L'ENSEIGNEMENT DES AFFAIRES WOBURN ET CASTANEDA

Léo Gerville-Réache[1], Vincent Couallier[2]

[1]*Université de Bordeaux, CNRS, UMR 5251, France, leo.gerville-reache@u-bordeaux.fr*
[2]*Université de Bordeaux, CNRS, UMR 5251, France, vincent.couallier@u-bordeaux.fr*

Résumé. Dans cette communication, nous revenons sur deux exercices emblématiques de l'enseignement de la statistique au lycée. Les affaires Woburn et Castaneda sont deux exemples où la justice a usé de la statistique pour se prononcer. Présente dans bons nombres de livres d'élève de première et terminale, l'analyse de ces affaires utilise la simulation numérique et les intervalles de fluctuation pour traiter de la question. Malheureusement, les modalisations et les questions qu'elles posent sont plus complexes qu'il n'y parait.

Mots-clés. Modélisation statistique, groupes comparables, échantillonnage, Woburn, Castaneda

Abstract. In this paper, two exercises, as mostly renown in secondary school's statistical instruction programs, are discussed. The Woburn and Castaneda affairs of justice are two examples where the Court used statistics to adjudicate. As shown in numerous secondary school end's books, analyzing those affairs leads to use numerical simulation and fluctuation intervals to address the issue. Unfortunately, modelization and arising questions are more complex than it looks.

Keywords. Statistical modeling, comparable groups, sampling, Woburn, Castaneda

1 Introduction

Parmi les objectifs de l'enseignement des statistiques et des probabilités au lycée, on peut lire dans le BO précisant le programme de terminale S : "L'étude et la comparaison de séries statistiques menées en classe de seconde se poursuivent avec la mise en place de nouveaux outils dans l'analyse de données. L'objectif est de faire réfléchir les élèves sur des données réelles, riches et variées (issues, par exemple, de fichiers mis à disposition par l'Insee). [...] Étudier une série statistique ou mener une comparaison pertinente de deux séries statistiques à l'aide d'un logiciel ou d'une calculatrice. [...] Exploiter l'intervalle de fluctuation à un seuil donné, déterminé à l'aide de la loi binomiale, pour rejeter ou non une hypothèse sur une proportion."

L'analyse de situations réelles est clairement une nécessité pour l'enseignement des statistiques. Pour autant, certains problèmes d'apparence simple posent de réelles difficultés de modélisation. Sans revenir sur les paradoxes probabiliste ou statistique qui représentent des situations pédagogiques particulièrement intéressantes, nous nous arrêtons dans cette communication sur deux problèmes particulièrement célèbres, enseignés en classe de première et/ou terminale : les affaires Woburn et Castaneda.

Après avoir rappelé les solutions usuelles des deux problèmes proposées par l'ensemble des acteurs de l'enseignement (Professeurs, IREM, Ministère), nous montrons pourquoi l'illustration pédagogique des méthodes statistiques sur ce type de situations concrètes pose plus de difficultés que d'éclaircissements. Nous en concluons que ces deux problèmes, qui semblent emblématiques pour démontrer l'intérêt de la statistique "à tous", ne sont pas adaptés au niveau lycée. En effet, si les solutions proposées sont simples et utilisent bien des compétences attendues, ces solutions ne sont nullement pertinentes et ne répondent en aucun cas aux questions sous-jacentes.

Les affaires Woburn et Castaneda permettent plutôt d'alimenter un débat sur l'utilisation critique du chiffre et des statistiques dans la société mais ne sont pas couvertes par le programme de probabilité et statistiques du lycée.

2 Inquiétudes à Woburn : jusqu'où croire au « hasard » ?

Cette "activité" est proposée dans EduSCOL (2012). C'est un problème présent dans de nombreux livres scolaires et cité par de nombreux acteurs des différents IREM. L'énoncé et la solution qui suivent, conforme aux nombreuses autres sources ont été trouvées dans Dutartre (2007).

Une petite ville des États-Unis a connu 9 cas de leucémies chez de jeunes garçons en l'espace de 10 années. Doit-on en « accuser le hasard » ?

Cet exemple montre l'importance des enjeux de la méthode statistique.

Woburn est une petite ville industrielle du Massachusetts, au Nord-Est des Etats-Unis. Du milieu à la fin des années 1970, la communauté locale s'émeut d'un grand nombre de leucémies infantiles survenant dans certains quartiers de la ville. Les familles se lancent alors dans l'exploration des causes et constatent la présence de décharges et de friches industrielles ainsi que l'existence de polluants. Dans un premier temps, les experts gouvernementaux concluent qu'il n'y a rien d'étrange. Mais les familles s'obstinent et saisissent leurs propres experts. Une étude statistique montre qu'il se passe sans doute quelque chose « d'étrange ».

Le tableau suivant résume les données statistiques concernant les garçons de moins de 15 ans, pour la période 1969-1979 (Source : Massachusetts Department of Public Health).

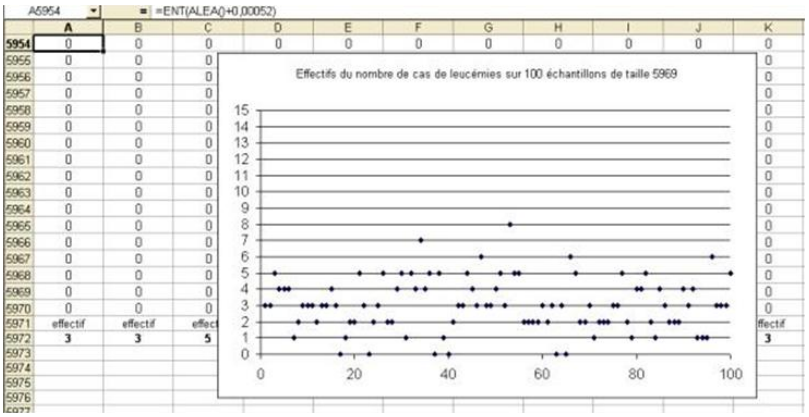
Population des garçons de moins de 15 ans à Woburn selon le recensement de 1970 : n	Nombre de cas de leucémie infantile observés à Woburn entre 1969 et 1979	Fréquence des leucémies aux Etats-Unis p
5969	9	0,00052

La question statistique qui se pose est de savoir si le hasard seul peut raisonnablement expliquer le nombre de leucémies observées chez les jeunes garçons de Woburn, considérés comme résultant d'un échantillon prélevé dans la population américaine.

La population des Etats-Unis étant très grande par rapport à celle de Woburn, on peut considérer que l'échantillon résulte d'un tirage avec remise et simuler des tirages de taille n avec le tableur.

On peut aisément simuler sur le tableur 100 échantillons de taille n = 5969 prélevés au hasard dans une population où p = 0,00052 en utilisant l'instruction : =ENT(ALEA()+0,00052) .

Sur chaque échantillon, en faisant la somme, on obtient le nombre de cas observés.



On peut représenter sur un graphique les 100 résultats observés sur les échantillons simulés.

Les simulations montrent que le nombre de cas observés à Woburn (9 cas) est extrêmement rare (moins de 1 % des simulations), si l'on ne considère que le hasard comme explication.

On ne peut donc pas raisonnablement attribuer au seul hasard le niveau très « significativement » élevé des leucémies infantiles observées chez les garçons de Woburn.

Ce taux anormalement élevé de leucémies est officiellement confirmé par le Département de Santé Publique du Massachusetts en avril 1980. Les soupçons se portent alors sur la qualité de l'eau de la nappe phréatique qui, par des forages, alimente la ville. On découvre ainsi le syndrome du trichloréthylène.

Commentaires :

L'étude montre que si on sélectionne au hasard 5969 individus dans une population où le taux de leucémie est de 0,00052, sur 100 simulations, on n'a observé aucun résultat supérieur à 8. Cette simulation produit une approximation de la probabilité qu'une variable aléatoire binomiale de paramètres p=0,00052 et n=5969 soit supérieur ou égale à 9. Cette probabilité vaut plus précisément 0,0047. Cette probabilité est effectivement petite mais, en quoi ce calcul permet-il de répondre à la question?

Dans l'énoncé, "*la question statistique qui se pose est de savoir si le hasard seul peut raisonnablement expliquer le nombre de leucémies observées chez les jeunes garçons de Woburn, considérés comme résultant d'un échantillon prélevé dans la population américaine.*" Effectivement le calcul précédent répond à cette question, prise dans son ensemble; si on s'accorde pour dire que le nombre de leucémies doit être "*considéré comme résultant d'un échantillon prélevé dans la population américaine"*.

Le problème est que cette considération est parfaitement inadaptée. La modélisation du "hasard" ne correspond aucunement à la situation et ne permet pas de répondre à la question statistique qui "*est de savoir si le hasard seul peut raisonnablement expliquer le nombre de leucémies observées chez les jeunes garçons de Woburn"*. Le problème fondamental de la question de Woburn est donc de comprendre "*comment, en observant un taux de leucémie qui me semble élevé à Woburn, je peux comparer statistiquement ce taux avec le taux national?*".

En l'état des informations, les groupements ne sont pas comparables. Ce concept de comparabilité est un des fondements de l'expérimentation. Dans ce cadre, la randomisation des individus dans les groupes expérimentaux est une condition nécessaire de comparabilité statistique. Dans le domaine de la statistique publique, L'INSEE applique, entre autre, le 14ème principe du code de bonnes pratiques de la statistique européenne :"cohérence et comparabilité".

En 1970, les USA contenaient environs 20 millions de garçon de moins de 15 ans. Cela correspond à environ 3350 groupes de 5969 individus (si toutes les villes des Etats Unis avaient le même nombre d'enfants de moins de quinze ans). L'analyse qui permettrait de donner un début de réponse à la question consisterait à calculer la probabilité d'avoir au moins un groupe de 5969 individus, parmi N groupes, ayant au moins 9 leucémies et de commenter cette probabilité en fonction de N.

Cette probabilité est celle qu'une variable binomiale de paramètre N et p=0,0047 (vu précédemment) ne soit pas nulle. Pour N=100, 200, 500, 1000, 3300, on obtient respectivement 0,38, 0,61, 0,90, 0,99 et 1-1,77^{E-07}. En outre, l'espérance du nombre de groupes, parmi 3350, ayant au moins 9 leucémies est ici de 16!

On voit alors clairement qu'avec un taux de leucémie nationale de 0,00052, le fait qu'une ville de 5969 garçons de moins de 15 ans ait connu 9 leucémies est parfaitement attribuable au hasard. Les phénomènes rares peuvent être, paradoxalement, très fréquents!

Ceci reprend en partie l'activité sur les dates d'anniversaire communes dans une classe de N élèves (décrit dans EduSCOL (2012)). Si la probabilité qu'un de mes camarades ait la même date d'anniversaire que moi est faible, la probabilité qu'il existe deux élèves de ma classe ayant la même date d'anniversaire l'est beaucoup moins. Converti ici : ayant choisi une ville (Woburn) ayant 5969 jeunes garçons, la probabilité de trouver plus de neuf cas de leucémie dans cette ville est faible, mais si je considère qu'il existe plus de 500 villes de plus de 5969 jeunes garçons, la probabilité de trouver au moins une ville dont le nombre de cas de leucémies est supérieures à 9 ne l'est pas (90% de chances) !

Répondre au problème de Woburn n'est pas aisé. En effet, plusieurs problèmes statistiquement délicats sont présents.

- Pourquoi "Woburn"?
- Pourquoi "La leucémie"?
- Pourquoi "Les garçon de moins de 15 ans"?

La réponse semble donnée dans l'énoncé : " *la communauté locale **s'émeut** d'un grand nombre de*

leucémies infantiles survenant dans certains quartiers de la ville". La question a été posée suite à l'observation du nombre de leucémies dans une ville précise et dans une sous population précise. **Il est statistiquement très difficile de répondre à une question sur la réalisation d'une variable aléatoire alors même que la question est posée parce que cette réalisation semble curieuse.**

Afin de ne pas conduire les élèves vers une conclusion erronée ou vers une modélisation bien délicate, il semble que le mieux et de ne plus aborder le problème de Woburn au lycée. Plus généralement, on peut se demander comment le Département de Santé Publique du Massachusetts est-il arrivé à valider cette analyse statistique? On pourrait en outre s'interroger sur le nombre de cas de leucémie observée à Woburn dans la décennie suivante ou chercher à produire une analyse confirmatoire sur des données indépendantes (soit sur une période différente, soir sur un autre groupe d'intérêt). Il existe en fait de nombreuses études plus ou moins rigoureuses de l'effet cancérigène des produits toxiques dans l'eau potable de Woburn, basées par exemple sur un schéma d'échantillonnage cas-témoin, ce qui a soulevé une importante polémique dans la communauté statistique : Cutler et al (1986), Lagakos et al. (1986), MacMahon et al. (1986), Byers et al. (1988), Durant et al. (1995), Costos et al. (1996), Costas et al. (2002). Le problème fondamental ici de la dépendance entre l'acquisition des données et l'engagement de l'analyse statistique est d'ailleurs soulevé dans les commentaires envoyés aux éditeurs de JASA suite à l'article de Lagakos et al (1986).

3 Justice et discrimination : « preuve » statistique

Cette activité est également très connue et très utilisée au lycée. L'énoncé et la solution qui suit, conforme aux nombreuses autres sources sont reprises dans Dutartre (2007).

Comment aux États-Unis, la statistique a-t-elle permis de mettre en évidence une discrimination conduisant à casser un jugement ?

Énoncé distribué aux élèves : *L'affaire Castaneda contre Partida.*

En Novembre 1976 dans un comté du sud du Texas, Rodrigo Partida était condamné à huit ans de prison.

Il attaqua ce jugement au motif que la désignation des jurés de ce comté était discriminante à l'égard des Américains d'origine mexicaine. Alors que 79,1% de la population du comté était d'origine mexicaine, sur les 870 personnes convoquées pour être jurés lors des 11 années précédentes, il n'y eût que 339 personnes d'origine mexicaine.

Produisez votre expertise statistique.

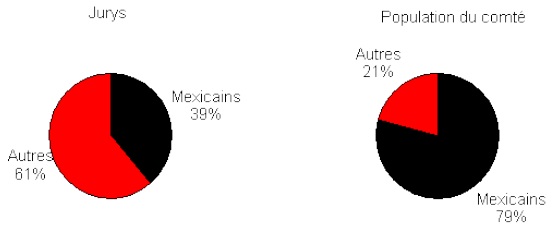
Devant la Cour Suprême, un expert statisticien produisit des arguments pour convaincre du bien fondé de la requête de l'accusé (les juges votèrent à 5 contre 4 en faveur de la requête).

En vous situant dans le rôle de cet expert, produisez à votre tour des calculs, des raisonnements, des graphiques... pour montrer que le hasard ne peut pas « raisonnablement » expliquer à lui seul la sous-représentation des américains d'origine mexicaine dans les jurys de ce comté.

· Vous commencez ce travail en binômes en utilisant les documents disponibles, la calculatrice, le tableur.

· Vous terminez la rédaction (arguments en français, calculs, graphiques...) en devoir individuel, à la maison.

Éléments de réponse : Une première partie du travail sur tableur a consisté en une analyse descriptive des données conduisant les uns ou les autres à produire des tableaux croisés, des histogrammes ou des camemberts.



La seconde partie du travail devait consister à utiliser les moyens informatiques pour montrer que l'écart

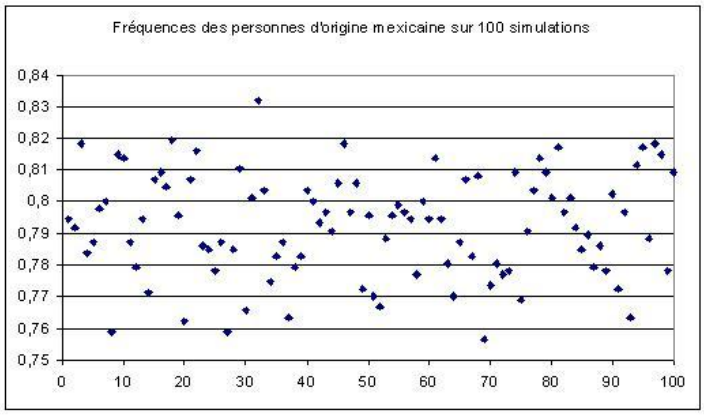
(important) observé ne peut raisonnablement pas s'expliquer par le seul hasard.

Pour cela, on peut simuler 870 tirages au sort dans la population du comté en utilisant la formule =ENT(ALEA()+0,791) .

Certains ont ensuite raisonné en effectifs, d'autres en fréquences de personnes d'origine mexicaine.

On peut vérifier que plus de 95 % des simulations fournissent une fréquence de personnes d'origine mexicaine comprise dans l'intervalle : [p – 1/racine(n) , p + 1/racine(n)] c'est-à-dire [0,791 – 1/racine(870) ; 0,791 + 1/racine(870)] c'est-à-dire environ [0,76 ; 0,82].

Voici un exemple de 100 simulations en fréquences des personnes d'origine mexicaine sur les échantillons de taille 870 :



La fréquence observée des personnes d'origine mexicaine dans les jurys est 0,39 qui est très éloignée de l'intervalle [0,76 ; 0,82]. Jamais les simulations n'ont permis d'observer un résultat aussi bas. Ceci permet de dire que le hasard n'est sans doute pas responsable de la sous-représentation des personnes d'origine mexicaine dans les jurys.

Reste à rechercher les causes de cette sous représentation (l'une des raisons est la nécessité d'une bonne connaissance de l'anglais écrit et parlé).

Commentaires :

La première remarque est similaire à celle largement développée dans l'affaire Woburn. En effet, dans quelle mesure ce comté du sud du Texas est un comté parmi d'autres? Dans quelle mesure la probabilité estimée par simulation montre-t-elle que le "hasard" n'est pas respecté?

Nous nous concentrons ici sur d'autres écueils qui relèvent de l'échantillonnage et de l'interprétation de la statistique.

Concernant l'échantillonnage, le processus de sélection d'un juré n'est pas uniquement basé sur un tirage aléatoire dans la population. Par exemple, il y a un âge minimum, la nécessité d'une bonne connaissance de l'anglais écrit et parlé... On doit alors comparer la proportion de jurés d'origine mexicaine par rapport à la proportion de personnes d'origine mexicaine remplissant les conditions de sélection. Cette question est évoquée en fin d'activité dans la recherche des raisons possibles.

On peut également se demander quel processus les avocats de Partida ont utilisé pour choisir de regarder la proportion de personnes d'origine mexicaine dans les jurys du comté ? Ont-ils choisi une caractéristique du condamné au hasard parmi l'ensemble (très grand) des caractéristiques du condamné ou ont-ils exploré un grand ensemble de ces caractéristiques et choisi celle dont l'écart à l'aléatoire théorique était le plus grand? Ce que la statistique nous dit c'est qu'en cherchant un écart "significatif" à l'aléatoire, on fini toujours par en trouver!

Afin de ne pas conduire les élèves vers une conclusion erronée ou vers une modélisation bien délicate, il semble que le mieux est de ne plus aborder le problème de Castaneda au lycée. Plus généralement, on peut se demander, une fois encore, comment la cours suprême des Etats-Unis a "validé" l'argument de l'expert de Partida. En effet, la décision de la cours suprême pose la question

suivante : considérant que sur les 11 années précédentes, la constitution des jurys de ce comté était discriminatoire, que doit-elle penser de la validité l'ensemble des jugements dans ce comté durant ces 11 années?

4 Conclusion

L'enseignement de la statistique au lycée est particulièrement délicat. Il ne s'agit pas d'une discipline mathématique mais d'une discipline qui utilise les mathématiques. Au lycée, les étapes de modélisation et d'interprétation doivent être au cœur de l'enseignement (Henri 2001).

Les affaires Woburn et Castaneda ne sont que des exemples symptomatiques de la difficulté de modéliser et d'interpréter une situation réelle. L'analyse de ces deux cas ne correspond pas à ce que l'on peut attendre d'un élève de lycée et devrait être retirée des enseignements.

Pour l'enseignant, la fourniture d'activités pédagogiques, permettant de "*montrer l'importance des enjeux de la méthode statistique*" est nécessaire. Peu formé aux statistiques, l'enseignant de mathématique du secondaire doit pourvoir se reposer sur des situations claires et sans ambigüité. Si bon nombre d'exercices proposés dans les livres scolaires ou les IREM sont raisonnablement modélisés et accessibles aux élèves, certains posent réellement question.

Bibliographie

[1] Bulletin officiel (2009). *Programme d'enseignement de mathématiques de la classe de seconde générale et technologique,* BO n°30 du 23 Juillet 2009

[2] Byers VS, Levin AS, Ozonoff DM, Baldwin RW (1988). *Association between clinical symptoms and lymphocyte abnormalities in a population with chronic domestic exposure to industrial solvent-contaminated domestic water supply and a high incidence of leukaemia*. Cancer Immunol Immunother; 27: 77-81.

[3] Costas, K., Knorr, R. S., & Condon, S. K. (2002). *A case–control study of childhood leukemia in Woburn, Massachusetts: the relationship between leukemia incidence and exposure to public drinking water*. Science of the Total Environment, 300(1), 23-35

[4] Costos K, Knorr R. (1996). *Woburn Childhood Leukemia Follow-up Study: Volume I, Analyses. Draft Final Report*. Boston, MA (USA): Massachusetts Department of Public Health.

[5] Dutartre P. (2007). *Présenter aux futur(e)s professeur(e)s une image positive de la statistique et de ses enjeux citoyens*, http://dutarte.perso.neuf.fr/statistique/corfem.htm

[6] Cutler JJ, Parker GS, Rosen S, Prenney B, Healey R, Caldwell GG. (1986). *Childhood leukemia inWoburn, Massachusetts*. Public Health Rep, 101: 201-5.

[7] Durant JL, Chen J, Hemond HF, Thilly WG. (1995). *Elevated incidence of childhood leukemia in Woburn, Massachusetts: NIEHS superfund basic research program searches for causes.* Environ Health Perspect. 103(Suppl 6): 93-8.

[8] EduSCOL (2012). Ressource pour la classe de première - *Probabilité et statistique*. http://cache.media.eduscol.education.fr/file/Mathematiques/59/6/Ressource_Statistiques_Probabilites_1eres_208596.pdf

[9] Henri M. (2001). *Autour de la modélisation en probabilité,* collection didactiques, pufc

[10] Lagakos SW, Wessen BJ, Zelen M. (1986). *An analysis of contaminated well water and health effects in Woburn, Massachusetts*. J Am Stat Assoc. 81: 583-96.

[11] MacMahon B, Prentice RL, Rogan WJ, Swan SH, Robins JM, Whittemore AS. (1986). *Comments and rejoinder on Lagakos, Wessen, and Zelen article on contaminated well water and health effects in Woburn, Massachusetts*. J Am Stat Assoc. 81: 597-614.