

Quelle culture mathématique pour le citoyen ?

Philippe Dutarte.
Académie de Créteil – IREM de Paris-Nord.
philippe-j.dutarte@ac-creteil.fr

Le groupe « statistique et citoyenneté » de l'IREM de Paris-Nord élabore et expérimente en classe des activités à partir de situations concrètes ayant une forte résonance sociologique, avec les objectifs suivants :

- montrer l'utilité d'une formation mathématique pour « décrypter » le monde moderne ;
- développer l'intérêt pour les mathématiques par des activités ayant une signification forte et favorisant l'interdisciplinarité ;
- privilégier l'autonomie des élèves à mettre en place une démarche « scientifique » dans l'analyse d'une situation : formulation d'hypothèse, « construction » d'un « modèle », expérimentation (simulations...), conclusions ;
- initier à l'aléatoire et aux notions de risques.

Nous développons ici deux activités qui peuvent être travaillées dès la troisième, dans le cadre des programmes de 2008.

Tabac et infarctus

Une campagne de la Fédération française de cardiologie donne l'information suivante : « 80% des victimes d'infarctus avant 45 ans sont fumeurs ». En quoi cette « donnée » indique-t-elle que fumer augmente le « risque » d'infarctus ? Il est en fait implicite que les fumeurs sont « sur-représentés » parmi les victimes d'infarctus. Comment peut-on calculer l'augmentation du risque d'infarctus due au tabac ? Quelques notions de mathématiques s'imposent.

Énoncé élève

La Fédération française de cardiologie affiche l'information ci-dessous dans une campagne de presse.

80% des victimes d'infarctus avant 45 ans sont fumeurs

Le tabac est un risque que vous pouvez éviter à votre cœur.
En 2007, les cardiologues font campagne pour un **cœur sans tabac**

La Fédération Française de Cardiologie est une association de cardiologues bénévoles qui ne reçoit aucune subvention de l'État. Notre association est uniquement soutenue par la générosité de ses fidèles donateurs.

www.fedecardio.com informations et dons

Fédération Française de Cardiologie

Société Française de Cardiologie

1. Pourquoi l'information donnée permet-elle de penser que fumer augmente le risque d'infarctus ?
2. Parmi les victimes d'infarctus ayant moins de 45 ans, quelle est la proportion de non-fumeurs ?
3. a) On note i le nombre de cas d'infarctus observés chez les moins de 45 ans. Exprimer en fonction de i le nombre d'infarctus parmi les fumeurs.
- b) On peut estimer qu'en France, parmi les moins de 45 ans, il y a environ 40% de fumeurs ou d'anciens fumeurs.

On note n le nombre de personnes de moins de 45 ans.

Montrer que proportion q d'infarctus parmi les fumeurs est $q = \frac{0,80 \times i}{0,40 \times n}$.

4. Donner, de même, l'expression de la proportion q' d'infarctus parmi les non-fumeurs.

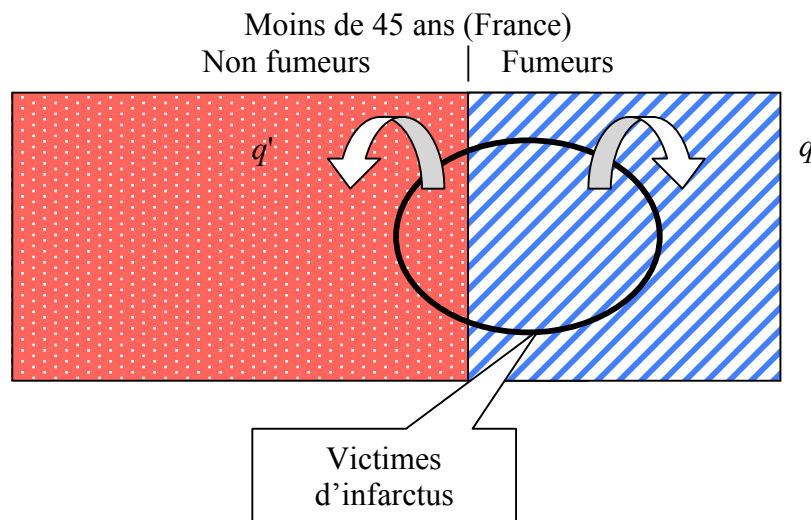
5. Montrer que $\frac{q}{q'} = 6$.

On peut interpréter ce résultat en disant que pour les moins de 45 ans, un fumeur a 6 fois plus de risques d'avoir un infarctus qu'un non-fumeur.

Sources : Fédération française de cardiologie – www.tabac.gouv.fr – INSEE.

Pour la proportion des fumeurs en France, les estimations sont variables : aux alentours de 34% de la population. Le site tabac.gouv.fr affirme, à partir des données de l'INSEE, que la proportion de fumeurs dans la classe d'âge 18-25 ans était de 48% en 2000, 40% en 2003 et est remontée en 2005-2006.

Éléments de réponse



1. On sait qu'il y a moins de 80% de fumeurs. On en déduit que les fumeurs sont sur-représentés parmi les cas d'infarctus.

2. 20%.

3. $q = \frac{0,8 \times i}{0,4 \times n}$.

4. $q' = \frac{0,2 \times i}{0,6 \times n}$.

$$5. \frac{q}{q'} = \frac{0,8 \times 0,6}{0,4 \times 0,2} = 6.$$

Commentaires

Même si les outils mathématiques sont relativement élémentaires (proportions et calcul littéral – on peut s'aider d'un dessin), le résultat obtenu n'est pas évident. Sans ces outils, on ne peut comprendre que 80% et 40% conduisent à un risque multiplié par 6.

Les calculs de proportions menés ici ressemblent, dans leur esprit, à ceux réellement menés en épidémiologie (notion d'odds ratio). Même s'il s'agit d'une simplification des modèles utilisés, cet exercice n'en est pas une trahison.

Remarquons que le « risque » d'infarctus lui-même, qui correspondrait à q (pour un fumeur) et à q' (pour un non fumeur), n'a pas été calculé et est sans doute assez faible (on peut penser que la population n est grande par rapport au nombre d'infarctus i). Pour un risque relativement faible, le message « risque multiplié par 6 » est sans doute plus frappant. Par ailleurs si le risque individuel est relativement faible (et cela se discute car il s'agit d'un « risque évitable » que l'on peut ajouter à d'autres), il n'en est pas de même globalement pour la société, au niveau de la santé publique.

Des naissances à pile ou face

L'exemple suivant, mêlant expérimentation réelle et simulation, peut constituer une initiation aux phénomènes aléatoires (on constate que le « hasard » suit des « lois ») et est propice à l'interdisciplinarité, dans le cadre par exemple des thèmes de convergence.

Énoncé élève (recherche « ouverte »)

On a relevé les données statistiques suivantes :

- En 2000, dans le village de *Xicun* (montagnes du sud de la province de Guangxi en Chine), il est né **20** enfants, parmi lesquels **16** garçons.
- En 2001, à *Louvres* (Val d'Oise), il est né **70** filles et **82** garçons.
- Dans la réserve indienne d'*Aamjiwnaag*, située au Canada à proximité d'industries chimiques, il est né entre 1999 et 2003, **132** enfants dont **46** garçons.

Ces observations sont-elles le fruit du hasard ?

On peut laisser quelque temps les élèves se débrouiller avec ces chiffres et saisir les pistes issues d'une « tempête de cerveaux ». On peut s'attendre à des calculs de statistique descriptive (pourcentage, fréquence), à une comparaison avec une proportion de 50%, et voir émerger la problématique du hasard.

On sait que la proportion de garçons à la naissance est habituellement d'environ 50%. En réalité elle est, de façon étonnamment stable, de 51,2%, mais on peut ici prendre sans inconvénient le modèle « d'une chance sur deux » pour analyser les données proposées.

Une première idée, pour comparer ces trois situations entre elles et les situer par rapport à un a priori de 50%, consiste à calculer le pourcentage de garçons (ou de filles) dans chaque cas :

	Xicun	Louvres	Aamjiwnaag
Pourcentage de garçons	80%	53,9%	34,8%

Dans au moins deux cas, on est loin du résultat « attendu ». Mais peut-être est-ce le hasard ? Pour préciser la réponse à la question posée, il faut étudier les propriétés du

hasard en supposant, pour simplifier, qu'à la naissance on a une chance sur deux d'avoir une fille ou un garçon.

Énoncé élève (expérimentation avec des pièces de monnaie)

1. Lancer 20 fois une pièce de monnaie et noter le nombre de « pile ».
2. Recommencer l'expérience une dizaine de fois ou regrouper les résultats obtenus dans la classe.
3. Comment peut-on utiliser ces expériences pour commenter les statistiques de Xicun ?
4. Pourquoi l'expérimentation avec des pièces ne permet-elle pas de répondre complètement au problème posé ?

Éléments de réponse

3. On peut décider que « pile » correspond à « garçon » et « face » à « fille » (ou le contraire). On constate que pratiquement aucune expérience ne donne 16 piles (ou plus), mais que cela peut se produire.

Le résultat chinois est très étonnant et laisse plutôt penser que « la pièce est truquée ».

Remarque pour le professeur :

On peut considérer que le nombre de pile correspond à la réalisation d'une variable aléatoire X de loi binomiale de paramètres $n = 20$ et $p = 0,5$. On a alors $P(X \leq 15) \approx 0,994$.

Ce résultat peut s'obtenir sur un tableur par l'instruction : =LOI.BINOMIALE(15;20;0,5;VRAI) . C'est-à-dire que l'observation d'un résultat comparable à celui du village chinois a 0,6% de chances de se produire (quand on suppose le hasard « une chance sur deux » à l'œuvre).

4. Pour Louvres et Aamjiwnaag, il faudrait lancer la pièce un trop grand nombre de fois.

On peut également dire qu'un plus grand nombre d'expériences permettrait de s'assurer des observations faites pour 20 lancers.

T.P. de simulation sur tableur – Énoncé élève

1. Simulation d'un lancer de pile ou face.

a) Entrer en cellule A1 la formule =ALEA() (avec des parenthèses vides) puis appuyer de nombreuses fois sur la touche F9.

Compléter la phrase : « La formule =ALEA() affiche un nombre décimal tiré au hasard entre et ».

b) Compléter la phrase : « La formule =ALEA()+0,5 affiche un nombre décimal tiré au hasard entre et ».

Vérifiez votre réponse en entrant la formule =ALEA()+0,5 en cellule A1.

c) En complétant la phrase suivante, trouver une règle permettant de simuler un lancer de pile ou face à l'aide de la fonction ALEA :

« Si en utilisant la formule =ALEA()+0,5 le tableur affiche un nombre de l'intervalle [0 ,[on considère que la pièce est tombée sur face, si au contraire le tableur affiche un nombre de l'intervalle [..... ,[on considère que l'on a obtenu pile. »

d) On souhaite améliorer la simulation en utilisant la fonction ENT qui affiche la partie devant la virgule d'un nombre positif (partie entière).

Entrer en A1 la formule =ENT(ALEA()+0,5) . Faire plusieurs fois F9.

Quels résultats obtient-on ? Quelle est leur signification dans le jeu de pile ou face ?

2. Simulation de 20 lancers de pile ou face.

Recopier le contenu de la cellule A1 jusqu'en A20 (pointeur de la souris en forme de croix noire).

En A21 entrer la formule =SOMME(A1:A20)/20 . Que calcule cette formule ?

3. Simulation de 100 expériences.

Sélectionner les cellules de A1 à A21 puis recopier vers la droite jusqu'en colonne CV.

Sélectionner la ligne 21 (en cliquant sur la tête de ligne) puis cliquer sur l'icône de l'assistant graphique et demander un « nuage de points ».

Faire de nombreuses fois F9.

D'après vos observations, sur 20 lancers, une fréquence de pile égale ou supérieure à 0,8 :

- ne se produit jamais ;
- se produit environ 20 fois sur 100 ;
- se produit environ 10 fois sur 100 ;
- se produit environ 1 fois sur 100.

4. Que pouvez-vous déduire de ces simulations à propos des naissances à Xicun en 2000 ?

5. Cas de 152 naissances.

a) Sur une autre feuille de calcul, simuler 100 expériences de 152 lancers de pile ou face et représenter les fréquences de « pile » comme précédemment.

D'après le graphique, donner un intervalle autour de 0,5 dans lequel se trouvent la « grande majorité » des points (quelques points sur 100 sont en dehors).

b) Que pouvez-vous déduire de ces simulations à propos des naissances à Louvres en 2001 ?

6. Cas de 132 naissances.

a) Sur une autre feuille de calcul, simuler 100 expériences de 132 lancers de pile ou face et représenter les fréquences de « pile » comme précédemment.

D'après le graphique, donner un intervalle autour de 0,5 dans lequel se trouvent la « grande majorité » des points.

b) Que pouvez-vous déduire de ces simulations à propos des naissances à Aamjiwnaag ?

Éléments de réponse

1. a) « La formule =ALEA() affiche un nombre décimal tiré au hasard entre 0 et 1 ».

b) « La formule =ALEA()+0,5 affiche un nombre décimal tiré au hasard entre 0,5 et 1,5 ».

c) « Si en utilisant la formule =ALEA()+0,5 le tableur affiche un nombre de l'intervalle [0,5 ; 1[on considère que la pièce est tombée sur face, si au contraire le tableur affiche un nombre de l'intervalle [1 ; 1,5[on considère que l'on a obtenu pile. »

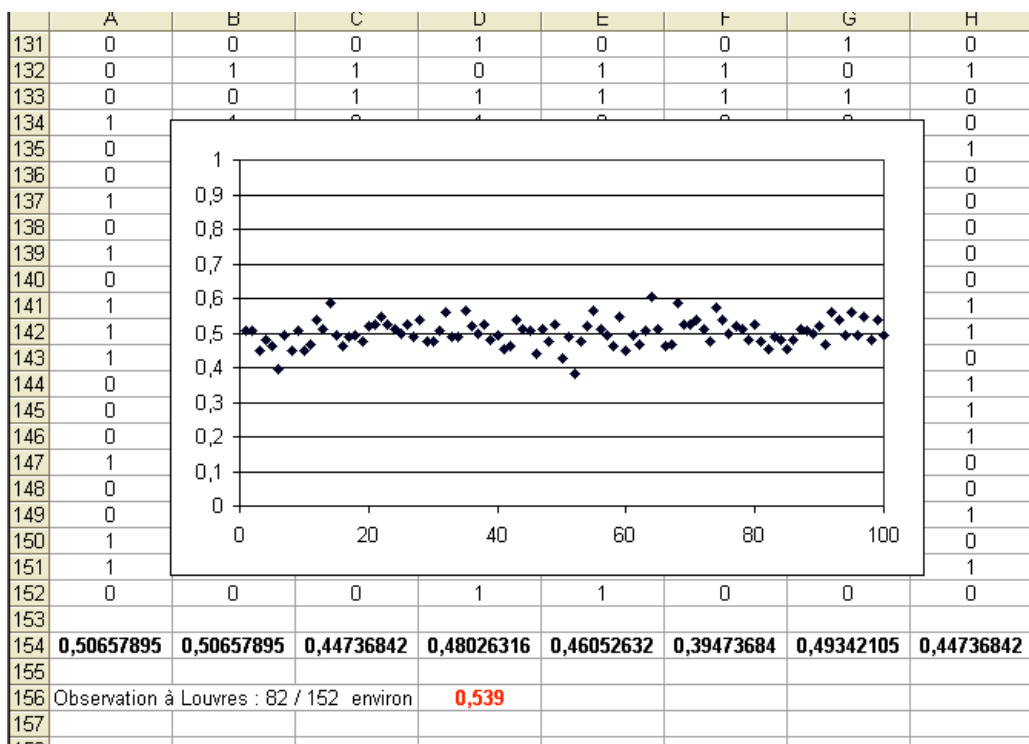
d) On obtient comme résultats 0 ou 1. Le résultat 0 correspond à face et 1 correspond à pile.

2. La formule calcule la fréquence des « pile » sur les 20 lancers.

3. D'après les observations, sur 20 lancers, une fréquence de pile égale ou supérieur à 0,8
P se produit environ 1 fois sur 100.

4. On peut en déduire que les statistiques de Xicun sont très étonnantes. Il est peu probable qu'elles s'expliquent par le seul hasard.

5. a) La grande majorité des points sont dans l'intervalle [0,4 ; 0,6].



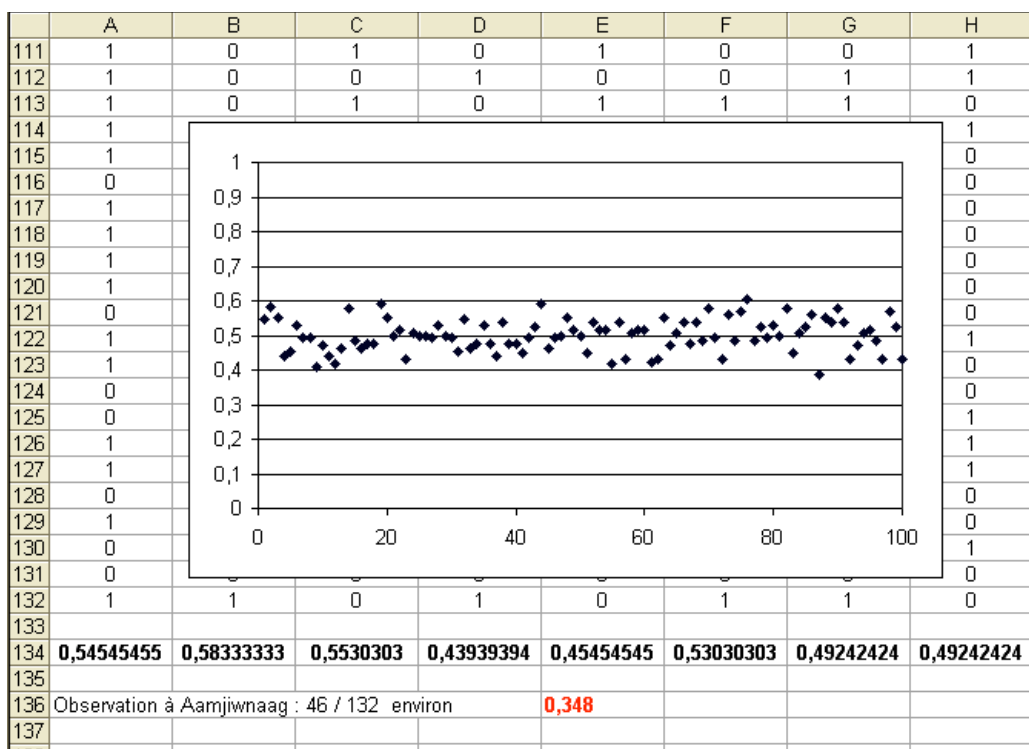
Remarque pour le professeur :

On peut calculer un intervalle de variabilité « naturelle », au niveau de 95% : $[0,5 - \frac{1}{\sqrt{152}} ; 0,5 + \frac{1}{\sqrt{152}}]$. Ce qui donne environ $[0,42 ; 0,58]$. C'est l'occasion

d'observer que le « hasard » obéit à des règles de calcul.

b) Le résultat 0,539 observé à Louvres se trouve dans l'intervalle $[0,4 ; 0,6]$. Il peut facilement s'expliquer par le hasard.

6. a) La « grande majorité » des points sont dans l'intervalle $[0,4 ; 0,6]$.



b) Le résultat 0,348 observé à Aamjiwnaag ne se trouve pas dans l'intervalle [0,4 ; 0,6]. Il est peu probable qu'il s'explique par le seul hasard.

Commentaires

Il est important de préciser que les « réponses » apportées par l'activité précédente ne sont que des réponses « statistiques ». Les résultats observés sur les naissances à Xicun et Aamjiwnaag sont « bizarres » (et préoccupants). On ne peut rien dire de plus quant aux causes, mais ces résultats doivent inciter à enquêter. C'est une obligation morale, car on a établi une « preuve statistique », rationnelle, qu'il se passe sans doute quelque chose d'inhabituel.

Pour le cas de Xicun, la cause probable est l'acquisition dans ce village (en 1999) d'une machine à ultras-sons bon marché, « made in China », permettant aux médecins de déterminer le sexe du fœtus. (Source : *Washington Post* du 29 mai 2001.)

Dans le cas d'Aamjiwnaag, une enquête sanitaire est menée. On connaît en effet, depuis Seveso, le rôle de certains polluants sur les déséquilibres du sex-ratio. (Sources : *Science et Vie* février 2006 – *Environmental Health Perspectives* octobre 2005 – Article en ligne.)

Bibliographie

- Brochures APMEP :
Brochure n° 156 – *Statistique au lycée volume 1* – APMEP 2005.
Brochure n° 167 – *Statistique au lycée volume 2* – APMEP 2007.
- Brochures IREM Paris-Nord :
Brochure n° 102 – *Simulation et statistique en seconde* – 2000.
Brochure n° 112 – *Enseigner la statistique au lycée : des enjeux aux méthodes* – 2001.
- Philippe Dutarte – *L'induction statistique au lycée illustrée par le tableur* – Didier 2005.
- Claudine Schwartz – *Pratiques de la statistique* – Vuibert 2006.