

LA STATISTIQUE DANS L'ENSEIGNEMENT SECONDAIRE EN FRANCE

Jean-Pierre RAOULT¹

TITLE

Statistics in secondary education in France

RÉSUMÉ

L'enseignement de la statistique dans les collèges français (qui accueillent les élèves pour leurs sixième à neuvième années d'études) est assez sommaire et purement descriptif. La conception de cet enseignement dans les lycées (dixième, onzième et, pour la majorité des élèves, douzième années d'études) est très différente selon qu'il s'agit des filières dites « professionnelles », « technologiques » ou « générales ». Dans ces dernières, cet enseignement est largement orienté vers des objectifs de statistique inférentielle (intervalles de confiance, embryons de tests d'hypothèses) plutôt que vers la statistique descriptive (pas de séries bidimensionnelles). Il a beaucoup évolué ces dernières années et évoluera sans doute encore prochainement, par une intensification de l'usage des moyens de calcul et par la mise à disposition des professeurs de ressources qui leur sont indispensables en raison du manque de formation à la statistique de nombre d'entre eux, surtout s'agissant d'exemples d'applications scientifiques, technologiques ou relevant de la vie sociale et civique.

Mots-clés : curriculum statistique au collège (France), curriculum statistique au lycée professionnel (France), curriculum statistique au lycée général et technologique (France), statistique descriptive, inférence statistique.

ABSTRACT

The statistics curriculum in French “collèges” (where pupils attend their sixth to ninth years of schooling) is rather cursory and purely descriptive. The curriculum in the “lycées” (tenth, eleventh and, in most cases, twelfth years of schooling) varies considerably depending on whether the students choose the “professionnel”, “technologique” or “général” stream. In the latter, the curriculum is strongly oriented towards inferential statistics (confidence intervals, basic notions of hypothesis testing), rather than towards descriptive statistics (no bivariate data). It has changed significantly these past few years and will certainly continue to do so as greater emphasis is gradually placed on computing, and resources become more widely available for teachers who need to compensate their general lack of training in statistics, notably with respect to its application to science, technology, and social and civic life.

Keywords: statistics curriculum in French “collèges”, “lycées professionnels” (France), and “lycées généraux et technologiques”, descriptive statistics, statistical inference.

1 Introduction

L'enseignement secondaire français accueille les élèves au sortir des cinq années passées en école primaire ; la norme est donc que les enfants y accèdent à l'âge de 11 ans. Il se décompose en quatre années de **collège** (dénommées, chronologiquement, sixième, cinquième, quatrième et troisième) et trois années de **lycée** (seconde, première et terminale).

¹ Professeur des universités honoraire, Université Paris-Descartes, jpraoult@orange.fr

Pour l'ensemble de l'enseignement obligatoire unifié, c'est-à-dire les neuf années d'école puis de collège, a été institué en 2006 un *Socle commun de connaissances et de compétences* et donc les éléments de statistique qui y figurent doivent être considérés comme faisant désormais partie d'un fonds commun partagé pratiquement par tous les futurs citoyens de ce pays. En revanche, l'analyse du curriculum relatif à la statistique (et, simultanément, pour une part, au calcul des probabilités, car les programmes officiels les traitent en commun) dans les lycées doit tenir compte de l'existence de trois grands groupes de filières, dites **professionnelles**, **technologiques** et **générales**, entre lesquelles se répartissent les élèves qui poursuivent des études au sortir du collège (ce qui est le cas pour la plupart d'entre eux). Tous ces enseignements ont connu récemment des réformes importantes². En particulier celles-ci se sont échelonnées, au fil des trois années que comporte la scolarité au lycée, de 2009-2010 à 2011-2012 pour les filières professionnelles et de 2010-2011 à 2012-2013 pour les filières générales et technologiques (sauf pour certaines de ces dernières pour lesquelles la réforme a été retardée d'un an).

Cette étude se composera de quatre sections : après celle consacrée au collège, nous considérerons, s'agissant du lycée, tout d'abord la voie professionnelle, puis simultanément les voies générale et technologique ; en effet ces deux dernières voies sont de plus en plus souvent réunies dans les mêmes établissements, alors que les lycées professionnels gardent en général leur autonomie ; de plus l'orientation entre les lycées professionnels d'une part et les lycées généraux et technologiques d'autre part s'effectue à l'entrée en seconde, alors que les cursus généraux et les cursus technologiques ne se différencient qu'à l'entrée en première. On le verra, la conception de l'enseignement de la statistique est aussi très différente selon ces différentes voies ; il est donc utile de savoir combien de jeunes sont concernés dans chaque cas ; nous retiendrons ici les chiffres (arrondis au millier) d'effectifs en fin de collège (classes de troisième à la rentrée 2010) et en lycée (classes de seconde et classes terminales à la rentrée 2011). Ces chiffres sont extraits des publications de la DEPP (Direction de l'Évaluation, de la Prospective et de la Performance) du ministère de l'éducation nationale (DEPP 2012).

Secondes : 762 000, dont :
 professionnelles : 190 000
 (plus 63 000 entrant en préparation au Certificat d'Études Professionnelles) ;
 technologiques et générales : 509 000.

Terminales : 646 000, dont :
 professionnelles : 184 000 ;
 technologiques : 141 000 ;
 générales : 321 000.

Dans une dernière section, nous donnerons quelques indications sur la situation dans les classes implantées dans les lycées mais accueillant des étudiants de niveau Enseignement Supérieur (Sections de Techniciens Supérieurs, Classes Préparatoires aux Grandes Ecoles Scientifiques).

Nous nous attacherons à décrire dans cette étude la situation actuelle, tant dans les motivations affichées que dans leur concrétisation, de l'enseignement de la statistique (et des probabilités dans la mesure où ils sont reliés). Nous nous abstenons de faire un historique

² L'auteur remercie Philippe Dutarte, inspecteur pédagogique régional de mathématiques dans l'académie de Créteil, pour ses fructueuses remarques sur une première version de cet article.

J.-P. Raoult

de ces enseignements, ne serait-ce que parce qu'ils ont connu des évolutions successives à un rythme assez soutenu et que les suivre nous entraînerait assez loin ; le passé (en fait très récent) n'interviendra donc que quand il sera indispensable d'y faire référence. Pour une mise en perspective de cet état des lieux, on se reportera en particulier aux études de Michel Henry (Henry, 2011) et de Philippe Dutarte (Dutarte, 2010), le titre de cette dernière traduisant bien l'évolution, encore en cours, de l'état d'esprit avec lequel est abordé l'enseignement de la statistique : *Evolutions de la pratique statistique dans l'enseignement du second degré en France : du calcul statistique au développement de la pensée statistique*.

Pour consulter l'intégralité des programmes officiels utilisés ici, on se reportera au site EDUSCOL du ministère français de l'éducation nationale : <http://eduscol.education.fr>.

2 Collège

Avant de traiter de la place de l'enseignement de la statistique dans le *socle commun de connaissances et de compétences* et dans les programmes des classes de collège, il faut signaler le cas particulier d'une minorité d'élèves qui sont scolarisés dans les **Sections d'Enseignement Général et Professionnel Adapté** (SEGPA), chargées d'accueillir, selon des textes administratifs datant de 1996, *des élèves en grande difficulté scolaire ou sociale* ; selon une circulaire de 2006, ils y sont admis quand, à un niveau donné, ils *ne maîtrisent pas toutes les compétences et connaissances définies dans le socle commun*. Durant l'année scolaire 2010-2011, ces sections ont accueilli :

- **en sixième** : 20 000 élèves, sur un total de 795 000 élèves pour l'ensemble des collèges (à titre de référence, précisons que, onze ans plus tôt, le nombre des naissances en France avait été d'environ 800 000) ;
- **en troisième** : 26 000 élèves, sur un total de 767 000 élèves pour l'ensemble des collèges (précisons qu'en quatrième et troisième les enseignements fondamentaux ne constituent qu'une part du cursus de ces élèves, aux côtés de stages d'initiation à la vie professionnelle).

Les programmes de mathématiques spécifiques pour ces sections consistent donc largement en une actualisation des capacités que ces élèves ne sont pas parvenus à acquérir antérieurement, en insistant sur celles qui sont le plus indispensables dans la vie courante. Un accent particulier est mis sur la numération, la proportionnalité et les figures géométriques élémentaires. Le « document d'accompagnement » édité par le ministère de l'Éducation Nationale est très largement inspiré par les recherches en didactique des mathématiques pour l'enseignement élémentaire, avec une insistance sur la pédagogie à base de résolution de problèmes. Dans ce contexte, un point de vue statistique n'est jamais explicité dans les textes relatifs à ces sections, si ce n'est, dans le programme de quatrième, dans la rubrique *proportionnalité*, des items intitulés : *calculer une moyenne* et *lire, interpréter et construire des graphiques*. On peut s'interroger sur l'intérêt qu'il y aurait, s'agissant d'élèves qui, sauf exceptions, ne poursuivront pas leurs études au-delà de ces sections, à faire une place plus importante à l'étude critique de statistiques ; en effet, leur interprétation concrète pourrait être, pour certains, une source de motivations dont la présentation actuelle des programmes nous paraît assez démunie. Des études en ce sens, ainsi éventuellement que des expérimentations dans certaines classes, seraient sans doute opportunes.

Pour ce qui est des élèves de l'enseignement primaire et de l'immense majorité des élèves de collèges, le **socle commun de connaissances et de compétences** est, aux termes d'une loi de 2006, le *cadre de référence de la scolarité obligatoire*, ambitieusement qualifié de *ciment de la nation*. Vu la place occupée par les statistiques et leurs diverses formes de représentations dans les médias, on aimerait qu'une initiation à leur analyse critique soit une des composantes de ce « ciment ». Ce « socle » s'articule autour de « sept grandes compétences » ; c'est au sein de celle intitulée *Les principaux éléments de mathématiques et la culture scientifique et technologique* qu'on peut trouver une modeste exigence de « littératie³ statistique », dans un préambule qui affirme : *dans chacun des domaines que sont le calcul, la géométrie et la gestion des données, les mathématiques fournissent des outils pour agir, choisir et décider dans la vie quotidienne*. On y relève des items qui, dans le cadre de *l'organisation et la gestion de données*, sont intitulés : *les représentations usuelles (tableaux, diagrammes, graphiques) et les notions fondamentales de statistique descriptive (maximum, minimum, fréquence, moyenne)*. On y trouve aussi un point libellé : *les notions de chance ou de probabilité*.

Un facteur influant pour l'introduction de ces nouveautés fut la volonté de prendre en compte les conclusions des évaluations internationales PISA (*Programme for International Student Assessment - Programme International pour le Suivi des Elèves*) ; celles-ci avaient fait apparaître une moindre compétence, dans ces domaines de la prise en compte de données et du traitement de l'aléatoire, chez les jeunes Français de 15 ans que chez les jeunes de même âge dans la plupart des autres pays concernés par ces études de l'OCDE.

La manière dont ces intentions se sont traduites dans l'enseignement primaire n'influe que peu sur l'enseignement au collège. En effet les programmes de l'enseignement primaire actuellement en vigueur (introduits en 2008) sont sur ce point en retrait par rapport aux précédents, datés de 2002. Dans ces derniers on trouvait une sous-rubrique intitulée *Organisation et représentation de données numériques* comportant les recommandations suivantes : *Les situations qui conduisent à utiliser diverses représentations d'un ensemble de données (tableaux, graphiques, diagrammes) s'appuient sur des données effectives : enquêtes, mesurages en physique ou en biologie (exemple de l'évolution de la taille d'un enfant, d'un animal ou d'une plante), documents en géographie... Quelques exemples de phénomènes aléatoires peuvent être proposés dans la perspective de faire apparaître des régularités (par exemple, lancers d'une pièce ou d'un dé, lancers de deux dés dont on fait la somme)*. Cette ouverture semble avoir suscité plus de travaux didactiques sur l'initiation à l'aléatoire que sur les présentations des données. Dans la version actuelle (BOEN 2008) du programme de mathématiques pour l'ensemble (dénommé administrativement « Cycle 3 ») des trois dernières années de l'enseignement primaire (enfants de 8, 9 et 10 ans) figure une rubrique *Organisation et gestion de données* où l'on se contente de lister des outils de représentation des données (*tableaux, graphiques*), des compétences (*trier des données, les classer...*) et des occasions de mettre en œuvre ces compétences et d'utiliser ces outils (*résoudre des problèmes de la vie courante ou tirés d'autres enseignements*). L'observation de la pratique dans les classes conduit à penser que cette formulation moins précise amène les professeurs des écoles à concentrer leurs leçons sur des tableaux et graphiques déjà élaborés plutôt que d'apprendre

³ **Littératie** : terme d'usage récent, mais rapidement croissant, en français, parfois orthographié « littéracie » ou « litéracie », issu de l'anglais « literacy » lequel, désignant à l'origine le fait de savoir lire et écrire, a vu son sens se généraliser pour désigner la *capacité d'emploi, pour un public donné, d'un niveau de culture assigné comme objectif dans un domaine donné*.

J.-P. Raoult

aux élèves à introduire ces outils à l'occasion de problèmes : l'enseignement de l'outil prend le pas sur l'intérêt de son utilisation pour traiter certaines questions, notamment venant d'autres disciplines.

Il importe d'analyser comment les indications rudimentaires du « Socle commun » se sont traduites dans les rédactions des programmes de collèges mis en application de 2006-2007 à 2009-2010, ainsi que dans la documentation (documents d'accompagnement officiels, manuels, initiatives diverses) mise à la disposition des enseignants. Ceux-ci, souvent peu préparés par leurs propres études à enseigner la statistique et le calcul des probabilités, ont accueilli ces branches du programme avec une inquiétude qui les a fréquemment amenés à leur donner peu d'ampleur dans leur enseignement.

De la sixième à la quatrième, le maître-mot dans les intitulés des chapitres du programme comprenant des notions de statistique (terme lui-même absent) est **données** : en sixième et cinquième il s'agit d'*organisation et représentation de données* et en quatrième de *traitement des données*, qui se déclinent en « compétences » comme suit :

- **Sixième** : représentations usuelles, tableaux, diagrammes en bâtons, diagrammes circulaires ou semi-circulaires ;
- **Cinquième** : effectifs, fréquences, classes, tableaux de données, représentations graphiques de données ;
- **Quatrième** : moyennes pondérées, calculer la moyenne d'une série de données, créer un graphique à partir des données d'une feuille de calcul.

Le terme *Statistique* apparaît, comme titre de chapitre, en classe de **troisième**, avec le contenu suivant : *Caractéristiques de position. Approche de caractéristiques de dispersion.*

Pour chacune de ces années, les commentaires présentent un triptyque : *lire, utiliser et interpréter.*

On remarque que les listes de « compétences » sont purement opératoires, impliquant des activités aisées et peu mathématisées, souvent vues comme de simples mises en jeu d'opérations informatisées stéréotypées. C'est là ce qui souvent met les professeurs de mathématiques mal à l'aise. Tout un travail préliminaire devrait être favorisé, insistant à l'aide d'exemples sur ce qui figure « en amont » des données dans les procédures de recueil et sur les questionnements pertinents et s'interrogeant en particulier, même succinctement et empiriquement, sur le « pourquoi » de la variabilité traduite dans les données présentées ; ceci serait utile pour donner du sens à l'objectif *interpréter* du programme ; or ce n'est qu'en classe de troisième qu'on voit apparaître timidement une « approche » de caractéristiques de dispersion. Le « document d'accompagnement » *Organisation et représentation de données*, diffusé par l'inspection générale de mathématiques et se situant dans la ligne de l'intention affirmée, au titre du socle commun, de prendre en compte des « tâches complexes », vise à favoriser ce questionnement. De plus ce document a l'avantage d'aborder, ce qui est insuffisamment présent dans la lettre du programme, les interactions des outils statistiques avec d'autres branches du programme de mathématiques (initiation au calcul, proportionnalité). Mais il ne peut présenter qu'un nombre très limité d'exemples (tonnages de pêches, observations météorologiques...) et il faut regretter que les exercices proposés dans nombre de manuels s'en tiennent à des manipulations non replacées dans le contexte originel des données et qui donc apparaissent comme très répétitives à la fois aux élèves et à leurs professeurs.

On doit citer ici l'existence, dans le programme de la classe de troisième, d'une section intitulée *Notion de probabilité* pour laquelle les compétences attendues sont : *Comprendre et utiliser des notions élémentaires de probabilité* et *Calculer des probabilités dans des contextes familiers*. Mais rien dans les « commentaires » officiels sur ce programme n'incite à faire le lien avec les notions de statistique descriptive enseignées par ailleurs.

Les réticences que nous venons d'exprimer, fondées sur la modestie de l'ambition des programmes de mathématiques du collège quant à l'introduction des outils statistiques et à leur positionnement dans le contexte de l'aléatoire, auraient dû pouvoir être en partie contredites par l'existence des **thèmes de convergence**⁴ qui sont destinés à fournir la base d'activités pluridisciplinaires⁵ et visent à communiquer à l'élève *une première représentation globale et cohérente du monde dans lequel il vit*. Les activités autour de ces « thèmes », qui *n'introduisent pas de compétences exigibles et ne font pas l'objet d'un enseignement spécifique*, doivent faire l'objet d'accords avec enseignants ; rien ne permet donc d'affirmer qu'est bien exploité celui des six thèmes proposés qui a le numéro 4 et est intitulé : **Mode de pensée statistique dans le regard scientifique sur le monde**⁶.

La fiche descriptive de ce thème fait la part belle au lien avec l'aléatoire en écrivant : *le domaine de l'aléatoire et les démarches d'observation sont intimement liés à la pensée statistique* ; elle aborde opportunément les diverses sources de la variabilité dans des sections intitulées *Le recueil des données en grand nombre lors de la réalisation d'expériences et leur traitement* ou *le problème de la variabilité de la mesure*. Il y a là des pistes intéressantes ; mais seule une très faible fraction des élèves de collèges a été amenée à profiter réellement des « thèmes de convergence »⁷ et donc en particulier à bénéficier du surcroît d'initiation à la statistique que le thème correspondant peut conférer⁸.

3 Lycée : la voie professionnelle

Un quart des élèves entrant dans l'enseignement professionnel⁹ au sortir du collège ne se dirigent pas vers le baccalauréat mais vers le CAP (Certificat d'Etudes Professionnelles) dont la préparation s'effectue en deux ans. L'essentiel de l'enseignement est de type

⁴ Référence du texte de base de 2005 régissant les « thèmes de convergence » : <ftp://trf.education.gouv.fr/pub/edutel/bo/2005/hs5/annexe5.pdf>.

⁵ Voir des ressources, élaborées sous l'autorité du statisticien Alain-Jacques Valleron, membre de l'Académie des Sciences, dans : <http://eduscol.education.fr/cid46087/le-pole-scientifique-au-cycle-central-du-college-les-themes-de-convergence.html>

⁶ Une mise à jour, en 2008, a fait passer le thème qui nous intéresse du numéro 4 au numéro 1 et a modifié son titre en : **Importance du mode de pensée statistique dans le regard scientifique sur le monde** (cf. page 5 du document : http://media.education.gouv.fr/file/special_6/52/5/Programme_math_33525.pdf). Il est difficile de savoir si ces signaux ont favorisé l'utilisation effective de ce thème de convergence dans les classes.

⁷ Les textes officiels affirment : *Les thèmes de convergence se prêtent particulièrement bien à une évaluation soit dans la discipline soit dans le cadre d'une pluridisciplinarité concertée*. Mais rien n'est précisé sur ce mode d'évaluation ni sur sa prise en compte dans l'évaluation globale des élèves, au contraire des points figurant explicitement dans le « socle », qui ont fait l'objet de « livrets de compétences » individuels à renseigner par les professeurs. Cette situation a incontestablement nui à la mise en place de ces thèmes de convergence.

⁸ L'auteur remercie Pierre Campet, professeur à l'IUFM de Paris, pour les indications qu'il lui a fournies sur la pratique dans l'enseignement des mathématiques en collège.

⁹ L'auteur remercie Jean-Luc Pernet et François Moussavou, de la commission « Lycée Professionnel » du réseau des IREM (Instituts de Recherche sur l'Enseignement des Mathématiques) pour l'aide apportée dans la rédaction de cette section.

J.-P. Raoult

préprofessionnel mais il y figure depuis longtemps des bases de statistique descriptive et, depuis 2010, quelques rudiments de probabilités proches de ceux enseignés dans les secondes professionnelles, afin de faciliter d'éventuelles intégrations en première pour les meilleurs élèves après leur CAP.

La préparation au baccalauréat professionnel n'a été que très récemment alignée sur la trame de celles du baccalauréat général ou technologique, soit en trois années, alors que précédemment (pour les élèves entrant dans ces études jusqu'en 2008) les lycées professionnels offraient d'abord un cursus en deux ans conduisant au BEP (« Brevet d'Etudes Professionnelles »), seuls certains élèves poursuivant, à nouveau en deux ans, en vue du Baccalauréat Professionnel¹⁰.

Une caractéristique essentielle des études en lycée professionnel, en ce qui nous concerne ici, est que, dans les filières à orientation « industrielle » les enseignements de mathématiques et de sciences physiques et chimiques y sont regroupés (la question ne se posant évidemment pas pour celles à orientation « tertiaire »). Nombre d'enseignants y sont donc formés à l'ensemble de ces disciplines et les instructions officielles leur enjoignent à tous de *privilégier une démarche d'investigation (s'appuyant) sur un questionnement des élèves relatif au monde réel*. La structuration du travail des élèves se fait non seulement par une classification de nature conceptuelle (qui distingue, en mathématiques, *statistique et probabilités, algèbre et analyse et géométrie*) mais aussi par « thématiques » ; la liste des thématiques peut être révisable ; lors de la mise en place de cette nouvelle structure, en 2009, elle comprenait cinq grands sujets : *développement durable ; prévention, santé et sécurité ; évolution des sciences et techniques ; vie sociale et loisirs ; vie économique et professionnelle*. L'échelonnement du travail des élèves selon les années est moins rigide que dans les autres ordres d'enseignement et laisse plus de latitude aux enseignants (il est dit dans les instructions : *Par année de formation, l'enseignant choisit au moins deux thématiques dans des sujets différents*).

Dans ces classes les études de mathématiques et, selon les cas, de sciences physiques ou de disciplines du secteur tertiaire (management, commerce, économie...) sont largement orientées vers l'apprentissage raisonné d'outils pour la vie professionnelle et aussi vers une formation civique. Il n'est donc pas étonnant qu'une place privilégiée soit conférée à la statistique. Les programmes sont présentés en « modules », dont, à partir de la classe de première, l'emploi varie selon les dénominations des baccalauréats professionnels préparés (il y en a plus de 80). Dans ce cadre, les modules placés sous le vocable *Statistique et probabilités* interviennent quelle que soit cette dénomination. Si on peut inférer de l'ordre dans lequel sont présentés les items des programmes une vision de l'importance respective des grandes tranches des mathématiques abordées, on remarquera que (contrairement à ce qui est le cas pour les lycées généraux) la section *Statistique et probabilités* est placée, pour chaque année d'études, en tête (avant *Algèbre et analyse* et *Géométrie*) ; cette place a été délibérément souhaitée par l'Inspection générale de mathématiques afin d'inciter les enseignants à ne pas reléguer cet enseignement en fin d'année scolaire et dans la précipitation, comme cela se pratiquait parfois auparavant. Le point de vue statistique est aussi présent dans le détail des propositions faites pour les « thématiques » évoquées plus haut ; c'est ainsi qu'on y trouve *les enjeux de l'évolution démographique* au titre du *développement durable, jouer avec le hasard*

¹⁰ Pour ménager la transition en évitant des sorties du système scolaire sans diplôme outre que le Diplôme National du Brevet (passé en fin de collège, mais dont l'obtention n'est pas indispensable pour l'entrée au lycée), la fin de l'année de première permet encore d'obtenir le BEP, mais celui-ci, de diplôme terminal d'un cycle, est devenu un « diplôme intermédiaire ».

et croire un sondage au titre de la *vie sociale et loisirs*, ou encore contrôler la *qualité* au titre de la *vie économique et professionnelle*.

Plus précisément, le libellé des programmes met constamment en avant l'aspect expérimental, la fluctuation d'échantillonnage apparaissant comme centrale en seconde et en première et y constituant le fondement de la notion de probabilité (voir à ce sujet l'article de F. Moussavou, 2010) ; en classe de première est proposé, pour la fréquence en situation d'échantillonnage dans une population où on observe un caractère à deux modalités, un intervalle de fluctuation approché (comme en seconde de la voie générale et technologique : voir ci-dessous) ; en terminale sont étudiées les statistiques à deux variables, avec pour but l'usage d'une droite d'ajustement pour *interpoler et extrapoler*. A aucun niveau ces programmes n'abordent explicitement la statistique inférentielle quoique des activités du type *deviner la composition d'une urne à partir de tirages successifs* s'y rapportent clairement et que les enseignants qui le désirent puissent utiliser l'intervalle de fluctuation approché proposé pour fabriquer un intervalle de confiance.

Il est essentiel de remarquer aussi que la volonté de centrer l'enseignement moins sur des concepts que sur des actions (exploiter des données ; identifier, classer, hiérarchiser l'information ; interpréter un résultat statistique ; gérer des situations simples relevant des probabilités) n'est pas contrecarrée par la nécessité de s'adapter aux contraintes inhérentes aux examens de fin de scolarité. En effet, le baccalauréat professionnel est attribué, en ce qui concerne les épreuves de mathématiques et sciences physiques et chimiques, sur la base d'un contrôle en cours de formation.

Enfin une des particularités des études en lycée professionnel est qu'y figurent deux dispositifs à contenu non ciblé, dénommés « accompagnement personnalisé » et « enseignement général lié à la spécialité ». Des mathématiques peuvent être enseignées dans chacun de ces deux dispositifs et il est possible d'y faire de la statistique ou des probabilités ; certains enseignants le font par exemple dans le cadre d'un projet pluridisciplinaire nécessitant un traitement de données, et, au sein du réseau des Instituts de Recherche sur l'Enseignement des Mathématiques, la Commission « Lycée professionnel » a favorisé des travaux en ce sens.

L'auteur de ces lignes souhaite aussi faire part de son constat d'une disponibilité, vis-à-vis de l'enseignement de la statistique, globalement différente chez les enseignants en lycée professionnel de ce qu'elle est chez les professeurs des filières plus fondamentales. Déjà habitués qu'ils sont souvent, par leur formation, à la bidisciplinarité, ils ne sont pas enclins au même titre à s'interroger sur le statut « véritablement mathématique » de ce qu'on les invite à enseigner. La continuité entre l'observation chiffrée et son traitement par des outils numériques qui l'éclairent s'effectue chez eux avec plus de naturel que chez les professeurs en lycée général, voire technologique. De plus les enseignants observent que les élèves en voie professionnelle peuvent être rendus conscients du fait que les statistiques sont omniprésentes dans l'entreprise ; comme l'étude de celles-ci est relativement nouvelle dans le cursus des élèves, leur traitement est souvent abordé par eux avec nettement moins d'appréhension que d'autres parties du cours de mathématiques ; en effet ces dernières leur paraissent plus obscures et surtout plus « ancrées » dans les enseignements des classes précédentes, enseignements vis-à-vis desquels ils s'étaient souvent trouvés en situation d'échec.

4 Lycée : les voies technologique et générale

C'est en 2013 que la plupart des élèves passeront un baccalauréat rénové, suite à une importante réforme de structure des lycées induisant une refonte des programmes (Eduscol, 2012).

Au sein de l'ensemble des lycées technologiques et généraux, des filières ne se distinguent qu'à partir de la classe de première : dans le programme commun de mathématiques de seconde on trouve une section intitulée *Statistiques et probabilités*. Des sections de même intitulé (ou parfois *Probabilités et statistique*) figurent au sein des programmes de mathématiques de première et terminale dans les filières dont la liste suit (pour chacune desquelles nous donnons entre parenthèses le nombre d'inscrits en terminale à la rentrée 2011, sauf pour STAV, filière gérée non par le ministère de l'Education Nationale mais par celui de l'Agriculture).

Baccalauréats technologiques :

- **STI2D** : Sciences et Technologies Industrielles et du Développement Durable (32 000) ;
- **STL** : *Sciences et Technologies de Laboratoire* (7 000) (en Terminale, la statistique est renforcée pour les élèves choisissant la spécialité intitulée *Biotechnologies*) ;
- **STMG** : *Sciences et Technologies du Management et de la Gestion* (74 000, programme rénové entrant en vigueur au baccalauréat 2014 seulement) ;
- **ST2S** : Sciences et Technologies de la Santé et du Social (22 000) ;
- **STAV** : Sciences et Technologies de l'Agronomie et du Vivant.

Il existe aussi des baccalauréats technologiques sans enseignement de statistique : **STD2A** (*Sciences et Technologies du Design et des Arts Appliqués*), **TMD** (*Techniques de la Musique et de la Danse*), **TH** (*Technologique de l'Hôtellerie*) (l'accès en classe de première dans ces deux derniers nécessitant d'avoir suivi une seconde spécifique).

Baccalauréats généraux :

- **S** : *Scientifique* (160 000) ;
- **ES** : *Economique et Social* (106 000) ;
- **L** : *Littéraire* (52 000 ; la statistique apparaît uniquement à titre optionnel, au sein d'un enseignement facultatif de mathématiques au programme identique à celui de ES).

L'objectif de la formation en statistique dans toutes ces filières est fortement conditionné par le fait qu'elle figure toujours au sein des programmes de mathématiques, d'où il résulte que les programmes ont été rédigés par des commissions ministérielles composées de mathématiciens (inspecteurs, enseignants, quelques chercheurs) ; nombre d'entre eux n'ont pas eu eux-mêmes l'occasion d'enseigner la statistique, ni surtout de se confronter à ses usages. Il s'agit là d'une caractéristique historique française : si le calcul des probabilités y a été, et est toujours, une branche florissante de la recherche mathématique, la statistique s'y est longtemps développée loin de cette recherche, malgré des efforts de quelques mathématiciens de valeur (citons la création de l'Institut de Statistique de l'Université de Paris, par Emile Borel en 1922).

Il en résulte que, même si la statistique inférentielle avait déjà fait son apparition lors des avant-dernières rédactions de programmes (aux alentours de 2000), il apparaissait naturel aux auteurs des programmes actuels d'organiser ceux-ci par des démarches initiées plus par les fondements mathématiques (en particulier ceux des probabilités) que par les usages de la démarche statistique.

Un tel choix était d'ailleurs assez inévitable, vu la conception de l'enseignement des mathématiques qui prévaut chez les enseignants français : mettre en place une démarche de modélisation de problèmes concrets, rendre sensible aux élèves les fluctuations d'échantillonnage, ou encore passer des théorèmes asymptotiques du calcul des probabilités aux approximations qu'ils autorisent (mais avec toujours une marge d'erreur) sont des démarches intellectuelles auxquelles il leur est souvent difficile de s'adapter. Il importait donc que la conception des curricula favorise autant que possible une bonne coordination de ces démarches avec l'ensemble de l'enseignement des mathématiques, tout en initiant à leur originalité ; on pourrait même espérer que cette mise en jeu des mathématiques pour les besoins du calcul des probabilités et de la statistique joue en retour un rôle de justification de l'intérêt de certains autres pans du programme.

Il faut reconnaître que cet objectif ne peut pas être totalement atteint avec les nouveaux programmes car ceux-ci sont contraints par des réformes de structure du lycée, réduisant le nombre d'heures d'enseignement affectées en propre aux différentes matières, au profit d'activités transversales. Ceci est particulièrement flagrant en première scientifique, où la réduction à 4 heures par semaine de l'horaire de mathématiques a conduit à restreindre le programme ; il en est résulté un certain émiettement du programme, qui s'effectue au détriment des interactions entre les branches des mathématiques impliquées ; comme le conseil donné aux professeurs est (de même qu'en terminale) de consacrer environ un quart de l'activité aux probabilités et à la statistique, ceci leur a paru d'autant plus difficile à réaliser qu'ils y voient souvent une cause supplémentaire de déficit en formation dans les secteurs qui leur paraissent plus « fondamentaux » du programme.

Pourtant un atout essentiel pour l'initiation à l'aléatoire est l'usage de plus en plus intensif en classe des moyens de calcul. On relève que le parti adopté par les programmes précédents, qui était de consacrer la totalité de l'année de seconde à la simulation d'échantillonnages pour observer des fluctuations, sur un mode assez heuristique, n'a pas été conservé ; cependant, l'usage de telles simulations reste une ligne de force du cursus. L'autre ligne de force est, nous allons le voir, l'accoutumance progressive à des « prises de décision » (donc, en fait, à de la statistique inférentielle).

Pour préciser ces considérations, nous allons tout d'abord observer comment s'organise la formation en probabilités et statistique, au cours de ses trois années au lycée, d'un élève qui en terminale est en section générale scientifique (S).

Dès la seconde (donc de même que tous les autres élèves de lycée), il complète sa formation en statistique descriptive (médiane, quartiles, moyenne (mais pas écart-type)) et il observe, à partir de simulations d'échantillons de taille n , une version grossière d'un intervalle de fluctuation, au seuil de 95%, de la fréquence observée f en contexte d'échantillonnage avec une proportion p (cet intervalle est $\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$). Le programme est ici assez ambitieux dans son vocabulaire, puisqu'il parle de *probabilité que la fréquence soit dans cet intervalle*, alors que la notion de probabilité n'a été manipulée au collège que dans des cas bien plus élémentaires (mais le point de vue fréquentiste adopté alors facilite ce glissement) ; il

J.-P. Raoult

recommande aussi en préambule de *mobiliser les élèves à la notion d'intervalle de confiance* mais n'explique pas ensuite cette démarche ; il introduit enfin, à propos de l'intervalle de fluctuation, la *prise de décision à partir d'un échantillon*, ouvrant ainsi une petite fenêtre vers la problématique des tests d'hypothèse. Ce qui est remarquable, c'est donc que les objectifs des programmes des années suivantes sont ici en germe ; comme il est de règle en mathématiques, il est possible (au prix d'un certain effort de la part du professeur) d'introduire un vocabulaire précis, mais les justifications de l'usage qu'on en fait ne peuvent évidemment pas être données ; en particulier, en probabilités, la notion d'indépendance et le conditionnement ne sont explicités qu'en première. Mais le programme n'incite pas à considérer ces notions dans le cadre d'exemples concrets.

Le programme de première S innove peu par rapport à celui de seconde, sinon qu'on y introduit la variance et l'écart-type en statistique descriptive, avec la recommandation de faire pratiquer ces notions sur des données *réelles, riches et variées* ; en probabilités, c'est alors que l'on présente l'espérance mathématique et la variance des variables aléatoires discrètes, en particulier de la loi binomiale. Le seul acquis proprement statistique est alors de calculer à l'aide de la loi binomiale un intervalle de fluctuation de la fréquence plus précis que celui vu en seconde et d'introduire la notion de *différence significative* entre cette fréquence et une valeur conjecturée du paramètre de la loi binomiale (tout en précisant que *le vocabulaire des tests est hors programme*).

Ce n'est qu'en terminale qu'on introduit les lois de probabilité définies par des densités (uniforme, normale, exponentielle) et qu'on énonce, sans démonstration, le théorème de Moivre-Laplace pour les fréquences. Les nouveautés, en statistique, sont alors l'introduction de l'intervalle de fluctuation asymptotique, meilleur que celui vu en seconde, justifié par ce théorème, et l'intervalle de confiance qui lui est associé, toujours sans possibilité de démonstration. De plus on recommande maintenant : *cette partie se prête particulièrement à l'étude de problèmes issus d'autres disciplines*.

Notons que, à aucun moment au cours des trois années, n'ont été abordés, dans cette filière S, les outils bidimensionnels (tableau de contingence de deux caractères qualitatifs, covariance en calcul des probabilités, régression en statistique descriptive).

On voit bien qu'il s'agit, dans ses intentions, d'un programme de statistique mathématique, acheminant année après année vers une conception de plus en plus complète des solutions apportées aux problèmes de test ou d'estimation par intervalle, en se limitant au cas d'une fréquence théorique inconnue¹¹. Ce nouvel édifice est donc cohérent mais sa réalisation se heurte au fait que, simultanément, le programme d'analyse a été fort réduit et que donc nombre d'outils manquent pour l'étayer, en particulier s'agissant de l'étude des fonctions (limites, intégration). Cette situation aggrave chez nombre de professeurs de mathématiques leur réticence face à cette branche du programme.

Ce n'est que marginalement que les instructions accompagnant le programme incitent à sortir d'un cadre purement formel ; mais ceci est en partie compensé par le fait que des exemples utilisables, de diverses origines, sont par ailleurs fournis aux enseignants en particulier dans les « documents ressources » élaborés et diffusés par le ministère de l'Éducation Nationale (Eduscol – Ressources), pour lesquels un accent particulier a justement été mis sur les probabilités et la statistique ; pareil souci se rencontre aussi dans des travaux

¹¹ La version antérieure du programme de terminale allait jusqu'au test du chi carré de l'hypothèse d'équipartition mais ceci a été jugé trop difficile pour être conservé.

menés au sein des Instituts de Recherche sur l'Enseignement des Mathématiques (voir sur leur « portail » les rubriques des Commissions « Lycée » et « Statistique et Probabilités » : www.univ-irem.fr).

Le programme des autres séries du lycée général, ES (Economique et Sociale) ainsi que, en option, L (Littéraire), ne se distingue pas, pour ce qui est de la statistique, de celui de S (en probabilités il y a quelques retraits, en particulier la loi exponentielle en Terminale).

En séries STI2D (Sciences et Technologies Industrielles et du Développement Durable) et STL (Sciences et Technologies de Laboratoire), le programme a une trame analogue à celle des séries générales, mais un plus fort accent est mis sur les données réelles. Sont ainsi évoqués, à propos de la statistique descriptive en première, *la maîtrise statistique des procédés, le contrôle de qualité, la fiabilité, le développement durable*. En terminale la loi exponentielle est mise en situation par le taux de désintégration en radioactivité, l'introduction de la loi normale *s'appuie sur des exemples issus des autres disciplines* et on en introduit la simulation. Les intervalles de confiance sont utilisés ici non seulement pour estimer une proportion inconnue mais aussi pour tester l'égalité de deux proportions, ce qui est présenté comme *une occasion d'étudier les méthodes statistiques pratiquées dans les disciplines scientifiques et technologiques*. Une mention particulière doit être faite du programme de la spécialité « Biotechnologie » en Terminale où on présente de plus la statistique descriptive bidimensionnelle (méthode des moindres carrés).

En série STMG (Sciences et Technologies du Management et de la Gestion), l'enseignement de la statistique descriptive peut être relié à une portion de programme, intitulée *information chiffrée*. Comme dans les séries générales, l'aboutissement du programme est l'intervalle de confiance au niveau de confiance à 95% pour une proportion, mais ici c'est seulement l'intervalle « grossier » $\left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}} \right]$, où f est la fréquence observée, associé à l'intervalle de fluctuation vu dès la seconde, qui est introduit. En terminale figure un passage sur la statistique descriptive à deux variables (ajustement affine).

En série ST2S (Sciences et Technologies de la Santé et du Social), depuis plusieurs années (en effet les programmes actuels reprennent exactement ceux de 2006) l'accent est résolument mis sur l'usage des tableaux ou graphiques. On y considère donc dès la première *l'étude simultanée de deux caractères* et, en terminale, la totalité du programme de statistique est consacré aux *séries statistiques à deux variables* ; dans cette série la statistique inférentielle n'est pas du tout abordée.

Le programme de la série STAV (Sciences et Technologies de l'Agronomie et du Vivant) est fort modeste et axé sur la statistique descriptive. Il est recommandé de le traiter sous forme d'activités encadrées ; en première on considère des séries statistiques à une variable et en terminale des *tableaux de contingence*.

5 Formations post-baccalauréat

Quoique les élèves y aient le statut d'étudiant, il nous a semblé légitime d'inclure dans cette étude des informations sur les sections post-baccalauréat implantées dans des lycées ; en effet tant par leur mode de fonctionnement (élèves regroupés en classes où ils suivent tous les enseignements en commun, avec en principe un seul enseignant pour chaque matière) que par leur contenu (programmes nationaux très détaillés et contraignants) ou encore par le statut de

J.-P. Raoult

leurs professeurs, ces sections constituent un prolongement encore très scolaire des études secondaires, et donc on devrait s'attendre à ce qu'il y ait en particulier une certaine continuité dans le cursus en statistique.

Ces sections se situent aux deux extrémités de l'éventail des formations supérieures : soit études courtes (qualifiantes en deux ans), soit amorce d'études longues au plus haut niveau.

Les **sections de techniciens supérieurs**, qui préparent en deux ans au BTS (*Brevet de Technicien Supérieur*), présentent une large palette de spécialités. A visée nettement professionnalisante, elles pratiquent des programmes qui sont conçus en liaison avec des représentants des secteurs qui offrent des débouchés aux élèves. Dans de nombreux cas ces utilisateurs sont conscients des besoins de nature statistique dans leur activité (contrôle industriel de qualité, analyses en laboratoires, suivi d'indices économiques ou commerciaux...). Il n'est donc pas étonnant que, comme le signale P. Dutarte (2011), ce secteur de l'éducation en France ait été précurseur dans l'implantation d'un enseignement de la statistique largement piloté par ses applications. Mais il reste essentiellement assuré par des professeurs de mathématiques, éventuellement en liaison avec leurs collègues d'autres disciplines ; les IREM (en particulier leur Commission « Lycée technologique ») ont accompagné les efforts de ces enseignants pour pénétrer dans une « pensée statistique » : on consultera notamment avec profit la rubrique consacrée aux enseignements technologiques (pré ou post baccalauréat) de l'IREM de Paris-Nord (IREM Paris-Nord).

La situation est diamétralement opposée pour les classes préparatoires aux grandes écoles, dans leurs sections à dominantes de mathématiques ou de sciences physiques. Là, jusqu'à présent, le calcul des probabilités et la statistique n'ont eu aucune place dans les programmes. Une refonte de ces programmes va entrer en application, en 2013 pour les premières années et en 2014 pour les secondes années, refonte rendue indispensable par l'arrivée au niveau du baccalauréat en 2013 d'élèves ayant connu les programmes rénovés des lycées. En première année est prévue une introduction limitée de calcul des probabilités (se bornant au cas des ensembles finis d'éventualités) et pas de statistique ; les projets pour la seconde année ne sont pas encore connus mais il semble évident qu'on ne pourra pas parler de véritable « curriculum de statistique » dans ce type d'études.

Il est important d'en voir l'impact sur la vie scientifique et économique de la nation. C'est par ces sections que passent pour leur énorme majorité les responsables à venir de l'enseignement des mathématiques (inspecteurs, professeurs de ces mêmes classes) ainsi qu'une fraction importante des professeurs de lycée de plus haut niveau (ceux titulaires de « l'agrégation »). Le risque est donc important que la statistique continue de leur apparaître largement comme un « mal nécessaire » dans les programmes de mathématiques, mal relié à l'ensemble de la discipline et par rapport auquel leur recul restera souvent insuffisant. C'est aussi dans ces sections qu'acquière une solide formation en mathématiques les futurs « managers » français : ingénieurs et, pour une part financiers ; ce n'est que récemment (durant la dernière décennie) que les grandes écoles d'ingénieurs ont fait une place à la statistique dans les cursus de leurs élèves, d'où une situation paradoxale de « trou » de deux ans entre l'apprentissage des premiers rudiments lors de la préparation au baccalauréat et une reprise (souvent à partir de zéro car les souvenirs des élèves étaient estompés) en grande école. Cette situation est peu favorable au développement d'une pensée statistique chez ces étudiants du plus haut niveau et on peut craindre que cette situation dommageable ne soit pas fondamentalement modifiée avec les nouveaux programmes.

Intermédiaire est la situation dans les classes préparatoires aux grandes écoles, dans leurs sections à dominantes biologiques ou bien de gestion, économie et commerce. Là les programmes en statistique ont été conçus en fonction du besoin des autres disciplines ; convaincre les étudiants de ce besoin n'apparaît pas toujours aisé, mais les enseignants de mathématiques en connaissent la nécessité (au rebours de nombre de leurs collègues enseignant dans les classes faisant l'objet du paragraphe précédent) et se sont donc consacrés à présenter un enseignement de la statistique qui soit motivant pour leurs élèves.

6 Conclusion

La situation que nous venons de décrire est loin d'être figée. Une meilleure articulation entre les écoles et les collèges est un des objectifs affichés par le ministère de l'Éducation Nationale et elle pourrait avoir des répercussions sur la mise en situation concrète des outils mathématiques et donc aussi sur la présentation de la statistique au collège. Les programmes de mathématiques des lycées, quoique tout récents, vont sans doute connaître prochainement de nouvelles modifications, non sans doute dans la place globale faite aux probabilités et à la statistique ni dans leur conception, mais dans les possibilités d'interaction avec les autres branches des mathématiques¹². Les professeurs de mathématiques vont s'approprier de plus en plus les documents existant ou à venir mis à leur disposition, en particulier s'agissant de situations plus ou moins concrètes à exploiter, notamment en collaboration avec les enseignants d'autres disciplines ; à ce sujet on peut espérer une intensification de l'usage de la statistique dans les études pluridisciplinaires que nous n'avons pas évoquées dans le corps de cet article car elles sont souvent optionnelles, mais qui devraient se généraliser, par exemple en seconde avec « l'enseignement de détermination » intitulé *Méthodes et Pratiques Scientifiques*. C'est en fait tout un climat qui évolue en France autour de l'enseignement de la statistique, avec, au-delà des justifications liées aux applications scientifiques ou techniques (déjà largement exploitées dans les sections technologiques de lycée ou celles de techniciens supérieurs), des motivations relevant de la formation du citoyen (on signalera à cet égard les travaux du groupe *Statistiques, informatique et citoyenneté* de l'IREM de Paris-Nord (IREM Paris-Nord) ou le site *Statistix (Centre de ressources, lieu de partage et de mutualisation pour l'enseignement de la statistique : www.statistix.fr)*). Cette évolution mérite d'être observée avec intérêt.

Références

- [1] BOEN (2008), Horaires et programmes de l'école primaire, Bulletin Officiel de l'Éducation Nationale, 19 juin 2008, <http://www.education.gouv.fr/bo/2008/hs3/default.htm>
- [2] DEPP (2012), L'éducation nationale en chiffres, année scolaire 2011-2012, *Direction de l'Évaluation, de la Prospective et de la Performance*, <http://education.gouv.fr/cid57111/1-education-nationale-en-chiffres.html>

¹² Le ministère de l'éducation nationale vient de mettre en place une « commission de suivi des programmes de mathématiques » ; son premier rapport sera consacré à la classe de seconde, choisie car elle concerne tous les élèves autres que ceux des lycées professionnels et que c'est pour elle que l'on dispose du plus de recul (mise en place du programme actuel en 2010).

J.-P. Raoult

- [3] Dutarte, P. (2011), Evolutions de la pratique statistique dans l'enseignement du second degré en France : du calcul statistique au développement de la pensée statistique, *Statistique et Enseignement*, **2**(1),
<http://publications-sfds.math.cnrs.fr/index.php/StatEns/article/view/72>
- [4] Eduscol (2012), Programmes du lycée, Portail national des professionnels de l'éducation,
<http://eduscol.education.fr/pid26017/programmes-du-lycee.html>
- [5] Eduscol-Ressources, Mathématiques, ressources pour faire la classe au collège et au lycée, *Portail national des professionnels de l'éducation*,
<http://eduscol.education.fr/cid45766/mathematiques-pour-le-college-et-le-lycee.html>
- [6] Henry, M. (2011), Evolution de l'enseignement secondaire français en statistique et probabilités, *Statistique et Enseignement*, **1**(1),
<http://publications-sfds.math.cnrs.fr/index.php/StatEns/article/view/4>
- [7] IREM Paris-Nord, Statistiques, informatique et citoyenneté,
<http://www-irem.univ-paris13.fr/spip/spip.php?rubrique15>
- [8] Moussavou, F. (2010), La fluctuation d'échantillonnage dans les nouveaux programmes de baccalauréat professionnel, *Mathematice*, 13,
<http://revue.sesamath.net/spip.php?article192>