

# L'UTILISATION DES LOGICIELS DANS LES ENSEIGNEMENTS DE STATISTIQUE À L'UNIVERSITÉ

Christine Keribin<sup>1</sup>

<sup>1</sup> *Université Paris Sud, Laboratoire de Mathématiques, 91405 Orsay.  
christine.keribin@math.u-psud.fr*

**Résumé.** Cet exposé se place dans le cadre d'une série de communications orales dédiées à des expériences commentées d'enseignements en statistique. Les logiciels sont maintenant devenus un élément constitutif du savoir statistique et l'enseignement en statistique ne se conçoit plus sans séances de travaux dirigés ou pratiques sur machine. Ils peuvent être utilisés à tout niveau à l'université, pour découvrir les principes statistiques, illustrer des méthodes ou préparer à l'insertion professionnelle. Par la présentation d'expériences d'enseignements sur différents logiciels, j'aborderai en particulier la question de l'adaptation de l'outil utilisé en fonction du public et des objectifs pédagogiques visés.

**Mots-clés.** Enseignement de la statistique, logiciels

**Abstract.** Softwares are now part of the statistical science and teaching statistics cannot be thought without practical experiences with softwares. Software can be used at each degree of the university, for theoretical or more technical courses, to discovering statistical principles, illustrating methods or preparing to a professional environment. This talk presents teaching experiences in statistics with various softwares, and discusses more particularly the choice of a software according to the audience and the pedagogical objectives.

**Keywords.** Teaching Statistics, statistical software

## 1 Contexte

*Cet exposé trouve sa place dans une série de communications orales dédiées à des expériences commentées d'enseignements en statistique.* L'utilisation de logiciels dans un enseignement de statistique peut répondre à différents objectifs: rendre concrètes des notions théoriques vues en cours, expérimenter des méthodes sur des cas d'études, rendre opérationnel pour une entrée dans la vie professionnelle. Lors de mes enseignements en statistique dans diverses filières au niveau licence (DUT informatique, licences professionnelles datamining ou environnement ) ou master (ingénierie mathématique, MIAGE), j'ai été amenée à adapter les enseignements pratiques en fonction du public et des objectifs pédagogiques visés. Cet exposé discute l'utilisation des logiciels et leur adaptation à différents publics et objectifs pédagogiques.

## 2 Choix du logiciel

Il existe une offre nombreuse en logiciels statistiques, du langage de programmation à l'interface utilisateur presse-bouton, cette palette pouvant même se retrouver au sein d'un même éditeur. Le choix pour l'enseignant peut dépendre de divers facteurs comme les fonctionnalités offertes par le logiciel, son prix, sa disponibilité pour les étudiants ou sa diffusion dans le monde professionnel. La portée des statistiques dans la formation considérée et les compétences informatiques initiales des étudiants sont également à prendre en compte. Enfin, l'historique de cet enseignement dans la formation ou les propres compétences de l'équipe enseignante sur l'un ou l'autre logiciel peuvent aussi influencer sur le choix du logiciel.

Dans la maquette globale d'une formation statistique, il est intéressant de confronter les étudiants à l'utilisation de différents logiciels. Mon optique est d'utiliser le logiciel comme un support à la compréhension et la manipulation de notions et modèles statistiques, tout en permettant à l'étudiant d'assimiler la philosophie du logiciel utilisé et ses mécanismes principaux: il ne s'agit pas de rendre les étudiants spécialistes de l'utilisation d'un logiciel donné (les éditeurs proposent des formations dédiées, et les entreprises ont souvent un cursus de formation initiale pour les nouveaux embauchés), mais plutôt de les rendre compétents en statistique et sachant être à l'aise avec un logiciel, connu ou inconnu. C'est ce qui les rend autonomes et opérationnels dans l'entreprise. Mais bien sûr, avoir manipulé tel ou tel autre logiciel permet de faciliter un premier contact lors d'une demande de stage ou d'emploi.

## 3 Quelques expériences

### 3.1 Le logiciel comme apprivoisement

Une grande partie des étudiants de *DUT informatique* voient les cours de mathématiques comme une contrainte. Certains viennent de filière STI et se considèrent souvent irrémédiablement perdus avec un intérêt définitivement absent. Le programme de formation est défini par un Programme Pédagogique National (PPN) qui comporte des probabilités et statistiques pour l'informatique de gestion et le traitement des données. L'objectif est la connaissance des lois de probabilité usuelles et la compréhension de l'incertitude, avec une recommandation: utiliser un logiciel afin d'illustrer les notions et outils introduits et pratiquer la simulation.

Des expérimentations sur Excel et R ont été menées. Elles ont montré l'intérêt de s'appuyer sur une compétence établie (ici la compétence de programmation) pour amener les étudiants à s'intéresser à des problématiques mathématiques (et statistiques). A noter cependant dans ce type de situation, qu'il faut faire attention à ce que les questions techniques de programmation ne prennent pas le pas sur les notions statistiques à manipuler.

## 3.2 Le logiciel comme illustration

La statistique est la seule discipline mathématique de l'année de formation de la licence professionnelle *Aménagement du territoire et urbanisme* spécialité *Gestion durable des eaux pluviales*. Les étudiants sont encore plus réticents que les précédents, et ils n'ont pas de compétences informatiques particulières sur lesquelles s'appuyer. Les objectifs pédagogiques du module de statistique sont la sensibilisation à l'aléatoire, la familiarisation avec le principe d'événements extrêmes et la connaissance de certaines procédures à savoir utiliser. Je présenterai les difficultés et réussites rencontrées lors de l'utilisation du logiciel dans ce contexte.

## 3.3 Le logiciel pour développer l'esprit critique et prendre du recul

Pour les étudiants de master plus solides en mathématique et/ou statistique, le logiciel est bien entendu un terrain incontournable d'expérimentation. Dans un tel cadre, l'utilisation d'un logiciel en ligne de commande est préférable à celle d'un progiciel presse-bouton qui initialise des paramètres par défaut sans que l'utilisateur n'en ait bien conscience. Pour aller plus loin, il n'est rien de tel que de programmer soi-même une méthode pour bien en comprendre la structure, les finesses et les limites.

L'utilisation en parallèle de plusieurs logiciels permet d'en comparer les sorties: les sorties d'une méthode statistique sont apparemment différentes suivant les logiciels, et pourtant, chaque solution est correcte. La confrontation permet d'approfondir la connaissance des méthodes utilisées, et de prendre du recul par rapport à l'utilisation.

Au delà de ces utilisations académiques, le travail sur logiciel permet de comparer des méthodes sur de vrais jeux de données, et de transformer des connaissances en compétences statistiques. C'est une étape importante participant à la professionnalisation.