Politique nataliste : Analyse d'un devoir sur une « régulation des naissances »

Jacques Verdier(*)

J'ai donné en classe de première STL Biochimie-Génie biologique(1) le devoir à la maison ci-dessous, inspiré de la fiche « Politique nataliste » éditée par le GEPS (ex-GTD) pour la classe de seconde.

Première STL. Devoir à la maison

Dans un pays imaginaire, le roi décide de la politique de limitation des naissances et décrète ce qui suit : *les familles devront cesser de procréer dès qu'elles auront eu un garçon ; en tout état de cause, les familles qui auront eu 4 filles ne devront plus procréer non plus*.
La question qui se pose est la suivante : **cette politique favorise-t-elle la naissance des garçons ?**

Pour résoudre ce problème, on fera l'hypothèse qu'à chaque naissance la détermination du sexe est aléatoire, que les deux possibilités « GARCON » et « FILLE » sont équiprobables, et qu'il n'y a pas de naissances gémellaires.

<u>Aide</u> : d'après le décret royal, les seules possibilités sont « G », « F G », « F F G », « F F F G » et « F F F F ». On calculera la probabilité de chacun de ces cinq événements.
On pourra éventuellement faire une simulation avec une pièce (Face = Fille, Pile = Garçon), si cela peut aider.
[On pourra également calculer, si cela peut aider, « l'espérance » de chacun des types de familles sur un très grand nombre de naissances](2).

L'exactitude des justifications et la qualité de leur rédaction représentent des critères importants qui seront pris en compte dans l'évaluation de la copie. Si cela est nécessaire, on « illustrera » le raisonnement par une figure, un arbre, un graphique, …

Ce devoir a été donné après avoir traité le chapitre de probabilité. J'ai énormément utilisé les arbres comme outil de travail au cours de ce chapitre (en particulier lorsque l'expérience aléatoire est décomposable en parties chronologiques successives,

(*) Lycée Arthur Varoquaux, 54-TOMBLAINE.
(1) C'est une classe qui ne peut être considérée comme « scientifique » (du moins en ce qui concerne les mathématiques) ; le programme est sensiblement équivalent à celui de Première STT-Gestion.
(2) Cette phrase a été supprimée dans la version 2003 de l'énoncé. Voir ci-après pourquoi.

comme lors des tirages de plusieurs boules dans une urne, ce qui est le même modèle qu'ici).

J'ai recensé, pendant deux années de suite, les réponses des élèves à ce devoir : 52 copies au total (23 copies sur 26 élèves en 2002, 29 copies sur 30 élèves en 2003).

Analyse des réponses

En 2002, tous les élèves avaient conclu que cette politique favorisait la naissance des filles ; ceci peut s'expliquer en partie par le fait que la moitié des élèves sont internes, et que les « rumeurs » circulent très vite. En 2003, les réponses sont beaucoup plus mitigées : 13 ont conclu que cette politique favorisait la naissance des filles, 8 qu'elle favorisait celle des garçons, et 7 qu'il y avait égalité.

En 2002, tous les élèves sauf un ont calculé les probabilités proposées dans le paragraphe « aide » de l'énoncé, et seulement 20 (sur 29) en 2003. Ces probabilités sont, pour mémoire, $p(G) = 1/2$; $p(FG) = 1/4$; $p(FFG) = 1/8$; $p(FFFG) = 1/16$ et $p(FFFF) = 1/16$.

Parmi les 9 qui n'ont pas calculé ces probabilités, tous on fait une simulation (en jetant des pièces), mais certains n'ont pas tenu compte du nombre total de garçons et de filles ainsi trouvé pour répondre (cf. infra). Un élève a cru calculer des probabilités : il a pris pour telles les fréquences trouvées dans sa simulation.

En 2002, 22 élèves avaient calculé correctement les probabilités, et seulement 13 en 2003 (bien que cela soit demandé par l'énoncé : ils ont peut-être estimé que ce calcul ne leur permettrait pas de répondre, et que seule la statistique allait les « sauver »…).

Parmi les 8 élèves qui ont calculé des probabilités mais qui ont trouvé un résultat faux :
- 2 ont postulé l'équiprobabilité des cinq événements proposés ($p = 1/5$) ;
- 2 ont trouvé cette équiprobabilité à partir d'un arbre faux ;
- 2 ont trouvé respectivement 1/2, 1/4, 1/6, 1/8 et 1/8 ;
- 1 a calculé, à partir d'un arbre totalement faux, que $p(G) \approx 1{,}17$ et $p(F) \approx 1{,}17$;
- 1 a trouvé respectivement 1/1, 1/2, 1/3, 1/4 et 1/5.

Parmi les 13 élèves de 2003(3) qui ont calculé les probabilités correctement :
- 3 n'arrivent pas à conclure (et s'arrêtent là, car ils n'ont pas fait de simulation) ;
- 3 concluent que les filles sont favorisées, à 10 contre 4 (voir ci-après en encadré le raisonnement d'Élodie) ;
- 1 conclut à 68% de garons contre 32% de filles (voir ci-après en encadré le tableau extrait de la copie de Camille) ;
- 4 concluent à 73 % de filles contre 27 % de garçons (voir ci-après, en encadré, le raisonnement de Guillaume) ;

(3) En 2002, je n'avais pas fait une analyse aussi fine des divers raisonnements ; cependant les raisons invoquées pour montrer que la politique favorisait la naissance des filles n'étaient pas explicites, sauf cinq exceptions : trois qui avaient utilisé l'espérance (voir en fin d'analyse) et deux qui donnaient les filles gagnantes à 10 contre 4 (même raisonnement qu'Élodie).

– les 2 derniers n'utilisent que les résultats de leur simulation, ne tenant aucun compte des probabilités qu'ils ont calculées.

Extrait de la copie d'Élodie

Sur les 14 naissances possible, FFFF ; FFFG ; FFG ; FG ; G :
- Pour la naissance de 4 filles aucun garçon ne né.
- Pour la naissance de 3 filles, 1 garçon né.
- Pour la naissance de 2 filles, 1 garçon né
- Pour la naissance d'1 fille, 1 garçon né
- Pour la naissance d'aucune fille, 1 garçon né

Donc il y a 10 filles sur 14 contre 4 garçons sur 14 naissances.
On peut donc conclure que cette politique ne favorise pas la naissance des garçons.
Lorsque 10 filles naissent seulement 4 garçons naissent.

Extrait de la copie de Camille

	G	FG	FFG	FFFG	FFFF	total
fille	0	12,5	8,33	4,6875	6,25	31,4675
garçon	50	12,5	4,17	1,5625	0	68,2325
total	50	25	12,5	6,25	6,25	100

Raisonnement de Guillaume

Une fois les probabilités des cinq événements (« G », « FG », « FFG », « FFFG », « FFFF ») calculées correctement à l'aide d'un arbre, il donne le tableau suivant, sans aucune explication :

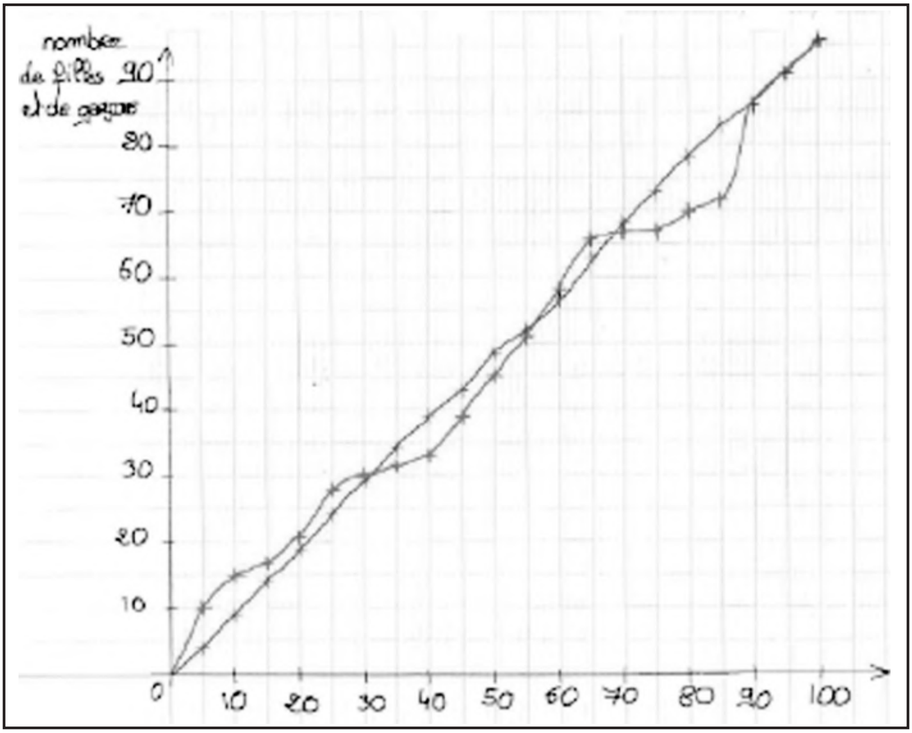
	G	F
G	100 %	0 %
FG	50 %	50 %
FFG	33,33 %	66,67 %
FFFG	25 %	75 %
FFFF	0 %	100 %
Total	27,08 %	72,92 %

Et conclut bien évidemment : « *Donc, comme la proportionnalité de G au total est à peine 1/4, et que la majorité est des filles, je conclus que cette politique ne favorise pas la naissance des garçons, mais celle des filles* ».

Autant il est facile d'interpréter les pourcentages des cinq premières lignes, autant je m'interroge sur la façon dont a pu être obtenu le « total »…

Analysons maintenant les démarches des 17 élèves(4) qui ont fait une simulation : elles vont de 10 familles à 275 familles (523 lancers de la pièce !).
Trois d'entre eux n'ont pas compté le nombre de garçons et de filles dans leurs statistiques … ce qui les rendait inutiles ; mais comme ils avaient calculé les probabilités, et que la plus forte de ces probabilités était p(« G ») = 1/2, et que cela était corroboré par l'expérimentation, ils en ont déduit que cette politique favorisait la naissance des garçons !

Restent 14 élèves qui ont tenu compte des résultats de leur simulation pour conclure (dont 8 l'on fait en traçant un graphique comme celui d'Aurélie, ci-dessous).



Conclusion d'Aurélie :

D'après les courbes, on constate qu'elles se superposent presque (...). Donc ce système ne favorise pas une naissance en particulier. Il y aurait autant de filles que de garçons.

Neuf d'entre eux concluent en fonction des résultats de leur simulation (cela favorise les naissances de garçons si leur simulation a donné plus de garçons que de filles, et vice versa).
Voici les conclusions de trois autres élèves, qui sont intéressantes :

(4) 15 en 2003 et seulement 2 en 2002.

Conclusion de Romain :
Sur les 100 familles, on compte : 93 garçons et 105 filles.
Donc d'après mes résultats, le décret ne favorise pas la naissance des garçons, mais équilibre à peu près les naissances.

Conclusion de Yazid :
La simulation porte sur 10 familles. Il a trouvé ainsi 9 filles et 10 garçons. Il conclut ceci :
Ce n'est qu'une simulation et ça ne donne pas forcément le résultat exact. Normalement, le nombre de filles est toujours supérieur au nombre de garçons.
Cette dernière partie de la conclusion n'est justifiée par aucun autre calcul ni raisonnement.

Raisonnement de Julien :
J'ai pris une pièce et je l'ai lancée (pile = garçon, face = fille). Je l'ai lancée jusqu'à ce que j'obtienne 30 familles [suit la liste des lancers]. *Je l'ai lancée 65 fois : 27 garçons et 38 filles.*
Admettons que je n'aie que des garçons : pour les 30 familles, cela aurait fait 30 garçons.
Admettons que je n'aie que des filles : pour les 30 familles, cela aurait fait 120 filles.
Donc cela favorise la naissance des filles.

En 2002, trois élèves avaient utilisé l'aide proposée en fin d'énoncé : « *On pourra également calculer, si cela peut aider, « l'espérance*(5) *de chacun des types de familles sur un très grand nombre de naissances* », phrase que j'avais supprimée dans la version 2003.
Le premier a travaillé sur 24 000 familles, et a trouvé le nombre « espéré » de familles correspondant à chacune des issues. Par exemple : 3 000 familles « FFG ». Mais là, au lieu de considérer que cela faisait 6 000 filles et 3 000 garçons, il a trouvé 2 000 filles et 1 000 garçons (il a partagé les 3 000 proportionnellement au nombre de lettres de FFG) ; c'est le même type de raisonnement que celui de Camille, cf. supra.
Les deux autres ont travaillé respectivement sur 10 000 et 1 000 000 de familles, ont calculé correctement le nombre de filles « espérées » (respectivement 9 375 et 937 500), mais ont calculé le nombre de garçons en retranchant le nombre de filles du nombre de familles (et ont alors trouvé respectivement 625 et 62 500). Leur conclusion est alors évidente !

(5) Il s'agissait d'une notion non définie en classe (elle n'est pas au programme, et il n'était pas question d'introduire les variables aléatoires), mais utilisée dans son sens courant : comme quand on dit que si on lance 1 200 pièces, on « **espère** » (a priori) 600 piles et 600 faces. C'est d'ailleurs ainsi que les trois élèves cités l'ont comprise. Cependant la formulation pouvant être ambiguë ou équivoque, je n'ai pas proposé cette aide en 2003.

Conclusion

On peut constater, sur cet exemple, combien peut être difficile l'apprentissage des probabilités pour des élèves qui n'ont aucune *culture de l'aléatoire*. Même la loi des grands nombres, qui semble implicitement acquise à cet âge-là(6), et qui est à la base de programme de Première STL, pose encore problème : certains élèves peuvent, dans la même copie, faire un calcul de probabilité et une simulation statistique dont chacun contredit l'autre (avec un seuil de risque infime).
D'autres, malgré les simulations faites en classe(7) avec des pièces ou des dés, utilisent les résultats de leur simulation pour conclure, ne pensant plus que les écarts observés entre garçons et filles (généralement faibles) peuvent être le fruit du hasard.

Correction en classe

Je me suis appuyé sur les raisonnements des trois élèves évoqués en fin d'analyse, en fournissant au rétroprojecteur le tableau ci-dessous (le rétroprojecteur me permettant de ne dévoiler les lignes qu'au fur et à mesure que ce qui précède est bien assimilé). Les deux dernières lignes n'étaient pas indispensables, mais j'en ai profité pour leur apporter ainsi la notion d'espérance mathématique(8).

<table>
<tr><td colspan="2">Issues</td><td>G</td><td>FG</td><td>FFG</td><td>FFFG</td><td>FFFF</td><td colspan="2">Totaux</td></tr>
<tr><td colspan="2">Probabilités</td><td>$\frac{1}{2}$</td><td>$\frac{1}{2}$</td><td>$\frac{1}{2}$</td><td>$\frac{1}{2}$</td><td>$\frac{1}{2}$</td><td colspan="2">1</td></tr>
<tr><td rowspan="3">Espérance théorique sur 10 000 familles</td><td>Nb. familles</td><td>5 000</td><td>2 500</td><td>1 250</td><td>625</td><td>625</td><td>10 000</td><td></td></tr>
<tr><td>Nb. Filles</td><td>0</td><td>2 500</td><td>2 500</td><td>1 875</td><td>2 500</td><td>9 375</td><td rowspan="2">Total 18 750 enfants</td></tr>
<tr><td>Nb. Garçons</td><td>5 000</td><td>2 500</td><td>1 250</td><td>625</td><td>0</td><td>9 375</td></tr>
<tr><td rowspan="3">Espérance théorique sur N familles</td><td>Nb. familles</td><td>$\frac{N}{2}$</td><td>$\frac{N}{4}$</td><td>$\frac{N}{8}$</td><td>$\frac{N}{16}$</td><td>$\frac{N}{16}$</td><td>N</td><td></td></tr>
<tr><td>Nb. Filles</td><td>0</td><td>$\frac{N}{4}$</td><td>$\frac{2N}{8}$</td><td>$\frac{3N}{16}$</td><td>$\frac{4N}{16}$</td><td>$\frac{15N}{16}$</td><td rowspan="2">Total $\frac{15N}{8}$ enfants</td></tr>
<tr><td>Nb. Garçons</td><td>$\frac{N}{2}$</td><td>$\frac{N}{4}$</td><td>$\frac{N}{8}$</td><td>$\frac{N}{16}$</td><td>0</td><td>$\frac{15N}{16}$</td></tr>
<tr><td rowspan="2">Espérance théorique sur 1 famille</td><td>Filles</td><td>0</td><td>$\frac{1}{4}$</td><td>$\frac{2}{8}$</td><td>$\frac{3}{16}$</td><td>$\frac{4}{16}$</td><td>$\frac{15}{16}$</td><td rowspan="2">$\frac{15}{8}$</td></tr>
<tr><td>Garçons</td><td>$\frac{1}{2}$</td><td>$\frac{1}{4}$</td><td>$\frac{1}{8}$</td><td>$\frac{1}{16}$</td><td>0</td><td>$\frac{15}{16}$</td></tr>
</table>

(6) Cf. brochure de l'IREM de Lorraine en bibliographie.
(7) Dans la mesure où elles sont effectivement faites. À part une exception, aucun de mes élèves n'avait travaillé en Seconde sur les fluctuations d'échantillonnage ; et l'exception ne consistait qu'en une seule activité : étude de la fréquence d'apparition des divers chiffres quand on observait le nombre affiché après appui sur la touche **Random** de sa calculatrice.
(8) La notion d'espérance n'est pas au programme de Première, mais à celui de Terminale. Je n'avais utilisé ce mot qu'une seule fois dans un exercice, à propos d'une espérance de gain dans une loterie, et en travaillant sur la totalité des billets (ce qui permettait de calculer le gain total et de le comparer à la dépense totale des joueurs).

Puisque bon nombre d'élèves avaient fait une simulation, je leur ai donné les résultats d'une simulation que j'ai moi-même faite, portant sur cinq séries de 1 000 (voir en annexe).

Cependant il faut que les élèves comprennent bien la différence entre la simulation et le calcul des probabilités, ce qui avait été le « pilier » des activités faites en classe, sur des exemples beaucoup plus élémentaires (jet d'une pièce, jet de 2 pièces, jet d'un dé, …) : elle est de la même nature que la différence, en géométrie, entre un objet mathématique et sa représentation par une figure matérielle. Dans la situation présentée ici (qui nécessitait d'« inventer » une démarche de résolution, puisque ce n'était pas du tout un exercice d'application), la simulation permettait d'énoncer une hypothèse. Le problème a été que certains élèves qui, lorsqu'ils avaient lancé 100 fois une pièce, obtenu 47 piles et 53 faces, avaient conclu que rien n'empêchait que le hasard ait pu donner ce résultat (et n'avaient donc pas remis en cause le modèle de pièce « équilibrée »), n'ont pas eu – contrairement à Romain – le même comportement en cherchant à résoudre ce problème : n'arrivant pas à calculer de probabilités, ils se sont contentés des résultats de la simulation…

Solution GEPS

J'ai proposé à mes élèves la « solution » par simulation proposée dans la fiche du GEPS (ex GTD).

Cela revient à tirer au hasard (à pile ou face) une série de F et G, par exemple :

F G G G F F G F G G F F G G G F F F G G F F G F F G G F G G F F F G F G G F G G …

et à fabriquer à partir de cette liste les « familles », en séparant dès que l'on a un garçon, ou 4 filles :

F G | G | G | F F G | F G | G | F F F G | G | F F F G | G | F F G | F F G | G | G G | G | F F F F | F G | G | F G | G | …

Cette démarche a été **unanimement refusée** par la classe, d'une part parce qu'elle venait après une démarche qu'ils avaient fini par comprendre et admettre (et ne voulaient donc pas remettre en cause), d'autre part par ce qu'elle impliquait pour eux une vision « chronologique » des choses (en ce sens que la deuxième famille devait attendre que la première soit constituée pour se constituer elle-même : c'est « ma » traduction de leurs remarques). Il s'agit pourtant bien, ici, du même modèle théorique sous-jacent : mais il n'était pas question d'aller sur ce terrain avec des élèves de Première STL !

Bibliographie :

Pascale POMBOURCQ, « Les statistiques dans le programme de seconde à la rentrée 2000 », Brochure A.P.M.E.P. n° 130, pages 56-57.

M.E.N., C.N.D.P. et G.T.D.-Mathématiques, fiche « Politique nataliste », extraite de « Onze fiches de statistiques », juin 2000, disponible sur
www.cndp.fr/gtd_maths/pdf/ESEMA004.pdf
ou sur
www.ac-strasbourg.fr/dynamic_html/158/350.html

I.R.E.M. de Lorraine, « L’enseignement des probabilités au collège et au lycée : exemples européens et propositions », 2001, I.S.B.N. 2-85406-168-3.

Paul-Louis HENNEQUIN, « Fille ou garçon », Hors série Tangente nº 9 (juillet 2000).

ANNEXE : SIMULATIONS SUR ORDINATEUR

Voici l’un des simulations proposées lors de la correction :

Nombre de simulations : 1000
Nombre de familles de type [G] : 510
Nombre de familles de type [FG] : 241
Nombre de familles de type [FFG] : 119
Nombre de familles de type [FFFG] : 54
Nombre de familles de type [FFFF] : 76
Nombre total d’enfants = 1869
Nombre total de filles = 945 (50,6 %)
Nombre total de garçons = 924 (49,4 %)