d'un événement). Cela permet de sortir du débat entre objectivistes et subjectivistes. C'est la démarche intégrée dans les programmes actuels.

Dans cette version, l'expérimentation passe par la simulation. L'exemple de franc-carreau a été présenté en plusieurs versions : a) version historique de Buffon (aiguille sur des lattes), qui utilise là la première loi continue de l'histoire (NB : dans l'idée naïve favorables/possibles, les probabilités sont des rationnels, l'intervention de pi dans une probabilité devrait paraître paradoxale.) ; b) version avec une aiguille sur un réseau carré : on aboutit à une intégrale triple bien retorse que seule la simulation semble permettre d'évaluer. (Question posée : pourquoi un rapport d'aires serait une probabilité? Réponse : *) Ca marche! **) Voir le protocole de discrétisation dans la thèse de Cileda Coutinho.) (Question aux stagiaires : comment imaginer qu'une représentation cartésienne du problème donne une probabilité par un calcul de rapport d'aires? Quid si on change de représentation, par exemple si on passe en polaires? Voir la foisonnante littérature sur "l'aiguille de Buffon".)

Cette partie 3) s'est conclue en présentant les programmes de troisième, seconde (déjà enseignés) et de première (cf. JO du 28/8/2010) et en montrant la cohérence entre ces programmes.

NB : l'enseignement des probabilités semblait une gageure lors de son introduction. Des progrès significatifs ont été faits depuis la fin des années 90 : on a une façon cohérente de présenter la notion et de lui donner du sens.

Commentaires : Les stagiaires ont offert une grande écoute et ont très envie d'en savoir plus. Il est à nouveau apparu que leur formation aux probabilités est très classique. (Ex. de question posée : pourquoi ne définit-on pas la probabilité comme une application des parties de Ω ou de la tribu vers $[0, 1]$? La réponse semble avoir été un peu évasive : *) parce qu'on est sorti de la période structuraliste ; **) pour rechercher un sens avant de plonger dans le formalisme.)

g) Intervention d'Annette Corpart et Nelly Lassalle

La session a été consacrée à la présentation de 8 activités "clés en main" sur les probabilités : 5 étaient des activités d'introduction, 3 relevaient du niveau lycée (première et terminale).

Pour la première activité, les animatrices ont hésité à faire manipuler [jeter des dés, en l'espèce], ce qui a conduit les stagiaires à poser des questions peu naturelles sur

le déroulement, qui ne seraient pas apparues s'ils avaient manipulé tout de suite. Le démarrage a donc été un peu lent. Néanmoins, les choses se sont arrangées et toutes les activités prévues ont été traitées.

Les stagiaires se sont montrés très intéressés par les activités liées à des applications industrielles : cartes de contrôle prises dans une vraie entreprise ; tests d'hypothèse.

La question de fabriquer du hasard et des échantillons a été naturellement soulevée, ce qui était justement le thème de la dernière activité : dans une feuille tapissée de cercles, approximer la moyenne des rayons en tirant 5 cercles "au hasard". Comment faire le tirage ? On peut par exemple procéder en pointant du doigt avec les yeux fermés ; en numérotant et en tirant 5 nombres au hasard par ailleurs ; en répartissant les cercles en strates selon leur rayon et en tirant "au hasard" dans chaque strate.

Au bilan, les stagiaires étaient contents, ce qui s'est montré en particulier par le grand nombre de questions posées, ils repartent avec des choses immédiatement exploitables.

On peut discuter ce point, et d'ailleurs un des stagiaires n'est pas d'accord, mais il a semblé opportun de mettre cette session à la fin du stage : d'une part, cela a évité certaines questions ("qu'est-ce qu'un test ?") qui auraient fait double emploi avec les exposés théoriques ; d'autre part, cela permet un regard critique sur les activités.

Les questions posées étaient dans l'ensemble plus techniques que celles des professeurs français en stage – bon signe à la fin d'un tel stage ! Les questions d'ordre didactique étaient, en revanche, très comparables à celles des Français.

h) Intervention de Jean-Pierre Raoult

La clôture du stage par Jean-Pierre Raoult a duré un peu moins d'une heure. Il a commencé par se présenter et expliquer pourquoi il est si intéressé par la problématique "Statistique et enseignement", titre d'une nouvelle revue dont il est co-rédacteur en chef.

L'enseignement des probabilités et de la statistique présente, pour les professeurs de mathématiques, des difficultés spécifiques, liées à la fois au fait que nombre d'entre eux n'ont pas été familiarisés avec ces disciplines durant leurs études et au fait que, comme d'autres branches du savoir scientifique, elles ne prennent leur sens qu'en liaison avec les phénomènes qu'elle permettent de formaliser pour en permettre un traitement mathématique et donc ne sauraient être enseignées sans la prise en compte de leurs applications. Il est heureux que, vers la fin de ce stage, ait été situé l'exposé de Michel Henry rendant compte des variations, au cours des années, dans les conceptions du traitement de ces difficultés au travers des programmes de l'enseignement en France. Cet exposé a mis en lumière les évolutions dans le choix de la "définition" du concept de probabilité, qui pose un problème de nature assez différente de la plupart des définitions figurant dans un cours de mathématiques, dans la mesure où on il s'agit d'un outil que l'on est amené à manipuler (dans le cadre scolaire ou extra-scolaire) sans qu'il ait été déjà indispensable d'en donner une définition formalisée. Plutôt que de "définition", il serait plus opportun de parler des différents points de vue sur "l'introduction" des probabilités qui peuvent apparaître au fil des programmes successifs.

Voici un exemple de trouble causé par ces questions : les programmes de 2000 recommandaient que les élèves tiennent un "cahier de statistique" dans lequel ils auraient rendu compte, à l'instar des cahiers de TP en sciences expérimentales, des activités statistiques menées dans l'année. Cette recommandation a fait long feu, ce qui témoigne de la difficulté à intégrer les statistiques dans la pratique quotidienne de la classe.

On arrive néanmoins, plusieurs années après l'introduction de ces matières, à un accord sur les buts de l'enseignement des probabilités et de la statistique :

- prise de conscience et maîtrise de la variabilité,

- précautions à prendre pour le choix d'un modèle,
- importance du contexte d'information pour une modélisation,
- place faite à l'usage de données réelles.

Sur ce dernier point, Jean-Pierre Raoult présente quelques considérations sur le choix et l'usage de ces données. Il évoque la difficulté qu'il y a, pour les enseignants, à aller recourir à des données "brutes" (ainsi, quoique les programmes recommandent l'emploi de données de l'INSEE, celles-ci doivent être manipulées avec précautions car leur contexte est souvent complexe –et certains manuels les présentent avec insuffisamment de précautions). Il cite des sites où des données sont fournies avec une "médiation" utile pour l'enseignant : **Statistix** (<http://www.statistix.fr/>) ; les fiches pédagogiques de l'INED (Institut national d'études démographiques). Il met en évidence l'intérêt d'actions pluridisciplinaires, dans le cadre desquelles les données sont en rapport avec les programmes d'autres disciplines et il attire l'attention des stagiaires sur les problèmes déontologiques que peut poser la collecte de données individuelles sur les élèves et critique à cet égard le programme canadien *Recensement à l'école/census at school*.

Les problèmes d'enseignement portent donc sur les fluctuations de la hiérarchie et l'ordre adopté pour le traitement des différents éléments. Jean-Pierre Raoult n'a pas de dogme personnel et insiste, en revanche, sur la nécessité de la cohérence de la démarche adoptée. Il fait une analogie avec l'introduction de l'exponentielle comme solution de $y' = y$ qui a fait l'objet d'une polémique intense il y a quelques années en France : dans ce que doit finalement avoir assimilé l'élève, peu importe le statut (définition ou théorème) des différentes caractérisations car ce qui compte, c'est de les avoir toutes.

Cet enseignement est, toutefois, un enseignement scientifique : il est donc soumis aux mêmes exigences de rigueur que le reste du corpus, en particulier pour ce qui est du respect du vocabulaire. C'est d'autant plus délicat, en la matière, que les mots utilisés sont ceux du vocabulaire courant :

- *erreur* : une erreur en général, c'est une faute ; en pédagogie, c'est de dire des choses contradictoires ou hors de portée des élèves ; ici, c'est tout une conclusion, obtenue dans une inférence effectuée à partir de données, non conforme à la réalité sous-jacente et qui n'est pas totalement évitable mais dont la probabilité doit être maîtrisée ;
- *échantillon* : l'emploi du mot par certains intervenants a été un peu critiqué parce que trop restrictif, réservé à des situations pédagogiques/idéales loin des données réelles ;
- *estimation*, qui a un sens très précis en statistique, loin d'être un synonyme de "approximation" ou "évaluation".

En conclusion, Jean-Pierre Raoult évoque certaines opinions, parfois avancées en France, qui tendraient à faire enseigner les outils statistiques de préférence par les professeurs des matières qui en font usage : sciences économiques, biologie, physique, géographie, instruction civique... Il pense que, tout au contraire, seuls les enseignants de mathématiques peuvent avoir le recul nécessaire pour présenter le calcul des probabilités et la statistique, aux yeux des élèves, comme des disciplines au caractère scientifique affirmé, et non des "recueils de recettes", fondées sur un substrat mathématique qui en assure l'universalité. Encore faut-il que les professeurs de mathématiques collaborent avec leurs collègues des autres disciplines pour assurer la cohérence entre leurs usages respectifs des données et de leur traitement.

Jérôme Germoni, directeur de l'IREM de Lyon