Travaux personnels encadrés : la vache folle

Daniel Daviaud(*)

Ce TPE débute en classe de Première S en octobre 2000. Trois élèves (Éva Daviaud, Laurène Marcelin et Claire Perrier) choisissent un sujet : la vache folle. Les deux disciplines concernées sont les SVT (ce qui va de soi) et les Maths (pour les aspects quantitatifs de l'épidémie).

J'informe les élèves qu'un professeur de l'université de Poitiers, Bernard Guennebaud, a développé des modèles mathématiques sur la propagation des maladies contagieuses et qu'il serait sans doute utile de s'y intéresser. B. G. a exposé son travail lors de la précédente assemblée générale de la Régionale de l'APMEP, et c'est là que j'en ai pris connaissance.

Mais y a t-il un modèle déjà prêt pour l'Encéphalopathie Spongiforme Bovine ? J'écris à B. G. pour lui demander si c'est vraiment une maladie « contagieuse » ? Il m'explique alors que, d'un certain point de vue, l'ESB peut être considérée contagieuse. Cependant, je me rends compte que cette maladie se distingue beaucoup des autres. D'une part, il paraît impossible de lier le nombre d'animaux nouvellement contaminés au nombre d'animaux déjà malades (car la maladie ne s'acquiert pas par la fréquentation des malades). D'autre part, la durée d'incubation semble varier selon les individus. Interrogé à ce sujet, B. G. me répond : « Il faut sans cesse modifier, adapter les modèles mathématiques. Les modèles que j'ai proposés sont des idées, pas des dogmes ». Ce conseil stimulant m'a encouragé à voler de mes propres ailes.

Pourquoi, demanderont certains, cette correspondance a-t-elle été « confisquée » par le professeur ? Ne revient-il pas aux élèves d'écrire au spécialiste ? Eh bien Non ! Pas dans ce cas, à mon avis. Car, dans un premier temps, les élèves se consacrent à l'étude biologique de la maladie … et tout n'est pas mâché d'avance. Ils ont aussi pour objectif de recueillir, sur Internet en l'occurrence, des données épidémiologiques quantitatives. Ils ont déjà fort à faire. Mais surtout, la discussion sur le choix et l'élaboration d'un modèle est prématurée puisqu'ils ignorent encore ce qu'est une modélisation ; ils l'apprendront au cours du TPE.

En conséquence, c'est l'enseignant qui leur propose, sous forme d'une activité dirigée, le modèle qu'il a conçu. J'utilise les notions mathématiques les plus simples : suite géométrique plutôt que fonction exponentielle car les élèves ne sont qu'en première ; pourcentages plutôt que probabilités (pour la même raison) alors qu'un modèle probabiliste serait peut-être opportun.

Je dois rédiger cette activité dirigée. Sinon, comment s'occuper simultanément de onze équipes d'élèves travaillant sur onze sujets différents ? Mon texte, écrit rapidement à la main, alterne de courts passages qui éclairent la démarche de modélisation, et, comme dans un problème, des questions de maths ou d'informatique, mais résolument aussi ouvertes que possible. Les élèves effectuent le

(*) Professeur au Lycée Jean Hyppolite, Jonzac.

travail proposé et, un peu pris par le temps, répondent séquentiellement aux questions, en ajoutant quelques commentaires.

En octobre 2001, les élèves sont passées en terminale, l'équipe est dispersée. Seule, Éva D. reprend le sujet. Elle veut approfondir certaines questions de biologie. Elle veut aussi profiter des données récentes pour vérifier l'efficacité prédictive du modèle et l'affiner un peu. Je lui demande aussi de reprendre intégralement la rédaction du dossier pour le présenter comme un article de revue scientifique, en mettant en valeur les problématiques abordées. Dans cette étape, je joue un rôle d'éditeur : relecture d'épreuves et proposition de corrections. Au final, ce travail de mise en forme aboutit, à quelques retouches près, à l'exposé qui suit.

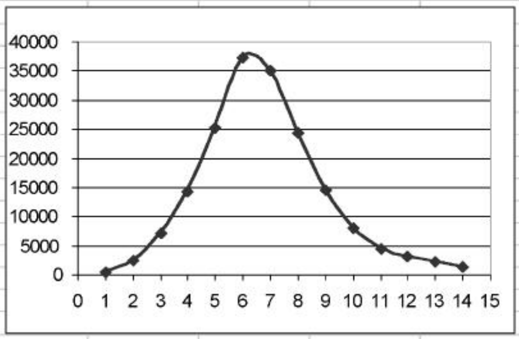
Précisons ici le cadre expérimental. Pour la partie mathématique de ce TPE, nous avons étudié exclusivement la propagation de l'ESB au Royaume-Uni. En effet, la maladie s'y est installée, à l'insu de tous, pendant une longue période, sans être contrariée par aucune mesure prophylactique. D'où l'apparition de cohortes nombreuses de vaches malades dont on peut espérer abstraire une loi d'évolution.

Comme l'a écrit Denis Gardes dans son article sur les TPE paru dans le bulletin n° 440 de mai-juin 2002, « le rôle de l'enseignant a été multiforme ». On est très loin de la définition étroite et inefficace donnée dans le BO du 11/01/2001 : « Les enseignants accompagnent les étapes du travail des élèves en leur prodiguant recommandations, avis et conseils ». Si le prof de maths ne dépasse pas ce niveau d'engagement, les élèves peuvent-ils produire autre chose que des graphiques banals accompagnés de calculs rudimentaires ?

A. La propagation de l'E.S.B. obéit-elle à une loi mathématique ?

Chaque année, le nombre des cas déclarés d'E.S.B. a été publié. Nous avons fondé notre travail sur les données recueillies à l'adresse www.oie.int/fr/info/fr_esbru.htm. Ce site web est régulièrement mis à jour. Nous l'avons consulté une dernière fois le 11/02/2002 pour rédiger ce rapport avec des données récentes.
L'objectif était de rechercher un modèle mathématique capable de décrire la propagation de la maladie. Afin de mieux visualiser l'évolution, nous traçons la courbe suivante :

Année	n	Cas d'ESB
1987	1	446
1988	2	2514
1989	3	7228
1990	4	14407
1991	5	25359
1992	6	37280
1993	7	35090
1994	8	24436
1995	9	14562
1996	10	8149
1997	11	4393
1998	12	3235
1999	13	2300
2000	14	1443
	To	180842



Nous notons M_n le nombre de cas déclarés au cours de l'année n.

Les points du graphique ne sont pas alignés : la suite (M_n) n'est donc pas arithmétique, même « localement ».

n	Cas d'ESB : Mn	Quotients successifs
1	446	5,637
2	2514	2,875
3	7228	1,993
4	14407	1,760
5	25359	1,470
6	37280	0,941
7	35090	**0,696**
8	24436	**0,596**
9	14562	**0,560**
10	8149	**0,539**
11	4393	**0,736**
12	3235	**0,711**
13	2300	**0,627**
14	1443	

Serait-elle géométrique dans sa partie croissante ou dans sa partie décroissante ? Pour le savoir, nous calculons les quotients permettant de passer de M_n à M_{n+1}.

Les quotients M_{n+1}/M_n, calculés pour $1 \leq n \leq 13$, paraissent à peu près constants entre les indices 7 et 13, c'est-à-dire dans la partie décroissante.

On peut alors considérer que la suite (M_n) est « presque » géométrique dans cet intervalle, avec une raison voisine de 0,6.

Grâce à la fonction LOGREG du logiciel EXCEL®, on peut trouver la raison et le premier terme (celui dont l'indice est 0) de la suite géométrique « la plus proche » de la suite (M_n) :

n	Cas d'ESB : Mn
7	35090
8	24436
9	14562
10	8149
11	4393
12	3235
13	2300
14	1443

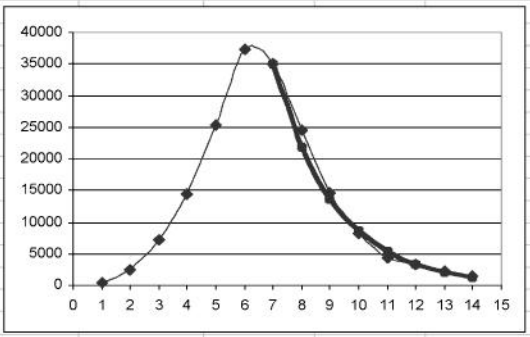
Raison	Premier terme
0,626	928561

La raison a été arrondie à 0,001 près et le premier terme à l'unité près.

La suite géométrique (M'_n) telle que $M'_n = 928\,561 \times 0{,}626^n$ approche (M_n) de $n = 7$ à $n = 13$.

Le graphique suivant permet d'apprécier visuellement la valeur de ce modèle.

Année	n	Mn	M'n
1987	1	446	
1988	2	2514	
1989	3	7228	
1990	4	14407	
1991	5	25359	
1992	6	37280	
1993	7	35090	**34981**
1994	8	24436	**21898**
1995	9	14562	**13708**
1996	10	8149	**8581**
1997	11	4393	**5372**
1998	12	3235	**3363**
1999	13	2300	**2105**
2000	14	1443	**1318**



Raison	Premier terme
0,626	928561

Ainsi, à partir de nos documents, nous avons trouvé **un modèle mathématique (une suite géométrique décroissante de raison 0,626) qui schématise la réalité, en la simplifiant un peu, mais qui permet de comprendre son organisation.**

Nous allons désormais essayer d'exploiter le modèle. Nous remplacerons les données originales par les valeurs calculées afin de conjecturer l'évolution future de l'épidémie, puis de tenter de retrouver comment elle s'est développée par le passé.

B. Peut-on expliquer le modèle découvert ?

L'étude de la maladie nous permet d'avancer certaines hypothèses. Celles-ci sont indispensables pour faire parler les chiffres.

- **Première hypothèse : l'incubation dure au moins quatre ans.**

Cette hypothèse est suggérée par l'âge des animaux malades : au moins quatre ans.

- **Seconde hypothèse : les bovins contaminés vivent pendant quatre ans sans développer la maladie.**

Nous utiliserons les notations suivantes :

$C0_n$ le nombre de bovins contaminés durant l'année n, avec <u>$n = 1$ en 1987</u>.

$C1_n$ le nombre de bovins contaminés durant l'année $n - 1$ (contaminés depuis 1 an).

$C2_n$ le nombre de bovins contaminés durant l'année $n - 2$ (contaminés depuis 2 ans).

$C3_n$ le nombre de bovins contaminés durant l'année $n - 3$ (contaminés depuis 3 ans).

$C4_n$ le nombre de bovins contaminés durant l'année $n - 4$ <u>ou avant</u> (contaminés depuis 4 ans ou plus)

M_n le nombre de bovins qui déclarent la maladie au cours de l'année n (et sont alors abattus).

Pour connaître l'évolution de l'épidémie, année après année, on se demande de quelle manière se fait le passage de $(C0_n, C1_n, C2_n, C3_n, C4_n, M_n)$ à $(C0_{n+1}, C1_{n+1}, C2_{n+1}, C3_{n+1}, C4_{n+1}, M_{n+1})$.

D'après la seconde hypothèse : $C0_n = C1_{n+1} = C2_{n+2} = C3_{n+3}$.

Et d'après la première hypothèse, les M_{n+1} bovins qui développent la maladie au cours de l'année $n + 1$ faisaient partie des $C4_n$ bovins qui, l'année n, étaient contaminés depuis 4 ans ou plus.

- **Troisième hypothèse : les derniers bovins contaminés l'ont été en 1988, date d'interdiction des farines animales au Royaume Uni (l'année $n = 2$ pour nous).**

On suppose donc que l'interdiction a été respectée, ou que la désobéissance a été négligeable.

Ainsi : pour $n \geq 3$, $C0_n = 0$,
pour $n \geq 4$, $C1_n = 0$,
pour $n \geq 5$, $C2_n = 0$,
pour $n \geq 6$, $C3_n = 0$.

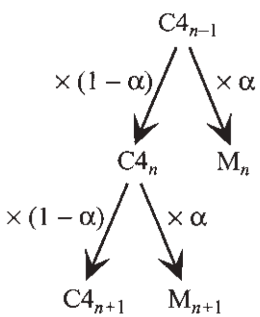
Donc, à partir de $n = 7$, $C4_n$ ne reçoit plus l'apport de $C3_{n-1}$.

Mais chaque année, les bovins contaminés depuis 4 ans au moins ne déclarent pas tous la maladie. Nous pouvons supposer qu'il y a une certaine probabilité pour qu'ils tombent malades. Plus simplement encore nous admettons qu'une certaine proportion d'entre eux développent la maladie.

- **Quatrième hypothèse : il existe un coefficient constant $\alpha \in\ ]0\ ;\ 1[$ tel que $\mathbf{M}_n = \alpha \times \mathbf{C4}_{n-1}$.**

Or, pour $n \geq 6$, $C3_n = 0$; donc, pour $n \geq 7$, $(C4_n)$ ne reçoit plus d'apport et diminue chaque année du nombre de nouveaux malades.
Pour $n \geq 7$: $C4_n = C4_{n-1} - M_n = C4_{n-1} - \alpha \times C4_{n-1} = (1 - \alpha) \times C4_{n-1}$.
Cette relation montre que **la suite $(C4_n)$ est géométrique de raison $1 - \alpha$ à partir de $n = 6$**.

On peut représenter la situation par un arbre.



On a alors $M_{n+1} = \alpha \times C4_n = \alpha \times (1 - \alpha) \times C4_{n-1}$
$= \alpha \times (1 - \alpha) \times (M_n/\alpha)$
$= (1 - \alpha) \times (M_n)$

Donc **la suite (M_n) est géométrique et de raison $1 - \alpha$ à partir de $n = 7$**.

Ainsi, les hypothèses énoncées expliquent la suite géométrique (M'_n) découverte précédemment.

Et comme la raison de cette suite est $q = 0{,}626$, alors : $\mathbf{1 - \alpha = 0{,}626}$ et donc $\alpha = \mathbf{0{,}374}$.

Chaque année, 37,4% des vaches contaminées depuis 4 ans au moins développent la maladie.

Les suites $(M_n)_{n\geq 7}$ et $(C4_n)_{n\geq 6}$ sont géométriques et de raison 0,626.

C. Quel avenir peut-on conjecturer ? (Prospective)

Commençons par un bilan de ce que nous savons en remplissant les cases du tableau suivant (nous verrons plus loin l'intérêt des premières lignes de ce tableau).
Tous les calculs sont arrondis à une unité près.

En maigre sur fond blanc, les valeurs de M_n pour $n < 7$ et les 0 résultant de l'arrêt des contaminations.

En gras sur fond blanc, les nombres de malades **calculés par le modèle**.

Par ailleurs on sait que $M_n = \alpha\, C4_{n-1}$ et que $\alpha = 0{,}374$.
On peut donc calculer la colonne des $C4_{n-1} = M_n/0{,}374$.
Exemples : $C4_0 = M_1/0{,}374 = 446/0{,}374 = 1193$. $C4_1 = 2514/0{,}374 = 6722$.
Etc. jusqu'à $C4_{12}$. Les résultats sont en maigre sur fond gris.

Pour calculer les valeurs à venir, il suffit de prolonger les suites $(C4_n)$ et (M_n) qui sont géométriques et de raison $1 - \alpha = 0{,}626$. Les résultats sont en gras sur fond gris.

alpha=	0,374						
Année	n	C0n	C1n	C2n	C3n	C4n	Mn
1982	-4					0	0
1983	-3					0	0
1984	-2					0	0
1985	-1					0	0
1986	0					1193	0
1987	1					6722	446
1988	2					19326	2514
1989	3	0				38521	7228
1990	4	0	0			67805	14407
1991	5	0	0	0		99679	25359
1992	6	0	0	0	0	93532	37280
1993	7	0	0	0	0	58551	**34981**
1994	8	0	0	0	0	36653	**21898**
1995	9	0	0	0	0	22945	**13708**
1996	10	0	0	0	0	14363	**8581**
1997	11	0	0	0	0	8991	**5372**
1998	12	0	0	0	0	5629	**3363**
1999	13	0	0	0	0	3524	**2105**
2000	14	0	0	0	0	**2206**	**1318**
2001	15	0	0	0	0	**1381**	**825**
2002	16	0	0	0	0	**864**	**516**
2003	17	0	0	0	0	**541**	**323**
2004	18	0	0	0	0	**339**	**202**
2005	19	0	0	0	0	**212**	**127**
2006	20	0	0	0	0	**133**	**79**
2007	21	0	0	0	0	**83**	**50**
2008	22	0	0	0	0	**52**	**31**

On observe que l'épidémie s'arrêtera d'elle-même s'il n'y a pas de nouvelles contaminations (l'E.S.B. est transmissible mais non contagieuse).
Enfin, il est inutile de poursuivre le calcul lorsque les vaches atteignent une vingtaine d'années.

Ces calculs ont été effectués sans connaître le nombre de cas en 2001. Aujourd'hui, c'est-à-dire en février 2002, le site web déjà cité mentionne 526 cas déclarés au 30/09/2001. On pourrait donc s'attendre à environ $526 \times 12/9 \approx 701$ cas pour l'année entière. Or notre modèle prévoit 825 cas. Mais ceci ne nous inquiète pas car nous avons pu constater que les nombres de cas déclarés sont souvent revus à la hausse.

D. Quel passé peut-on reconstituer ? (Rétrospective)

Reconstituer le passé, c'est retrouver les nombres de vaches initialement contaminées en remontant leur histoire. Il s'agit donc de compléter le haut du tableau précédent.

Calcul de $C3_5$.
On rappelle que $C4_6 = C3_5 + (1 - \alpha) \times C4_5$
où $C3_5$ est le nombre de vaches ayant juste 4 ans de contamination
et où $(1 - \alpha) \times C4_5$ est le nombre de vaches ayant 5 ans de contamination au moins et qui ne sont pas malades.
Donc $C3_5 = 93\,532 - 0{,}626 \times 99\,679 = 31\,133$.
Et par suite $C2_4 = C1_3 = C0_2 = 31\,133$. En gras sur fond blanc sur le tableau suivant.

Calcul de $C3_4$.
Comme ci-dessus, on a $C4_5 = C3_4 + 0{,}626 \times C4_4$.
D'où $C3_4 = 99\,679 - 0{,}626 \times 67\,805 = 57\,233$.

On en déduit la diagonale des résultats en maigre sur fond gris.

Derniers calculs.
De la même manière on trouve les valeurs affichées en gras sur fond gris.

On découvre que les contaminations ont commencé sournoisement dès 1982. Elles ont augmenté jusqu'en 1987, date des premiers cas déclarés. Elles ont diminué, puis disparu en 1988 car c'est l'année où les farines ont été interdites. Mais l'épidémie ne faisait alors que commencer. Elle allait s'amplifier pendant 4 ans, à mesure que les bêtes contaminées achevaient leur période d'incubation. Enfin le déclin s'est amorcé, en décroissance exponentielle. Au total, plus de 180 000 bêtes auront été contaminées et abattues, sans compter celles qu'on a abattues par précaution.
Le modèle mathématique nous apprend finalement beaucoup de choses. Mais celles-ci ne peuvent être vérifiées que si nos hypothèses sont bonnes.

Pour terminer, nous avons préparé une feuille de calcul EXCEL® qui montre le devenir de chaque contingent de vaches contaminées année après année. Par cumul, on aboutit au tableau suivant, aux erreurs d'arrondis près. C'est un peu comme la reconstitution d'un crime, pour vérifier que les résultats de l'enquête sont cohérents.

alpha=	0,374						
Année	**n**	**C0n**	**C1n**	**C2n**	**C3n**	**C4n**	**Mn**
1982	*-4*	**1193**	**0**	**0**	**0**	0	0
1983	*-3*	**5975**	**1193**	**0**	**0**	0	0
1984	*-2*	**15118**	**5975**	**1193**	**0**	0	0
1985	*-1*	**26423**	**15118**	**5975**	**1193**	0	0
1986	*0*	**43690**	**26423**	**15118**	**5975**	1193	0
1987	*1*	57233	**43690**	**26423**	**15118**	6722	446
1988	*2*	**31133**	57233	**43690**	**26423**	19326	2514
1989	*3*	0	**31133**	57233	**43690**	38521	7228
1990	*4*	0	0	**31133**	57233	67805	14407
1991	*5*	0	0	0	**31133**	99679	25359
1992	*6*	0	0	0	0	93532	37280
1993	*7*	0	0	0	0	58551	34981
1994	*8*	0	0	0	0	36653	21898
1995	*9*	0	0	0	0	22945	13708
1996	*10*	0	0	0	0	14363	8581
1997	*11*	0	0	0	0	8992	5372
1998	*12*	0	0	0	0	5629	3363
1999	*13*	0	0	0	0	3524	2105
2000	*14*	0	0	0	0	2206	1318
2001	*15*	0	0	0	0	1381	825
2002	*16*	0	0	0	0	864	516
2003	*17*	0	0	0	0	541	323
2004	*18*	0	0	0	0	339	202
2005	*19*	0	0	0	0	212	127
2006	*20*	0	0	0	0	133	79
2007	*21*	0	0	0	0	83	50
2008	*22*	0	0	0	0	52	31

Selon notre modèle, l'épidémie pourrait donc s'achever lorsque toutes les vaches contaminées auront déclaré la maladie ou seront mortes de vieillesse (en 2008, les vaches contaminées en 1988 auront 20 ans).