



**HAL**  
open science

## Un tableur grapheur pour enseigner les statistiques en sciences humaines et sociales.

Eric Roditi

► **To cite this version:**

Eric Roditi. Un tableur grapheur pour enseigner les statistiques en sciences humaines et sociales.. G.-L. Baron, E. Bruillard, L.-O. Pochon. Informatique et progiciels en éducation et en formation, INRP, pp.257-275, 2009. halshs-00609639

**HAL Id: halshs-00609639**

**<https://shs.hal.science/halshs-00609639>**

Submitted on 19 Jul 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Un tableur grapheur pour enseigner les statistiques en sciences humaines et sociales

*Éric Roditi*

---

UNIVERSITÉ PARIS DESCARTES, ÉQUIPE EDA (EA 4071), FRANCE

Une feuille de calcul modélisant une situation permet à l'utilisateur auquel elle est proposée d'identifier et de mettre en relation les variables en jeu. Les enseignants peuvent ainsi concevoir, à l'aide d'un tableur grapheur, des artefacts didactisés (Baron, 2006) pour leur enseignement, y compris si les savoirs visés ne sont pas informatiques. Le texte qui suit rend compte d'une recherche qui porte notamment sur l'utilisation du tableur grapheur pour l'enseignement des statistiques à des étudiants de sciences humaines et sociales. Elle est menée au sein de l'équipe EDA de l'université Paris Descartes.

Dans cette recherche, l'objectif de l'enseignement n'est pas que les étudiants deviennent des utilisateurs experts de logiciels de traitements statistiques plus ou moins élaborés. Il s'agit plutôt de leur permettre d'étudier des concepts et des outils mathématiques et qu'ils soient capables de les mettre en œuvre pour résoudre des problèmes concrets. On le comprend, la recherche dont il est question est d'abord une recherche en didactique des mathématiques où les « *enjeux didactiques des progiciels* » sont particulièrement forts, même s'ils le sont moins sur l'apprentissage du tableur que sur l'apprentissage de mathématiques.

Après avoir précisé les éléments de contextes spécifiques à l'enseignement des statistiques en sciences humaines et sociales, nous analysons les arguments qui sous-tendent l'utilisation du tableur pour concevoir des artefacts destinés soit plutôt à l'enseignant soit plutôt à l'étudiant. En traitant l'exemple du test statistique du  $\chi^2$  (lire ki-deux ou ki-carré) et après quelques rappels succincts sur ce test, nous présentons deux types de feuilles de calcul qui ont été conçues, les unes pour une utilisation par l'enseignant pendant les cours, et les autres pour des étudiants qui les emploient, pas seulement comme un logiciel de statistique qui donnerait les résultats après qu'ils ont indiqué leurs données, mais d'abord comme un moyen d'apprentissage et donc de questionnement mathématique.

## **Contexte de la recherche et éléments de problématique**

L'enseignement de statistiques inférentielles aux étudiants de sciences humaines et sociales est classique, mais il reste problématique. Les moyens informatiques de calcul et de

représentation, aujourd'hui très accessibles, semblent apporter des possibilités nouvelles pour améliorer cet enseignement.

### Les statistiques en SHS, un contenu problématique à bien des égards

Les statistiques inférentielles requièrent notamment des probabilités, du calcul intégral, du calcul matriciel, de la géométrie des espaces vectoriels euclidiens et de la topologie. Ces mathématiques font l'objet d'enseignement dans les cursus de mathématiques, éventuellement appliquées aux sciences sociales, mais restent difficiles à acquérir, même pour ces étudiants scientifiques. Autant dire qu'elles ne peuvent faire l'objet d'un enseignement pour des étudiants de sciences humaines. Les auteurs des manuels qui leur sont adressés prétendent rester à un niveau mathématique élémentaire, mais la plupart de ces étudiants ont suivi un programme d'enseignement secondaire où les outils nécessaires à leur lecture sont peu présents. Le niveau qu'ils ont atteint est rarement suffisant et leur rapport à cette discipline étudiée au collège et au lycée est souvent difficile. Leur passé scolaire ne les incite donc pas à suivre volontiers des cours qui comportent des mathématiques, cela les inquiète toujours, les motive rarement et les rebute souvent. Ils vivent généralement l'enseignement de statistique comme une épreuve qu'ils ne peuvent contourner et qu'ils essaieront de passer, sans trop d'espérance quant aux apprentissages qu'ils pourront développer.

Dans leur cursus, ces étudiants sont pourtant amenés à lire des articles de recherche ou des ouvrages qui comportent des références à des notions ou à des outils de statistiques descriptives ou inférentielles. Les plus avancés, pour mener leurs propres travaux de recherche ou pour en communiquer les résultats, ont généralement besoin de recourir à de telles méthodes. Cela leur permet par exemple de résumer une situation, de rendre compte de l'évolution d'un phénomène ou d'effectuer des comparaisons.

En outre, les statistiques, indépendamment des mathématiques qu'elles convoquent, posent des difficultés d'apprentissage spécifiques montrées par de nombreux travaux (Bihan-Poudec et al, 2005 ; Lahanier, 1999 ; Regnier, 2002 ; Wozniak, 2005). Nous reviendrons sur ces difficultés en abordant celles qui concernent précisément le contenu d'enseignement visé par notre recherche : le test du  $\chi^2$ .

### La diffusion des logiciels ne résout pas les problèmes d'enseignement

Le développement puis la diffusion d'outils informatisés à partir des années soixante-dix ont profondément changé le rapport aux statistiques. Des calculs très importants ont en effet été rendus possibles par l'automatisation et la puissance des ordinateurs. Puis des outils logiciels très performants proposant des traitements numériques et graphiques ont rendu les méthodes statistiques accessibles de manière autonome. La plupart des logiciels produits visent en effet à guider et à simplifier le travail des utilisateurs en effectuant automatiquement les traitements et en proposant des résultats sous différentes formes. Reste alors à l'utilisateur le travail d'interprétation.

À l'université, en sciences humaines et sociales, de tels logiciels sont souvent proposés aux étudiants ; cela aurait pu répondre aux problèmes d'enseignement, mais on constate que, faute d'une formation suffisamment solide en statistique, leur capacité à utiliser correctement les

logiciels et à interpréter les résultats obtenus est top faible. Force est d'admettre que l'apprentissage des logiciels ne dispense pas d'une formation en statistique si l'on souhaite que les étudiants puissent répondre aux questions qu'ils rencontrent dans leur cursus comme dans leurs propres recherches.

### Le tableur grapheur, une piste explorée pour enseigner les statistiques

Les recherches sur les effets de l'utilisation de logiciels pour l'enseignement – en particulier des mathématiques – montrent qu'elle modifie les activités d'apprentissage : elle complexifie la situation didactique (il y a nécessité d'appropriation préalable de certains schèmes d'activité) et elle conduit les apprenants à mettre en œuvre des démarches nouvelles, des démarches exploratoires notamment, quasiment expérimentales ou plus simplement fondées sur des essais et erreurs.

Dans l'enseignement des mathématiques, on peut distinguer trois catégories de logiciels : ceux qui ont été conçus pour les mathématiciens et qui sont introduits dans l'enseignement, ceux qui ont été conçus pour l'enseignement des mathématiques et enfin les logiciels plus généraux qui sont utilisés dans l'enseignement. Dans la première catégorie on trouve par exemple les logiciels de calcul formel qui sont utilisés à un certain niveau de l'enseignement. Dans la catégorie des logiciels conçus pour l'enseignement des mathématiques, on peut penser aux logiciels de géométrie dynamique qui font l'objet de nombreux travaux de recherche (Dahan, 2005)<sup>1</sup>. En ce qui concerne les logiciels qui n'ont pas été conçus pour l'enseignement nous citerons le cas des tableurs grapheurs auxquels sont consacrées notamment des recherches sur l'apprentissage de l'algèbre et le rôle du tableur dans les apprentissages visés (Abramovitch<sup>2</sup> ; Arzarello *et al.*, 2002 ; Capponi, 1990 ; Haspekian, 2003).

Dans le domaine des statistiques, des possibilités offertes par le tableur ont déjà été explorées. Notre recherche se situe dans la continuité de travaux menés sur l'enseignement des statistiques notamment par la commission inter-IREM « Statistiques et probabilités »<sup>3</sup> (Chaput & Henry, 2005, 2007). Nous nous intéressons à l'usage du tableur dans un enseignement qui porte à la fois sur l'apprentissage de notions et de méthodes statistiques, sur la mise en œuvre pratique de ces méthodes dans des situations concrètes, et sur l'interprétation des résultats obtenus, tout en prenant en compte la réalité des connaissances mathématiques des étudiants. Nous cherchons notamment à analyser les effets de l'utilisation par l'enseignant et par les étudiants de feuilles de calcul conçues en fonction d'objectifs précis d'apprentissage et nous avons commencé par l'exemple très classique du test du  $\chi^2$ . Deux

---

<sup>1</sup> Voir : [http://www-cabri.imag.fr/cabri2/publications/ouvrages\\_en\\_francais.html](http://www-cabri.imag.fr/cabri2/publications/ouvrages_en_francais.html)

<sup>2</sup> Voir ses publications : [http://www2.potsdam.edu/honors/greek/clubs/EDUC/abramovs/publications\\_list.htm](http://www2.potsdam.edu/honors/greek/clubs/EDUC/abramovs/publications_list.htm).

<sup>3</sup> Voir sur le site des IREM : <http://www.univ-irem.fr/spip.php?rubrique25>.

autres membres de notre équipe de recherche travaillent sur des questions liées aux nôtres avec d'autres paradigmes : Georges-Louis Baron qui interroge le renforcement réciproque de la compréhension des statistiques et de la conceptualisation des processus de traitement de l'information avec un tableur, et Philippe Chaussecourte qui interroge, avec une approche clinique d'orientation psychanalytique, les phénomènes inconscients à l'œuvre dans ces situations d'enseignement et d'apprentissage où mathématiques, statistiques et informatique sont en jeu.

## **Le test du $\chi^2$ et ses difficultés d'apprentissage**

Dans ce paragraphe nous rappelons brièvement mais précisément au lecteur en quoi consiste un test du  $\chi^2$  et quels sont les enjeux d'enseignement d'un tel test en sciences humaines et sociales. Le lecteur familier de ces questions pourra passer directement au paragraphe suivant où sont étudiées des feuilles de calcul proposées dans l'enseignement de ce test.

### La méthode et ses difficultés d'apprentissage

Partons d'une situation fictive pour illustrer l'intérêt et les difficultés d'apprentissage de la méthode statistique. Dans une filière où l'insertion professionnelle des étudiants est fonction de la qualité de leur formation, on évalue les étudiants en fin de cursus par un examen auquel on constate 20% d'échec, 48% de réussite sans mention et 32% de réussite avec mention. Un établissement propose une formation dans cette filière et affiche 25% d'échec, 25% de réussite sans mention et 50% de réussite avec mention. Pour évaluer les résultats de cet établissement, on se propose de comparer les pourcentages de réussite obtenus avec ceux qui sont globalement constatés dans la filière. Trois possibilités apparaissent assez vite compte tenu de la situation pour laquelle l'obtention d'une mention est importante : comparer les répartitions échec / réussite (20% / 80% contre 25% / 75%), comparer les répartitions échec ou réussite sans mention / réussite avec mention (68% / 32% contre 50% / 50%), comparer les répartitions complètes (20% / 48% / 32% contre 25% / 25% / 50%). Le choix entre ces possibilités revient à celui qui évalue les résultats de l'établissement, la mise en œuvre de méthodes statistiques commence après.

Choisissons d'évaluer l'établissement en considérant le critère de la mention à l'examen, cela revient à comparer la répartition 68% / 32% à la répartition 50% / 50%. L'étudiant qui n'a pas encore appris de statistiques inférentielles conclut facilement que les candidats de l'établissement sont mieux formés qu'ils ne le sont généralement dans la filière car ils sont 50% à obtenir la mention lorsqu'ils sortent de l'établissement alors que généralement, ils ne l'obtiennent que pour 32% d'entre eux. Dès les premiers cours, il comprend (c'est une figure de style car il y a là une véritable difficulté d'apprentissage) que les pourcentages ne se comparent pas si facilement. On ne peut pas dire en effet qu'une femme qui a fait trois enfants dont deux garçons (soit 66% de garçons) réussit mieux à faire des garçons que son amie qui a fait quatre enfants dont deux garçons et deux filles (soit 50% de garçons). Il en irait en revanche tout autrement s'il s'agissait de comparer les naissances de deux pays où, pour le premier on compterait 300 000 naissances dont 200 000 garçons (soit aussi 66% de garçons) et pour le second 400 000 naissances dont 200 000 garçons (soit aussi 50% de garçons) ! Il

convient de tenir compte de la répartition des effectifs et pas seulement de celle des pourcentages.

Continuons donc d'explorer la méthode statistique mise en œuvre pour comparer les résultats des étudiants d'un établissement à ceux de toute la filière. Admettons que le nombre de candidats formés par l'établissement soit 28 ; il y a donc eu à l'examen 7 échecs, 7 réussites sans mention et 14 réussites avec mention. Pour comparer ces résultats avec ceux de la filière, la méthode statistique repose sur un raisonnement pour le moins original :

premièrement on déclare vouloir montrer que ces résultats sont exceptionnels (on précisera le sens de cet adjectif plus loin) et donc que l'établissement forme ses candidats de façon tout à fait particulière ;

deuxièmement on suppose exactement le contraire de ce qu'on déclare vouloir prouver : on pose en effet l'hypothèse que les 28 candidats ont été tirés totalement au hasard dans la population formée par tous les candidats à l'examen (cette hypothèse est couramment appelée « hypothèse nulle » car on cherche à l'invalider – *to nullify* en anglais) ;

troisièmement on évalue la probabilité  $p$  pour que, dans de telles conditions, les scores de réussite des candidats formés dans l'établissement soient à ce point éloignés (là encore il faudra préciser le sens de cet adjectif) de ceux constatés dans la filière ;

quatrièmement on tire une conclusion en fonction de la valeur de cette probabilité, en se disant que plus elle est faible, moins l'hypothèse que les 28 candidats soient quelconques est plausible et donc plus vraisemblable est l'existence d'un effet spécifique de l'établissement sur les résultats de ses candidats.

Au bilan, cette méthode répond à la question : la différence de résultats peut-elle être raisonnablement attribuée au hasard ? Elle repose sur un raisonnement qui ressemble fort à un raisonnement « par l'absurde » très difficile à apprendre (Arsac, 1987), comme l'est plus généralement le raisonnement conditionnel (les travaux sont nombreux à ce sujet en psychologie cognitive et en psychologie du raisonnement), et comme le sont aussi les probabilités conditionnelles (Nabbout, 2006) sur lesquelles reposent implicitement le calcul de la probabilité  $p$ . La méthode est par ailleurs déconcertante : il faut admettre que la réponse n'est pas sûre, qu'elle est entachée d'un risque, puisque 28 candidats tirés au hasard pourraient tous obtenir la mention, c'est peu probable, mais ce n'est pas impossible.

### Mise en œuvre et autres difficultés d'apprentissage

Alors que nous sommes partis d'une situation bien déterminée, les taux de réussites à un examen dans une filière et les résultats d'un établissement à cet examen, la méthode statistique nous invite à prendre en compte cette situation déterminée, mais aussi ce qui pourrait arriver si la situation était aléatoire. Des probabilités sont convoquées, nous allons les expliciter sommairement en décrivant la mise en œuvre du test de comparaison et nous indiquerons ce faisant le sens que nous accordons aux adjectifs « exceptionnels » et « éloignés » qui ont été utilisés précédemment.

On considère la population des candidats de la filière, on rappelle que dans cette population les scores sont ainsi répartis : 20% d'échec, 48% de réussite sans mention et 32% de réussite

avec mention. Les 28 candidats formés dans l'établissement constituent un échantillon de cette population pour lequel les scores sont 7 échecs, 7 réussites sans mention et 14 réussites avec mention. Le but du test est de décider si l'échantillon étudié « ressemble » ou non à la population de référence. Autrement dit, on cherche à savoir de combien les résultats de l'établissement sont « éloignés » des résultats de la filière, et par conséquent s'il est probable ou non que cet échantillon ait pu être un échantillon aléatoire de cette population.

L'évaluation de la différence entre les résultats de l'établissement et ceux de la filière conduit à évaluer *l'écart* entre la distribution des effectifs observée sur l'établissement et une autre distribution d'effectifs, qualifiée de théorique parce qu'imaginaire, correspondant à un échantillon de même taille, qui serait le reflet exact de la population de référence. Un calcul de pourcentage élémentaire montre que la distribution théorique des 28 candidats, conforme à la répartition 20%, 48%, 32% est 5,6 échecs, 13,44 réussites sans mention et 8,96 réussites avec mention. Deux difficultés à cette étape, une mineure qui vient du fait que le calcul de pourcentage qui n'est pas maîtrisé par tous les étudiants, et une autre plus importante qui vient du fait que les résultats obtenus ne sont pas des nombres entiers : comment une famille peut elle avoir en moyenne 2,1 enfants ? Puis vient le calcul de l'écart entre ces deux listes de nombres. La formule de calcul mise au point par les statisticiens repose sur des connaissances mathématiques qui relèvent seulement de l'algèbre élémentaire, sauf précisément le choix de la formule car la distance euclidienne usuelle ne convient pas (l'explication de ces raisons demande un certain développement et ne peut être présentée dans ce texte). Cela constitue une difficulté supplémentaire pour l'apprentissage.

La formule permet donc de dire précisément, par un nombre, de combien les scores obtenus sur l'établissement sont éloignés de ceux de la filière. Mais savoir un nombre ne suffit pas pour savoir s'il est grand : 2 secondes est une durée bien déterminée, c'est très peu s'il s'agit du temps perdu sur un déplacement de 400 m réalisé pour aller de son domicile à la station de métro la plus proche, c'est énorme en revanche s'il s'agit de la différence de temps entre deux coureurs du 400 m ! La même durée sur la même distance est appréciée différemment en fonction du contexte. Il vient alors une autre étape fondamentale de la mise en œuvre du test : évaluer l'importance de la valeur de l'écart obtenu. C'est à ce moment que les probabilités interviennent. L'écart effectivement constaté entre l'établissement et la filière est considéré comme une valeur particulière de l'écart considéré cette fois comme une variable aléatoire soumise aux fluctuations d'échantillonnage, c'est-à-dire dont la valeur varie à chaque fois qu'un échantillon de 28 candidats est tiré au sort dans la population. Sous certaines conditions, la loi de cette variable aléatoire est approximativement une loi bien connue en probabilité et en statistique qu'on appelle la loi de  $\chi^2$ . On détermine par conséquent, approximativement mais suffisamment précisément à l'usage, la probabilité  $p$  d'obtenir, pour l'écart entre les scores, une valeur égale ou supérieure à celle de l'écart effectivement constaté. N'entrons pas davantage dans les détails, les bases mathématiques des étudiants en sciences humaines et sociales ne suffisent pas à comprendre le calcul de cette probabilité, et de loin. La méthode, dans ses principes, n'est pas non plus facile à acquérir : elle propose un modèle aléatoire pour évaluer si une distance est importante ou non. C'est ce modèle aléatoire qui donne son sens au terme « exceptionnel » que nous avons utilisé : si la probabilité est très faible, les résultats sont considérés comme exceptionnels. On choisit conventionnellement 5% ou 1% comme seuil – la valeur de ce seuil est appelée  $\alpha$  (alpha). Ainsi, avec un seuil  $\alpha$  de 5%,

l'écart entre les scores de l'établissement et ceux de la filière sera jugé important – car exceptionnel – s'il y a moins de 5 chances sur 100 d'en obtenir un aussi grand en tirant au sort 28 candidats dans la population.

Grâce à ce critère, on prend une décision sur les résultats. Si la probabilité  $p$  est inférieure à  $\alpha$  l'hypothèse nulle est rejetée et les résultats de l'établissement sont jugés différents de ceux de la filière. On conclut alors, au seuil  $\alpha$ , que le test est significatif ou que la différence des résultats est significative ou encore qu'il y a un effet significatif de l'établissement sur les scores des candidats. On précise parfois la valeur de la probabilité  $p$  qu'on appelle alors le degré de signification du test. Inversement, si la probabilité  $p$  est égale ou supérieure à 5%, la valeur de l'écart n'est pas jugée suffisamment importante et l'hypothèse nulle n'est pas rejetée. Autrement dit, rien n'interdit de penser que les résultats obtenus par l'établissement aient pu être obtenus par hasard par un groupe quelconque de 28 candidats. La différence de résultats n'est pas significative. Deux difficultés d'apprentissage sont encore à signaler à cette étape de la mise en œuvre du test : le sens du seuil  $\alpha$  est très mal interprété par les étudiants, de même que le sens du terme significatif, certains étudiants se refusant par exemple d'interpréter les résultats d'un test négatif pensant que le test lui-même n'a alors aucune signification.

Précisons pour terminer que certains auteurs présentent la règle de décision un peu autrement, notamment lorsqu'ils font référence à une table de valeurs de la loi de  $\chi^2$  dans laquelle on lit la valeur qui correspond à une probabilité de 5%, cette valeur est alors appelée  $\chi^2$  théorique tandis que l'écart entre les scores est appelé  $\chi^2$  empirique (ou  $\chi^2$  calculé). La règle de décision est alors de rejeter l'hypothèse nulle si la valeur du  $\chi^2$  empirique est supérieure à celle du  $\chi^2$  théorique ; il n'y a pas de calcul de la probabilité  $p$  c'est-à-dire du degré de signification.

### Enjeux d'un tel enseignement en sciences humaines et sociales

L'enseignement de statistique inférentielle en sciences humaines et sociales répond au besoin très fréquent de comparaison. En outre, la loi de  $\chi^2$  a une importance particulière car elle intervient, à un niveau plus élevé, dans l'analyse de données conduisant à classer des individus et à dégager des proximités entre individus.

Le test du  $\chi^2$ , tel qu'il est enseigné classiquement, permet de répondre à trois questions de nature différente. Comme nous l'avons vu en détail précédemment, le test permet de comparer un groupe d'individus à un groupe de référence pour leur répartition entre différentes modalités d'un caractère (test de conformité) ; le caractère était le résultat à l'examen et les trois modalités étaient l'échec, la réussite sans mention et la réussite avec mention. Le test permet aussi de comparer deux groupes d'individus entre eux (test d'homogénéité) ou encore d'apprécier, dans un groupe, l'indépendance de deux caractères (test d'indépendance), ce serait le cas par exemple si l'on comparait les résultats à l'examen des garçons et des filles.

On comprend que les étudiants de sciences humaines et sociales aient à connaître ces outils très utilisés dans ces disciplines, aussi pour en comprendre la portée et les limites.

## Analyse de deux types d'artefacts didactisés

Plusieurs difficultés d'apprentissage ont donc été repérées. Pour trois d'entre elles, nous avons conçu des feuilles de calcul et des situations à étudier qui devraient aider les étudiants à les surmonter. Certaines feuilles visent à faire expérimenter les fluctuations d'échantillonnage, nous ne les détaillerons pas car les arguments essentiels sont identiques à ceux développés par l'équipe de l'IREM Paris Nord (Dutarte et al., 1998, 2000). D'autres proposent de montrer comment une probabilité peut être associée au  $\chi^2$  empirique, c'est-à-dire à l'écart entre les scores de l'établissement et ceux de la filière. Les dernières enfin constituent une alternative didactique au logiciel de statistiques : elles permettent de traiter des situations pratiques, mais elles montrent les différentes étapes de calcul qui sont généralement masquées à l'utilisateur d'un tel logiciel. Ces feuilles de calculs sont utilisables par l'enseignant et par les apprenants, elles constituent en ce sens de véritables *artefacts didactisés* (Baron, 2006).

### Un artefact didactisé, utilisé principalement par l'enseignant

Reprenons la situation d'examen exposée précédemment. Comment montrer aux étudiants qu'on peut associer une probabilité à l'écart calculé entre les résultats des 28 candidats de l'établissement d'une part et de la filière d'autre part ? Nous avons décidé de recourir à une approche fréquentiste de cette probabilité, et d'en obtenir une approximation en simulant un grand nombre de tirages au sort fictifs d'échantillons de 28 candidats. Le nombre de tirages nécessaires pour que les variations liées aux fluctuations d'échantillonnage ne soient pas sensibles étant très grand, il n'était pas question de réaliser une simulation matérielle, nous avons donc opté pour une simulation informatique sur feuille de calcul. Le passage par la simulation permet d'éviter la modélisation mathématique nécessaire au calcul de cette probabilité et de rendre sensible la dimension aléatoire du test du  $\chi^2$ .

Nous savons que nos choix didactiques soulèvent aussi des difficultés spécifiques. Citons par exemple, les conceptions sous-jacentes à l'approche fréquentiste des probabilités auxquelles la simulation fait référence (Reuter, 1999 ; Bernier & al., 2000 ; Piednoir 2006). Citons encore l'explicitation du théorème central limite qui justifie la méthode, et qui est très délicate pour les étudiants concernés. Nous avons estimé néanmoins, compte tenu de travaux antérieurs sur cette question, que la simulation offre ici un potentiel didactique intéressant (Bordier, 1991 ; Dutarte, 2002 ; Girard & Henry, 2005 ; Pichard, 2005).

La feuille de calcul ainsi programmée constitue un *ostensif* numérique et graphique de la détermination de la fréquence obtenue. Comme la feuille comporte une simulation de nombreux tirages aléatoires, l'ostensif n'est pas figé, il comporte une part variable, tout comme une figure de géométrie comporte une part variable (de forme, de dimension, de position, etc.). Cette variabilité peut être expérimentée par les usagers, comme en géométrie dynamique, et elle favorise ce faisant le passage de la fréquence à la probabilité. Une telle feuille de calcul est utilisable dans un cours ordinaire assorti de séances de travail dirigé en tirant parti à la fois des possibilités de projection d'écrans d'ordinateur et du fait que tous les ordinateurs peuvent être équipés sans frais d'un tableur grapheur.

Comme le montre la figure n°1, la feuille de calcul comporte trois zones. La barre de formule qui n'apparaît pas dans les figures ci-dessous, permet aux étudiants de comprendre comment les calculs sont effectués.

Dans la première zone figurent les données de la situation et le calcul de l'écart entre la distribution théorique (conforme aux résultats de la filière) et la distribution empirique (constaté dans l'établissement). Le carré de cet écart est de 6,27 environ.

<b>Théorique</b>				
	<b>Echec</b>	<b>RSM</b>	<b>RAM</b>	<b>Total</b>
en %	20%	48%	32%	100%
pour 28 candidats	5,60	13,44	8,96	28
<b>Empirique</b>				
	<b>Echec</b>	<b>RSM</b>	<b>RAM</b>	<b>Total</b>
	7	7	14	28
<b>Ecart théorique - empirique</b>				
Ecart <sup>2</sup>	0,350000	3,085833	2,835000	<b>6,27083</b>

Figure n°1

La figure n°2 indique une deuxième zone qui est celle des tirages aléatoires, on n'en voit qu'une partie car il y a 2 000 tirages de 28 candidats ! On a supposé la population constituée de 5 000 individus numérotés par ordre croissant de leur résultat, ces résultats correspondent à la distribution « théorique » puisque cette distribution est celle de la population. Vingt-huit individus sont tirés au sort l'un après l'autre et constituent un échantillon aléatoire de candidats de la population. Pour celui représenté dans la figure n°2, on constate 7 échecs, 15 réussites sans mention et 6 réussites avec mention. Le carré de l'écart est de 1,51 environ.

Simulations						
Individu	Centile	Candidat	Echec	RSM	RAM	aléa
1619	32	1	0	1	0	0,32363
2667	53	2	0	1	0	0,53339
3332	57	3	0	1	0	0,56622
538	14	4	1	0	0	0,13397
4420	88	5	0	1	1	0,88391
531	14	5	1	0	0	0,13808
1519	30	7	0	1	0	0,30290
2927	51	8	0	1	0	0,50636
1145	23	9	0	1	0	0,22900
1344	27	10	0	1	0	0,26880
263	5	11	1	0	0	0,09291
4738	35	12	0	0	1	0,34753
4733	35	13	0	0	1	0,34657
2186	44	14	0	1	0	0,44371
838	17	15	1	0	0	0,17351
1373	27	15	0	1	0	0,27453
1804	35	17	0	1	0	0,35070
3282	55	18	0	1	0	0,55026
1804	35	19	0	1	0	0,35070
595	12	20	1	0	0	0,11899
1732	36	21	0	1	0	0,36337
3706	74	22	0	0	1	0,74105
163	3	23	1	0	0	0,03291
2972	51	24	0	1	0	0,51423
2318	46	25	0	1	0	0,46336
4050	81	26	0	0	1	0,80996
352	7	27	1	0	0	0,07224
2400	62	28	0	0	1	0,62095
		Bilan	7	15	6	28
		Ecart <sup>2</sup>	0,35	0,18107143	0,97785714	1,50892857
		1	0	0	0	5,2
		2	0	0	0	13,8
		3	0	0	1	73,5
		4	0	1	0	31,3
		5	0	0	1	63,0
		6	0	0	1	39,1
		7	0	0	1	35,4
		8	0	0	1	87,1

Figure n°2

La figure n°3 montre une troisième zone de la feuille de calcul qui propose un bilan de la simulation de 2 000 échantillons. Dans l'exemple présenté par la figure n°3, on a obtenu 90 échantillons sur les 2 000 pour lesquels le carré de l'écart est supérieur à 6,27 celui constaté pour l'établissement. Cela représente une fréquence de 4,5% des échantillons. En effectuant une nouvelle simulation de 2 000 échantillons, on constate peu de variation de la fréquence liée aux fluctuations d'échantillonnage, cela laisse penser que la probabilité cherchée est bien inférieure à 5%. En outre, la troisième zone propose un tableau de distribution des carrés des écarts ainsi qu'un diagramme représentant cette distribution.

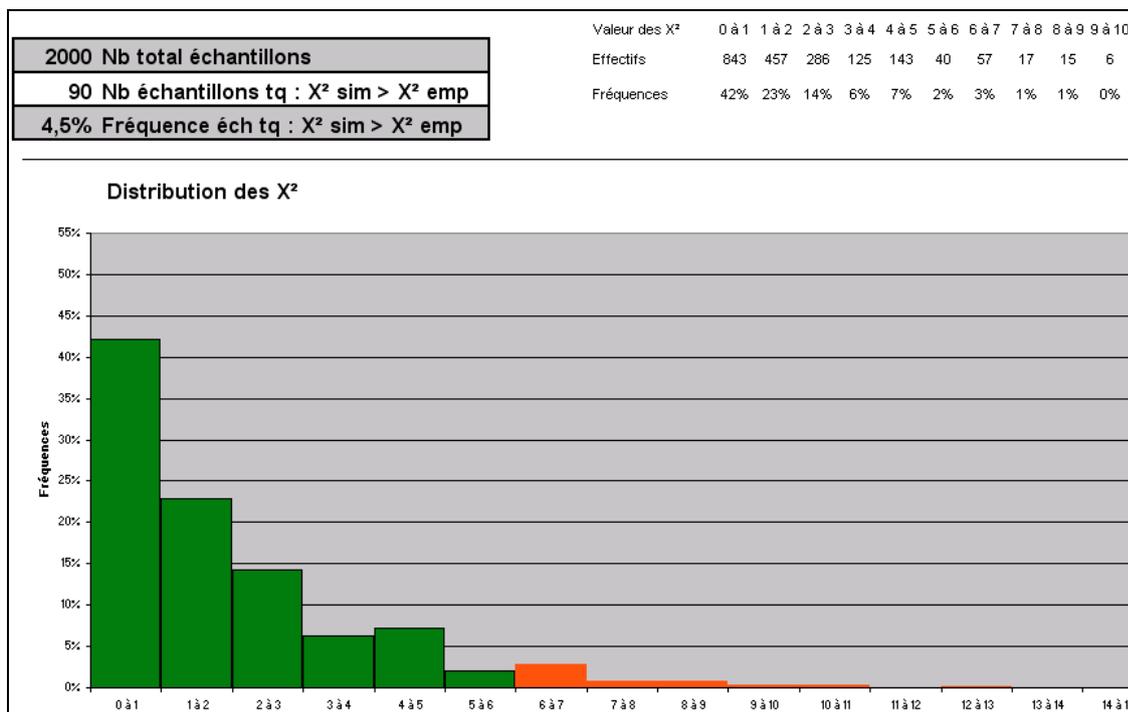


Figure n°3

En cours, la feuille de calcul est montrée aux étudiants comme un moyen de résoudre le problème de l'estimation d'une probabilité, problème dont la résolution par la simulation est accessible aux étudiants alors que la résolution par la voie mathématique traditionnelle ne l'est absolument pas. Entre leurs mains, après le cours, cette feuille peut être aussi considérée comme un environnement interactif d'apprentissage ouvert, c'est-à-dire ne guidant pas les interactions. Le travail sur les formules utilisées permettant, par exemple, de comprendre comment sont conçus les échantillons aléatoires et donc ce qu'est un échantillon aléatoire.

### Un artefact didactisé, utilisé principalement par les étudiants

Voyons maintenant un exemple du troisième type de feuilles de calcul, des feuilles qui constituent une sorte de logiciel permettant de traiter les situations qui conduisent à un test de  $\chi^2$ , des feuilles dans lesquelles les traitements numériques sont toujours effectués de façon « visible » si bien que l'artefact constitue simultanément un outil pratique et un *milieu* (au sens didactique du terme) pour une situation d'apprentissage. La copie d'écran proposée dans la figure n°4 de la feuille de calcul permettant de réaliser un test de conformité donne une idée de l'utilisation et des questionnements possibles.

Ainsi, l'utilisateur ayant saisi les données de son problème et le seuil  $\alpha$  connaît le résultat du test, mais pas seulement. Par exemple, le tableur exécutant de façon « visible » le calcul de l'écart entre la distribution observée et la distribution théorique, il permet aux étudiants de questionner les propriétés de cette nouvelle « distance ». Dans le cas de la situation des résultats à l'examen, on constate par exemple que le carré de l'écart est de 6,27 environ qui se répartit entre les trois modalités du caractère avec les contributions suivantes : 0,35 ; 3,09 et 2,84. On en déduit que la différence de pourcentage des échecs à l'examen contribue très peu

à l'écart entre les résultats de l'établissement et ceux de la filière. Il serait donc faux d'affirmer que, de manière significative au seuil de 5%, les candidats formés par l'établissement échouent davantage à l'examen que c'est généralement le cas dans la filière. Puis les étudiants peuvent utiliser la feuille de calcul réaliser diverses tâches proposées par l'enseignant, comme changer les valeurs de la distribution observée sans changer l'écart, les changer pour augmenter l'écart lié à une modalité donnée, etc.

Comparaison d'une distribution observée à une distribution de référence									
Saisir dans les zones vertes : 1. Le seuil du risque accepté - 2. Les modalités et leurs effectifs observés sur l'échantillon - 3. Les fréquences de référence pour ces modalités.									
Règle de décision									
Seuil alpha	5%	ddl = 2		X <sup>2</sup> théorique = 5,9914764					
Distribution empirique des effectifs									
Modalités	Echec	Passable	Mention						Total
Effectifs	7	7	14						28
Distribution théorique des fréquences									
Modalités	Echec	Passable	Mention						Total
Fréquences	20,0%	48,0%	32,0%						100%
Conclusion du test		Ecart Significatif						p = 4,3482%	
Distribution empirique des fréquences									
Modalités	Echec	Passable	Mention						Total
Fréquences	25%	25%	50%						100%
Distribution théorique des effectifs									
Modalités	Echec	Passable	Mention						Total
Effectifs	5,6	13,44	8,96						28
Ecart entre les distributions des effectifs (X <sup>2</sup> empirique)									
Modalités	Echec	Passable	Mention						X <sup>2</sup> empirique
	0,35	3,085833333	2,835						6,27083
Recherche d'effectifs théoriques inférieurs à 5									
Modalités	Echec	Passable	Mention						Bilan
	1	1	1						1

Figure n°4

Les étudiants peuvent interroger les valeurs des différents paramètres, variables et fonctions qui sont affichées ainsi que leurs relations, notamment les relations entre le seuil de signification et le  $\chi^2$  théorique, entre le degré de signification et le  $\chi^2$  empirique, etc. Dans l'enseignement, différentes tâches sont ainsi proposées qui conduisent à des activités favorisant d'une part l'apprentissage des tests statistiques et d'autre part l'apprentissage de l'interprétation de la signification du test, notamment en expérimentant les effets de variations des valeurs de la distribution observée sur ce résultat. Caractéristique essentielle, l'environnement mis à la disposition des étudiants leur permet de se l'approprier et de mener avec lui des genèses instrumentales.

## Conclusion

Dans la recherche que nous menons sur l'enseignement des statistiques inférentielles en sciences humaines et sociales, le tableur est utilisé comme élément essentiel pour répondre aux difficultés spécifiques majeures de leur apprentissage. Son rôle est très particulier : certaines feuilles de calcul sont utilisées face aux apprenants et constituent un ostensif doté d'une dynamique qui invite à reconsidérer le point de vue sur l'ostension telle qu'il apparaît dans les situations traditionnelles, c'est-à-dire comme fondant le développement des connaissances mathématiques sur l'observation et l'écoute du maître. D'autres feuilles de calcul sont utilisées par les étudiants, à la fois pour apprendre des savoirs de statistique et pour traiter des problèmes. Ces feuilles de calcul sont communiquées aux étudiants afin qu'ils puissent expérimenter eux-mêmes les situations et leur traitement, soit en partant de jeux de données qui leur ont été proposés, soit en devant les modifier ou en inventer.

Ces artefacts didactisés ne correspondent pas à une volonté d'intégration de technologies numériques dans un système d'enseignement existant, ils ne correspondent pas non plus à un enseignement de l'utilisation d'un outil pour la pratique, ils sont en revanche utilisés parce qu'ils rendent possibles la conception par l'enseignant de tâches pour de nouvelles activités des étudiants leur permettant de questionner les concepts mathématiques ainsi que de lier concepts mathématiques et problèmes concrets. Dans son ensemble, le travail de conception ne mobilise que des compétences assez élémentaires dans le domaine informatique, le travail de conception est donc essentiellement didactique.

La recherche est encore en cours, d'autres questions restent posées portant sur l'apprentissage de statistiques et sur les genèses instrumentales concernant le tableur : comment les étudiants accèdent-ils aux moyens de modélisation qui leur sont proposés ? Quelles conceptions construisent-ils des concepts dont l'enseignement ne vise pas un apprentissage mathématique tel qu'on le comprend habituellement en didactique des mathématiques ? Comment interprètent-ils les résultats des tests statistiques mis en œuvre ? Des questions portent aussi sur l'enseignement en classe, notamment quant aux interactions avec les étudiants et aux aides, tant sur les activités mathématiques que sur l'usage des instruments.

## Références bibliographiques

- ARSAC, G. (1987), L'origine de la démonstration : essai d'épistémologie didactique, *Recherches en didactique des mathématiques*, n°8/3, p. 267-309.
- ARZARELLO, F., BAZZINI, L. & CHIAPPINI, G. (2002). La pensée algébrique dans une perspective sémiotique : l'environnement du tableur, *Sciences et Techniques éducatives*, n°9/1-2.
- BARON, G.-L. (2006). De l'informatique à « l'outil informatique » : considérations historiques et didactiques sur les progiciels. Le cas particulier des logiciels de traitement de tableaux, in : L.-O. POCHON, E. BRUILLARD, & A. MARECHAL, *Apprendre (avec) les progiciels. Entre apprentissages scolaires et pratiques professionnelles*. Neuchâtel : IRDP, Lyon : INRP, p. 39-54.
- BERNIER J., PARENT E., BOREUX J.-J. (2000). *Statistiques pour l'Environnement. Traitement Bayésien de l'Incertitude*. Londres, Paris, New York : Technique & Documentation.
- BIHAN-POUDEC, A. et al (Eds), (2005). Mesurer – Actes du Symposium Pédagogie de la statistique à l'université ». Paris : L'Harmattan.

- BORDIER, J. (1991). Un modèle didactique utilisant la simulation sur ordinateur, pour l'enseignement de la probabilité, Thèse de Doctorat, Université Paris 7.
- CAPPONI, B. (1990). *Calcul algébrique et programmation dans un tableur : le cas de Multiplan*. Thèse de doctorat. Université Joseph Fourier, Grenoble.
- CHAPUT, B. et HENRY, M. (Eds), (2005). *Statistique au lycée*, vol. 1. Brochure 156. Paris : APMEP.
- CHAPUT, B. et HENRY, M. (Eds), (2007). *Statistique au lycée*, vol. 2. Brochure 167. Paris : APMEP.
- DAHAN, J.-J. (2005). La démarche de découverte expérimentalement médiée par cabri-géomètre en mathématiques - Un essai de formalisation à partir de l'analyse de démarches de résolutions de problèmes de boîtes noires. Thèse de doctorat. Université Joseph Fourier, Grenoble.
- DUTARTE, P. (2002). La simulation en statistique, *Repères-IREM*, n°47. p. 93-111.
- DUTARTE, P. et al. (1998). *Simulation d'expériences aléatoires*, Commission inter-IREM Lycées technologiques, IREM Paris Nord.
- DUTARTE, P. et al. (2000). *Simulation et statistique en seconde*, IREM de Paris-Nord.
- GIRARD J.-C. & HENRY, M. (2005). Modélisation et simulation en classe, quel statut didactique ? In B. Chaput et M. Henry (Eds), *Statistique au lycée* volume 1, Brochure APMEP n° 156.
- HASPEKIAN, M. (2003). Entre arithmétique et algèbre : un espace pour le tableur ? Perspectives didactiques et réalités. *Actes du colloque européen ITEM*, Reims, France.
- LAHANIER, D. (1999). Conceptions du hasard et enseignement des probabilités et. statistiques, Paris, P.U.F.
- NABBOU, M. (2006). Enseignement des probabilités en classe terminale au Liban : étude de représentations et de pratiques de professeurs dans des situations aménagées. Thèse de Doctorat, Université Paris 5.
- PICHARD, J.-F. (2005). Théorie des erreurs, courbes en cloche et normalité, In B. Chaput et M. Henry (Eds), *Statistique au lycée* volume 1, Brochure APMEP n° 156.
- PIEDNOIR J.-L. (2006). La statistique Bayésienne, *Bulletin de l'APMEP*, n°464. p. 373-388.
- REGNIER, J.-C. (2002). A propos de la formation en statistique. Approches praxéologiques et épistémologiques de questions du champ de la didactique de la statistique. *Questions éducatives. Revue du Centre de Recherche en Éducation*, n°22-23, p. 157-201.
- WOZNIAK, F. (2005), Conditions et contraintes de l'enseignement de la statistique en classe de seconde générale. Un repérage didactique. Thèse de Doctorat. Université Claude Bernard, Lyon 1.