

# Une vision prospective de l'enseignement de la statistique

**Claudine Schwartz**

Université Joseph Fourier, Grenoble, Laboratoire de Modélisation et Calcul, IREM

**Résumé :** Après avoir rappelé les enjeux de l'enseignement de la statistique, nous en proposerons une vision prospective, en termes de grands objectifs. Nous illustrerons, à travers des exemples, quelques aspects du monde de l'aléatoire qui peuvent fonder un enseignement, du primaire au lycée, dans diverses disciplines.

Nous évoquerons aussi les problèmes que cela pose, notamment pour la formation initiale et continue, et les dynamiques qu'on peut envisager de créer pour apporter des solutions.

Ce texte est le résumé d'une conférence dont le but était de soumettre à la réflexion de la communauté des professeurs des thèmes pour l'enseignement de la statistique.

## A propos de la statistique et de son enseignement

Les thèmes d'enseignement qui seront proposés sont calés sur une vision de la statistique dont on dégage ci-dessous quelques éléments.

Le mot statistique est pris dans un sens large : statistique descriptive, modélisation, théorie des probabilités. La statistique préside au recueil de données (les « data »), à leur codage, leur représentation synthétique graphique ou numérique, leur mise en rapport et leur mise sous forme d'objets construits (indices) ; elle établit aussi des liens entre les données et la théorie des probabilités. Nous excluons cependant ce qui concerne les notions de probabilités subjectives utilisées autrefois pour décrire des degrés de croyance.

En ce sens élargi, la statistique est un mode (et un monde) de pensée, avec son langage, ses concepts, ses choix consensuels et discutables, ses processus de validation.

Le déploiement des outils de la description et de l'inférence statistique accompagne celui de l'informatique, tant au niveau du stockage des données que de leur traitement à l'aide de logiciels adaptés. Il convient de comprendre ces outils et techniques (tests, fourchette de sondage par exemple), tout en étant vigilant à ne pas limiter la statistique à leur utilisation.

La statistique est au carrefour d'un champ de la pensée empirico-inductive (description et interprétation des données) et d'un champ de pensée hypothético-déductif (celui de la théorie, celui des mathématiques achevées) : ce seul fait en rend l'enseignement déroutant tant pour les professeurs de mathématiques que pour les professeurs des sciences expérimentales (sciences de la vie et de la terre, physique chimie). Mais déroutant n'est pas impossible, et crée même un enjeu pédagogique particulièrement intéressant.

La statistique est aussi en tension entre la création de savoirs (notamment au niveau macro-social) et l'exercice de pouvoirs fondés sur l'usage d'outils qui encadrent l'action (choix de médicaments, d'une politique sociale) : là aussi cette tension est déroutante tant pour des professeurs de mathématiques qui vivent du côté du savoir et de la résolution *certaine* de problèmes que de celle des professeurs de sciences économiques et sociales qui travaillent sur des données complexes, avec des stratégies d'action basées sur des modèles mathématiques qui ne le sont pas moins.

La statistique propose des preuves, dont la nature spécifique peut-être éclairée par la locution anglaise « piece of evidence » ; la preuve statistique n'est qu'un élément de preuve qui confirme ou infirme une hypothèse théorique, qui valide ou rejette un modèle. Ces éléments de preuve sont d'autant *plus puissants* qu'ils reposent sur l'analyse de données nombreuses ou particulièrement parlantes en regard des hypothèses en jeu. Comme la preuve statistique ne produit pas à elle seule un résultat « ferme et définitif », elle est élaborée sous une forme qui permet la décision (avec un niveau de confiance ou une quantification du risque). Les registres de la preuve en statistique (preuve *basée sur des faits* expérimentaux) et en mathématique (démonstration sans appel dans un monde théorique axiomatisé) opèrent dans des mondes distincts, et peuvent s'éclairer mutuellement.

Il convient à ce niveau de dire clairement que la statistique....peut s'apprendre ! Et oui, c'est un scoop : on a plutôt tendance à la présenter comme un chemin parsemé d'embûches et de pièges où les meilleurs esprits s'égarent ; il en est ainsi précisément parce-qu'il n'y a pas eu d'apprentissage - et non parce qu'on serait face à une forme d'intelligence « naturelle » (!) dont seuls quelques originaux auraient été dotés à la naissance. La confiance dans les nombres s'acquiert, la philosophie de la réalité que construit la statistique se discute et se confronte aux faits dont elle est issue. A l'inverse, la suspicion vis-à-vis de l'information chiffrée est enfant de l'ignorance et voir la statistique comme l'art du mensonge n'est souvent qu'une simple posture de confort pour qui ne cherche pas à savoir.

La statistique, en créant des outils performants de description du monde ouvre à de larges débats : l'enseignement se doit de permettre au plus grand nombre de pouvoir y participer.

## Des objectifs d'enseignement

### Pour l'école primaire

La pensée en milieu aléatoire gagne à être abordée tôt, les enfants y étant dès leur plus jeune âge confrontés. Mieux vaut faire en sorte d'encadrer l'élaboration de leurs premières intuitions et de les organiser, pour accueillir ultérieurement une construction régulière et continue de connaissances.

Les thèmes suivants peuvent être abordés à l'école primaire à travers diverses activités :

- Chances égales (inégaies)
- Hasard équitable (pas équitable)
- Le dé n'a pas de mémoire (indépendance des expériences)
- La chance aux dés n'existe que pendant un temps court
- Histogrammes (« horizontaux » ?)
- « Hasard » : l'intervention du hasard ne rend pas la réflexion inutile



**Il convient à ce niveau de laisser les élèves s'exprimer ; ainsi dans une classe de CM1, ils ont énoncé que les chances égales, c'était quand on faisait pas à pas le dessin (l'histogramme ci contre en est une étape) des lancers, il y avait toujours une face gagnante, mais cela pouvait être n'importe laquelle des six faces.**

L'école primaire est un juste lieu pour expérimenter avec des dés pour deux raisons au moins :

- L'indépendance des résultats des lancers d'un dé y est vue sous l'angle « absence de mémoire » ; les enfants comprennent alors très vite que si on a obtenu deux fois de suite la face 3, les chances sont inchangées pour le lancer suivant. C'est d'autant plus simple à comprendre pour eux qu'ils n'ont pas d'image mentale associée à la loi des grands nombres, image qui fait paradoxe avec cette absence de mémoire (comment peut-il y avoir absence de mémoire et régulation des fréquences à long terme ?).
- de nouvelles calculatrices de poche ont une touche dé qui remplacera les dés pour les collégiens (ou alors les jeux de hasard seront directement vendus avec des dés électroniques). Avec un dé électronique, parler d'absence de mémoire au collège relève de l'argument d'autorité et non plus d'une expérience personnelle...

Pour plus de détails, on pourra consulter :

<http://www.statistix.fr/spip/spip.php?rubrique15>

### **Pour le collège**

- l'enseignement de la statistique doit être partagé entre de nombreuses disciplines
  - Le travail sur le discours oral et le discours écrit est important et participe de l'apprentissage de la langue française.
  - Recueil simple de données par les élèves et mutualisation des résultats. On évitera le recueil de données personnelles (les notes en font partie). On pourra aussi travailler sur des extraits de gros fichiers trouvés sur Internet, chaque binôme ayant un sous-fichier différent : on observe ainsi la fluctuation des moyennes.
  - Travail interclasse (classes de niveaux différents ou de sections différentes qui se communiquent leurs résultats)
- emploi du mot probabilité ( 3 chance sur 4, ou 6 chances sur 8, c'est pareil, c'est la probabilité 3/4 ) ;
  - simulation avec des dés électroniques ; somme de deux dés
  - simulation de promenades aléatoires
  - représentations graphiques simples, lecture de tableaux à double entrée
  - moyenne, étendue, médiane, quartile, déciles

### **Pour le lycée**

Certains éléments du programme de seconde actuels seront traités au collège et on pourra alors aborder la notion de loi de probabilité dès la seconde. Dans le cadre d'un enseignement partagé entre les disciplines, on pourra introduire la loi de Gauss, la notion de processus, définir l'espérance de vie, la notion de norme statistique et de hors norme.

On trouvera des exemples de thèmes interdisciplinaires sur le site Statistix (<http://www.statistix.fr>).

### **En guise de conclusion :**

La statistique actuelle est issue de la synthèse de différents courants, notamment celui de la statistique administrative et celui des probabilités. Il reste cependant encore un écart culturel entre les

statisticiens universitaires ou ingénieurs et ceux de la statistique publique (INSEE<sup>1</sup> et INED<sup>2</sup> notamment). Un enseignement harmonieusement partagé entre les différentes disciplines statistiques, une formation des maîtres qui inclue théorie, modélisation et une bonne connaissance des sites de la statistique publique permettront à terme de réduire cet écart. L'INSEE et l'INED ont fait ces dernières années, à travers leur site, des efforts particuliers de pédagogie vis-à-vis du milieu enseignant. Comme ce sont des sites riches, l'habitude d'y circuler demande du temps ; c'est en formation initiale que cette habitude pourrait s'acquérir, si on veut que les enseignants en poste trouvent ensuite le temps de les exploiter avec leurs élèves.

---

<sup>1</sup> Institut National de la Statistique et des Etudes Economiques

<sup>2</sup> Institut National des Etudes Démographiques