Le rôle créateur des mathématiques en sciences de la vie

Guy Rumelhard,
professeur, Lycée Condorcet, Paris ; INRP

Nécessité de l'interdisciplinarité et obstacles

Parmi les objectifs proposés aux travaux personnels encadrés (TPE), j'ai choisi de privilégier l'interdisciplinarité entre les mathématiques et les sciences de la vie pour les classes de première S, car cet objectif m'a semblé le plus difficile à réaliser, compte tenu des nombreux obstacles qui s'opposent à la rencontre de ces deux disciplines.

Du côté des sciences de la vie

Dans le travail de recherche les biologistes font, de manière très libre, des emprunts à d'autres disciplines, emprunts de techniques expérimentales (microscopie, électrophorèse, chromatographie…), de modèles analogiques (régulateur à boule ou couveuse thermostatée pour expliquer les régulations), de concepts (milieu, cycle, spécificité, affinité chimique), de méthodes pour organiser les expériences (carré latin, tirage au sort). Les découvertes ou plus simplement les avancées sont souvent faites par des savants provenant d'autres disciplines : Pasteur est chimiste comme J. Monod, A. Jacquard est polytechnicien comme D. Schwartz, Schrödinger est physicien, d'autres sont médecins, etc. Mais ce caractère interdisciplinaire peut disparaître dans la mesure où certains biologistes tendent à réduire la biologie à une biochimie et à considérer que le progrès des découvertes provient entièrement d'un progrès des techniques. Pour l'enseignant de biologie, le caractère interdisciplinaire peut également disparaître si l'on continue à croire que toute connaissance résulte directement d'une observation. Dans cette optique, connaître c'est voir. Pour de nombreux enseignants le caractère équiprobable de la séparation des chromosomes homologues au moment de la méiose ne relève pas d'une "conceptualisation" mais d'une "observation" au microscope sur des préparations fixées et colorées.

Par ailleurs, les problèmes concrets qui sont à l'origine de la recherche ne relèvent pas *a priori* d'une seule discipline. Les maladies sont bien souvent complexes, mêmes celles dont l'explication semble se réduire à la présence d'un microbe. Les problèmes écologiques, telle la pollution des sols par les engrais azotés, sont également complexes et concernent les chaînes alimentaires, la chimie des sols, mais aussi les questions économiques, le rendement, l'efficacité d'une production, donc des décisions politiques. Plusieurs maîtres mots caractérisent alors ces questions : prise de risque, prévention, dépistage précoce, prédiction, anticipation.

L'avancée des connaissances conduit à remanier ou à superposer les clivages entre les disciplines biologiques. Si l'on parle d'embryologie, de génétique, d'écologie, d'éthologie, d'immunologie, etc., on découpe également les questions en termes de structure, de fonction, de milieu, d'histoire au double sens de développement et d'évolution. On peut également découper les questions en fonction du concept d'organisation qui se décline en termes de système, de régulation, d'information. On peut aussi utiliser le concept de niveau d'organisation et distinguer les écosystèmes, les populations, les cellules, les molécules (génétique des populations, biologie cellulaire).

Mais il y a plus, car la biologie moléculaire se présente comme le terme ultime de l'analyse biologique et, chaque fois que l'on peut atteindre ce terme, l'approche mathématique semble obsolète. On peut soupçonner que le tabac est cause de certains cancers du poumon en calculant des corrélations, mais depuis que l'on connaît les molécules en cause et leur mécanisme d'action, le calcul de corrélations semble inutile. Il n'en est rien en réalité car une corrélation ne désigne pas nécessairement une cause unique, ni une cause directe, et il reste que le tabac n'est ni une cause nécessaire (on peut avoir un cancer sans jamais fumer), ni une cause suffisante (on peut fumer sans avoir le cancer).

Dernière difficulté pour rapprocher les mathématiques et les sciences de la vie, les physiciens peuvent parler avec pertinence de "physique théorique", mais il n'est pas encore certain que le concept de "biologie théorique" utilisé depuis le début des années 1970 ait un contenu équivalent.

Actuellement l'approche "expérimentale" domine, ce qui rend difficile d'admettre l'existence, dans les sciences de la vie et de la santé, de deux traditions qui paradoxalement s'opposent alors qu'elles dérivent toutes les deux des travaux de Pierre-Simon Laplace (1749-1827) : d'une part, la recherche d'un *déterminisme* strict qui se présente alors comme une science des certitudes et des lois, d'autre part un calcul des *probabilités* qui se présente comme une science de l'incertain et de la variabilité. Plusieurs titres de livres ou d'articles réunissent ces deux aspects sous une forme qui joue sur ce paradoxe pour créer un "oxymoron" : les certitudes du hasard, les lois du hasard, les structures du hasard [1]. Un recul historique illustre ces deux traditions à travers deux figures de savants exactement contemporains l'un de l'autre mais qui s'ignorent : Claude Bernard et G. Mendel. Claude Bernard (1865) [2], à la suite de Magendie, se préoccupe de sortir la médecine d'un certain empirisme en créant une physiologie expérimentale déterministe qui établit des invariants, des constantes et des lois. Mais ces recherches butent sur un obstacle. En étant très attentif à la variabilité des cas individuels, Claude Bernard se montre très réservé à l'égard de tout calcul de moyenne statistique, seul procédé permettant de réduire cette variabilité qui empêche de trouver des lois, des invariants ou des constantes. Selon lui, la moyenne annule les cas extrêmes, les cas rares et les oscillations. Et il explique qu'il aurait manqué la découverte fondamentale de la fonction glycogénique du foie en procédant ainsi et en faisant la moyenne des glycémies. G. Mendel (1865) [3] conçoit la transmission des "caractères" (nous disons actuellement des gènes) comme une combinatoire. Celle-ci peut faire l'objet de calculs en s'appuyant sur le développement du binôme. Les deux termes du binôme sont les deux allèles de chaque gène. Il établit des règles de combinaison que, paradoxalement, le XX^{ème} siècle tentera de transformer en "lois", selon l'injonction positiviste. En fonction de cela, il refuse de considérer les travaux de ses prédécesseurs sur les hybridations car ils n'ont pas été faits sur un nombre suffisant de cas. Selon lui, tous les cas possibles n'ont pas été réalisés. On ne peut donc les observer.

Du côté des mathématiques

Même si le point de départ qui suscite le travail mathématique est un problème pratique de production (augmenter le rendement d'une culture, rechercher l'action d'un facteur), ou la recherche d'une explication scientifique (mode de transmission des gènes), l'étude nécessite bien souvent un assez long *détour*. L'absence de retour à la question initiale, faute de temps ou de volonté, risque d'effacer l'aspect interdisciplinaire (analyse des développements du binôme pour la génétique, puis "retour" à l'étude de n couples d'allèles).

N'étant pas professeur de mathématiques, je n'avancerai pas d'explication justifiant la réticence vis-à-vis des statistiques. Or celles-ci constituent bien souvent l'essentiel du travail demandé en biologie et peuvent donner lieu à des calculs fastidieux si l'on privilégie l'étude de longues séries de données. La relation entre les deux disciplines se vit alors sur le mode de *l'outil*, de la discipline de service, du procédé de calcul (pour les mathématiques), ou bien sur le mode de *l'illustration* d'un concept ou d'une technique (pour les sciences de la vie).

Le titre de cet exposé soulignant le "rôle créateur des mathématiques" vise à sortir de cette alternative en minimisant par ailleurs la place des calculs, et en insistant sur les concepts. Il copie le titre d'un article concernant les relations entre mathématiques et physique [4]. Traditionnellement le terme d'opérateur était utilisé pour désigner le rôle des mathématiques. Selon Dominique Lecourt [5], qui analyse en 1969 de manière synthétique l'ensemble des travaux de Gaston Bachelard, la compréhension du fonctionnement des sciences trouve sa pierre de touche dans le rôle attribué aux mathématiques. La pensée commune parle des mathématiques comme d'un *langage*, un simple moyen d'expression qui "met en forme" qui accueille une réalité qui préexiste. Pour G. Bachelard, les mathématiques jouent un *rôle d'opérateur*. Elles constituent, elles construisent les observations et les faits expérimentaux. Elles opèrent comme le scalpel du chirurgien qui coupe et sépare, ou l'aiguille qui coud et relie, mais également comme le ressort tendu, elles relancent la recherche.

Du côté de la didactique

Chaque discipline définit un *ordre de progression* selon ce qui lui semble "simple" ou élémentaire et ce qui lui semble plus complexe. Mais ces ordres ne sont pas les mêmes selon les disciplines et ce qui est simple en mathématiques est bien souvent un cas rare ou singulier en biologie. Ce qui est familier ou d'observation immédiate en biologie est complexe en mathématiques. Bien souvent le fréquent ou le familier, consistant par exemple à comparer l'efficacité de divers engrais sur la croissance de plantes, n'est pas simple. Il y a donc un travail à faire pour choisir les exemples qui favorisent les rencontres. Il faut également savoir abandonner les ordres traditionnels de progression au profit *d'accès directs*, et ceci est souvent mal accepté en mathématiques comme en sciences de la vie. Par exemple, utiliser des tables de résultats, des abaques, des procédés graphiques pour résoudre des problèmes mathématiques sans établir ou démontrer ces résultats, ou bien modéliser un résultat biologique tout en méconnaissant les mécanismes biochimiques, génétiques ou chromosomiques qui l'expliquent et rendent parfois obsolète la modélisation ! (modéliser les changements de forme au cours du développement embryonnaire ou la transmission des gènes).

Éléments d'épistémologie historique

Un deuxième groupe de raisons tenant à l'épistémologie "spontanée" des enseignants des deux disciplines peut expliquer certaines difficultés de la rencontre. En sciences de la vie, la production de connaissances nouvelles résulte d'un *travail théorique* (concepts, modèles analogiques ou formalisés) largement sous estimé par les enseignants, et d'un *travail d'expérimentation*. Quand l'expérimentation n'est pas possible, on peut observer des variations à l'aide de comparaisons entre espèces (comparaisons géographiques, historiques à l'aide des fossiles et au cours du développement embryonnaire ; comparaison de l'anatomie, de l'éthologie, de l'écologie). Ce terme d'expérience est tellement invoqué de manière incantatoire ou comme slogan pour "défendre la discipline" - toutes les sciences étant supposées à tort expérimentales ! – on oublie sa signification. De manière très large, on pourrait dire que *expérimenter c'est faire varier*, et le mathématicien précisera immédiatement : variation absolue, variation relative, vitesse de variation, effet cumulé (relation de proportionnalité, de dérivation, d'intégration, PID en bref). On pourrait également retenir l'expression : il faut *transformer pour connaître*. Et le mathématicien proposera de nombreuses transformations (ou anamorphoses), en particulier pour comprendre les changements de formes qui ne peuvent donner lieu à un travail d'analyse au sens de décomposition. Une forme est un tout qui ne se décompose pas, mais qui se transforme. Par ailleurs, en physiologie, l'analyse ne suffit pas. Elle doit être suivie d'un travail de mise en relation des éléments séparés. Le terme *d'intégration* (emprunté aux mathématiques) prend alors une grande importance. On parle, à propos des systèmes de régulation, de *"biologie intégrative"*.
L'approche historique développée par cette université d'été ne se sépare pas d'une approche épistémologique. Réciproquement, toute épistémologie qui ne veut pas tomber dans un pur formalisme ou une pure logique se doit de mettre ses présupposés à l'épreuve de l'histoire des sciences. Sans développer longuement ce chapitre d'épistémologie historique, trois autres termes étroitement reliés entre eux méritent qu'on s'y arrête précisément pour les dissocier et recomposer leur ordre d'enchaînement : *décrire, expliquer, prévoir*.

Dans une optique naturaliste *décrire* semble être le premier terme incontournable de tout travail scientifique [6]. On peut être surpris que l'introduction d'un programme de mathématiques statistiques en première D et terminale D en 1966 ait reproduit le même clivage et le même ordre de succession :

mathématiques descriptives en première D (ce qui est particulièrement fastidieux), statistiques "explicatives" ou permettant les prises de décision en terminale. Or les descriptions empiriques sont bien souvent un obstacle à la connaissance. Cela a été pendant longtemps la plus fréquente erreur des médecins. Il ne suffit pas d'enregistrer le nombre de malades présents dans un hôpital et de le mettre en relation avec un traitement donné ou de rechercher des facteurs de risque. Il faut d'abord "construire" l'échantillon à observer en appliquant des règles ou une modélisation et donc surmonter *l'obstacle de l'empirisme* profondément ancré chez les naturalistes. Le savoir est conquis (contre les représentations qui font obstacle) et construit avant d'être simplement constaté. Le naturaliste admet volontiers que doivent se succéder dans le temps de la recherche une "embryologie descriptive", puis une "embryologie causale", et enfin, comme terme ultime supposé, une "embryologie moléculaire" qui oublie parfois le concept de milieu et redécouvre bien tardivement une "embryologie mathématique" présente dans les travaux singuliers (au double sens du mot) de D'Arcy Thompson dans la première moitié du XX$^{\text{è}}$ siècle [7].

Expliquer se décline de plusieurs façons : *conceptualiser, modéliser* (de manière analogique ou formalisée), rechercher des *causes*. Ce dernier terme est récusé par les physiciens positivistes qui ont longtemps parlé de lois (terme d'origine juridique ou religieuse) et utilisent actuellement le terme ambivalent de modèle qui a deux sens exactement inverses, ce qui permet l'intrusion de l'idéologie au cœur des explications scientifiques.

Prévoir est bien souvent couplé de manière indissociable à pouvoir. Les formules positivistes imprègnent implicitement l'enseignement : *connaître pour transformer, savoir pour prévoir, pour pouvoir*. Disons d'un mot que dans le domaine médical il existe de nombreuses situations dans lesquelles on peut prendre des décisions d'action sans connaître les mécanismes en jeu, et inversement, disposer d'explications parfaitement connues qui ne conduisent à aucun remède possible.

Analyse d'exemples du rôle créateur des mathématiques

Les limites imposées à cet exposé ne permettent pas de multiplier les exemples, ni même de développer longuement chacun de ceux que j'expose. Mon but est d'illustrer les difficultés repérées et de proposer des solutions directement dans ce texte ou par l'intermédiaire d'une bibliographie citée en référence.

Les grands domaines des sciences de la vie qui peuvent donner lieu à une rencontre avec les mathématiques peuvent se marquer par quelques concepts importants :

– la *variabilité*, à condition de bien marquer que plusieurs attitudes contradictoires, qui font donc obstacle les unes aux autres, sont ici possibles : réduire cette variabilité en minimisant son importance, en la considérant comme accidentelle, comme regrettable et en se contentant d'un calcul de moyennes (travaux de Quételet sur l'homme moyen à partir de l'observation de la taille des conscrits) dans le but de rechercher des lois ; prendre en compte cette variabilité en la décrivant (variance, écart) ; considérer, à la suite de Darwin, que cette variabilité contient l'explication de la transformation des espèces (nous évitons ici le mot évolution car Darwin parle de "descendance avec transformation") ; expliquer cette variabilité par une combinatoire génétique ;

– la *croissance*, propriété essentielle et spécifique du vivant (même si les minéraux croissent aussi mais de manière très différente), concept que l'on peut appliquer aux individus et à la dynamique des populations (populations animales, végétales, cellulaires, moléculaires) ;

– la *vie et la mort*, c'est à dire l'analyse des facteurs de risque et la modification de l'espérance de vie pour un individu donné ou pour une population en relation avec le concept darwinien de sélection ;

– le *fonctionnement* d'un organisme ou d'un ensemble d'organismes à travers les concepts de fonction, de compartimentation, de régulation, de cycle, de flux, de transfert d'information, etc. ;

– la *normativité biologique ou la créativité du vivant*, qui interdit précisément toute normalisation, toute formalisation *a priori* [8].

Chacun des concepts précédents peut s'analyser dans une optique *déterministe* (recherche de certitudes, de lois) ou *probabiliste* (prise en compte du hasard, d'une incertitude), le terme de modèle masquant cette distinction s'il n'est pas suivi d'un qualificatif (déterministe, stochastique).

Des situations binaires : garçons et filles, hommes et femmes

En statistiques, la demande du mathématicien portera sur un exemple "simple" c'est-à-dire binaire. Le biologiste peut proposer la répartition des sexes ou la répartition de telle maladie selon les sexes. Dans l'optique des TPE, il est aisé de proposer cette situation aux élèves car elle conduit à modéliser la répartition des chromosomes X et Y au moment de la méiose, à distinguer deux types de spermatozoïdes et donc à ouvrir la possibilité d'un choix des sexes, sinon d'un eugénisme (élimination des filles en Chine par exemple). Les anomalies chromosomiques interviennent également pour soutenir l'intérêt de cette question. Les types XXX, XXY, XYY, etc. sont encore dans l'actualité puisque le tueur Francis Heaulme a tenté de plaider la présence du syndrome de Klinfelter (XXY) pour expliquer son comportement. Par ailleurs, le sexe des sportifs s'analyse désormais au niveau chromosomique. Un sujet de baccalauréat avait repris cette idée d'anomalie intellectuelle liée à la trisomie des chromosomes sexuels ce qui est faux, et la notion de "chromosome du crime" n'a pas entièrement disparu. Il existe enfin des familles à filles et des familles à garçons qui contredisent l'équipartition et qu'il n'est pas aisé d'expliquer.

Sur cet exemple le mathématicien pourra introduire les *fluctuations d'échantillonnage* à partir d'une loi d'équipartition (nombre de garçons dans une famille de huit enfants, ce qui est devenu rare !), *intervalle de confiance* d'un pourcentage sur un petit effectif (huit à dix enfants) et sur une très grande population (totalité des naissances pendant une année dans un pays donné), *écart* entre le modèle et la réalité observée, recherche de la *loi de répartition réelle* impliquant la définition des *deux risques* de première espèce (risque α connu d'éliminer des cas vrais, il est choisi à 5%) et de deuxième espèce (risque β inconnu qu'il faut réduire au maximum pour espérer trouver la loi de répartition). On trouvera des données chiffrées dans le livre de Vessereau [9] et dans celui de Yves et Maurice Girault [10]. C'est ici également l'occasion d'utiliser, selon le moment de l'année et les connaissances des élèves, des tables ou des abaques donnant l'intervalle de confiance à 5% ou à 1% en fonction de diverses tailles d'échantillon, ce qui permet une première étude des résultats sans calcul précis [11]. Par la suite on peut donner la formule après l'avoir établie.

Nombre de garçons et de filles à la naissance (Vessereau. Que sais-je ?)

Garçons :	221 023	0,515
Filles :	208 417	0,485
Total :	429 440	1,0

Nombre de garçons dans des familles de huit enfants

0	215
1	1 485
2	5 331
3	10 649
4	14 959
5	11 929
6	6 678
7	2 092
8	342
total	53 680

Tableau 1

On développera de la même façon le cas de maladies liées au sexe (ou liées aux chromosomes sexuels, ce qui n'est pas la même chose) et dont la prévalence n'est pas la même dans les deux sexes. L'étude se fera alors en comparant la répartition observée à l'hypothèse (dite nulle) d'une égale répartition pour apprécier une éventuelle différence. Il faut prendre en compte la taille de l'échantillon et le cas des non malades, mais aussi d'autres facteurs tel l'âge par exemple.

Les cas rares dans lesquels le polymorphisme d'un caractère est réduit à deux phénotypes peuvent donner lieu à ce même genre d'étude. C'est le cas classique, développé dans tous les manuels de lycée et souvent posé au baccalauréat depuis 1973, de la Phalène du bouleau (Biston betularia) étudié par Ketelwell en Angleterre dans les années 1950 pour soutenir l'idée darwinienne de sélection naturelle et publié dans le livre de *Génétique écologique* de Ford [12].

Imputation causale entre tabac et cancer des poumons

Cette question qui donne lieu à de nombreuses campagnes de la part des comités d'éducation à la santé suscite facilement un vif intérêt, sinon même de violentes polémiques. Il ne s'agit pas ici de prendre parti dans un sens ou l'autre mais de chercher comment une imputation causale a pu être mise en évidence initialement à l'aide d'un procédé mathématique : l'établissement d'une *corrélation*, puis la recherche de *facteurs de risque* dont l'importance peut être appréciée à partir des modifications de *l'espérance de vie*. Dans la mesure où chacun sait que l'on peut être malade avec ou sans tabagisme, et non malade que l'on fume ou pas, un tableau à double entrée croisera les deux situations : malade / non malade et fumeur / non fumeur. Le tabac est une *condition qui n'est ni nécessaire et ni suffisante*. C'est d'ailleurs le cas de beaucoup de facteurs impliqués dans beaucoup de maladies. Pour les maladies infectieuses, la condition nécessaire "présence du microbe" n'est pas suffisante. Nous ne sommes pas égaux devant la tuberculose ou le Sida, même si la présence du bacille de Koch ou du virus HIV est indispensable. D'autres co-facteurs génétiques ou du milieu interviennent. La méthode des corrélations est une recherche de causes éventuelles, mais n'en constitue pas *une preuve*. La corrélation établie peut n'avoir aucun sens, et en tout cas, une corrélation ne signe pas nécessairement une cause unique, ni une cause directe. Il peut toujours exister un facteur qui détermine les deux éléments mis en relation. Pédagogiquement, il peut être intéressant de confronter cette situation aux fausses corrélations qui conduisent à des conclusions visiblement absurdes. Les exemples sont aisés à trouver [13].

Organiser *a priori* une observation ou une expérience

Le *comptage* des globules rouges dans une goutte de sang, de cellules de levures dans une culture en évolution, de cellules d'algues microscopiques croissant en fonction de divers facteurs implique une procédure précise, faute de quoi les valeurs trouvées n'auront aucune fiabilité. Il faut diluer la solution qui contient en général beaucoup trop de cellules par mm^3 et s'assurer que cette dilution est bien homogène. Pour compter sous le microscope, on utilise une lame dont le volume sous la lamelle est parfaitement connu, cette lamelle étant divisée en petits carrés qui servent de repère pour compter. La dilution doit être telle que le nombre de cellules vivantes par carré soit aisément comptable (0 à 10 environ). Mais il faudrait surtout vérifier que la répartition des cellules dans les divers carrés soit aléatoire, c'est-à-dire obéit à une *loi de Poisson*, de façon à prendre en compte le résultat trouvé que l'on multiplie ensuite par le coefficient de dilution [14].

L'organisation d'un *plan d'expérience* peut dépendre également d'une procédure aléatoire, faute de quoi des biais peuvent s'introduire conduisant à des résultats erronés. L'utilisation du *carré latin* en est un bon exemple. Si les individus sur lesquels on expérimente un traitement sont pris eux-mêmes comme témoin et ne sont pas confrontés à un groupe témoin extérieur, il faut combiner les éléments de l'étude, c'est-à-dire un traitement donné (A), sur un individu donné (B), à un moment donné de l'évolution de la maladie (C), dans une quantité donnée (D), à un endroit donné de l'organisme (E), (voie digestive, intramusculaire, intraveineuse) de manière aléatoire. On peut nommer A,B,C,D,E, ces divers éléments. Une combinaison en carré permet d'associer ces éléments :

A B C D E	D C A B E
B C D E A	E D B C A
C D E A B	B E C A D
D E A B C	C A D E B
E A B C D	A B E D C

Tableau 2

Le premier carré obtenu par permutation systématique introduit un biais car les éléments A et B par exemple, se suivent quatre fois. Si on le construit en tirant au sort chaque case, on n'est pas assuré d'avoir les cinq éléments dans chaque ligne ou chaque colonne. On peut donc permuter au hasard ces lignes et ces colonnes. On obtient alors le second carré qui permet de guider l'organisation des expériences [15].

D'autres procédures de "randomisation" peuvent s'appuyer sur des chiffres engendrés au hasard et que l'on peut lire dans des tables (tables de 0 et 1 ; tables de séries de dix chiffres : 1 2 3 4 5 6 7 8 9 0), ou produire à l'aide des calculatrices dont les élèves disposent.

Modéliser la dynamique de populations animales

Plusieurs situations sont à considérer [16]. Si on observe la totalité de la population à l'œil nu, le dénombrement est plus ou moins aisé (nombre d'enfants trisomiques observés à la naissance ; nombre de cerfs vivants dans une forêt enclose ; nombre de vaches folles atteinte d'ESB en supposant qu'elles n'ont pas été abattues avant l'apparition de la maladie). Bien souvent, la population totale n'est pas visible (animaux souterrains, nocturnes, poissons…) et il faut alors des procédures indirectes : captures simples par analyse des pêches en fonction d'une maille de filet pour déterminer les quota acceptables. Marquage après capture, puis lâcher et nouvelle capture en admettant que les animaux marqués se "diluent" de manière aléatoire dans la population. Les mathématiques fournissent alors des fonctions décrivant l'évolution et permettant d'extrapoler (ou d'interpoler). Le concept important est alors celui de *modèle*. Idéalement, telle population devrait suivre une *croissance géométrique* : 1, 2, 4, 8, 16, 32, 64, etc. On peut alors évaluer l'influence de tel facteur par rapport à ce modèle d'évolution. Présence d'un facteur limitant, tel que la quantité de nourriture.

Le cas des vaches atteintes d'ESB est d'actualité car une maladie humaine nouvelle lui semble liée : la variante de la maladie de Creutzfeldt Jacob (vMCJ). L'évolution de l'ESB et de beaucoup d'autres infections et intoxications serait modélisée par une courbe de Laplace-Gauss (première ligne du tableau) [17]. Connaissant le début décalé de plus de huit ans de la vMCJ (avant 1995, on ne sait pas distinguer les deux formes classique et nouvelle), on peut donner diverses valeurs à la moyenne et à l'écart-type et extrapoler l'évolution à venir de la maladie. Les nombres actuels croissent de 3 à 25 cas par an en cinq ans. On serait ainsi au début de la courbe de Gauss, mais on ne connaît pas le délai réel d'incubation (8, 10 ou 15 ans ou plus ?). Certains prévoient plus de 100 000 malades au total d'ici 10 ans, au plus fort de la courbe de Gauss. Le nombre de MCJ "classique" qui atteint les adultes âgés semble, quant à lui, ne pas varier (troisième ligne du tableau).

1987	1988	1989	1990	1991	1992	1993	1994
132	1910	6863	12289	22613	34712	36271	25579

1995	1996	1997	1998	1999	2000	
15453	3672					ESB
3	10	10	18	15	25	vMCJ
35	40	59	63	61	38	MCJ

Tableau 3

Prédire l'espérance de vie et ses facteurs de variation

Notion très largement vulgarisée, l'espérance de vie n'en est pas moins très délicate à définir, plus encore à calculer, soit pour un individu donné à un moment donné de sa vie, soit pour une cohorte d'individus nés la même année, soit pour une population vivant dans un pays donné. Si les tables de mortalité et les *causes de la mort* ont été correctement enregistrées, cette notion trouve une définition idéale une fois que tous les individus nés une année donnée sont morts (soit au maximum 125 ans après car la durée maximale de vie ne varie pas). Mais cette notion n'intéresse que l'historien ; tant que l'on est vivant, on s'intéresse à la *valeur prédictive* du concept. De même pour les responsables d'un pays. De même pour une population animale dans une optique darwinienne de sélection des espèces en reliant cette notion à l'efficacité des diverses formes de reproduction et aux divers facteurs de risque. Pour avancer davantage, il faudrait reproduire ici de fastidieuses tables de mortalité. Nous renvoyons donc à la bibliographie [18].

Décrire et normaliser la taille et le poids chez l'homme et prédire l'évolution

Cette question largement développée dans l'atelier *: L'homme moyen et la normalisation de la taille et du poids* [dans *Statistiques et probabilités*] [19] est l'occasion de rappeler que la courbe de Gauss n'est que l'approximation de la loi binomiale quand "n" est suffisamment grand. De plus, le facteur intéressant est "sous la courbe". Autrement dit l'explication génétique de la taille fait appel à des éléments discontinus et nombreux, sans que le nombre puisse actuellement être précisé et qui concernent une population d'individus. La loi binomiale suffit donc à fournir une explication théorique. Cette explication se présente comme le modèle de la variabilité de la taille si elle n'était déterminée que génétiquement. De nombreux facteurs du milieu vont modifier cette taille dans un sens ou l'autre et éloigner les observations réelles du modèle. On pourra grâce à cela apprécier éventuellement l'effet de chaque facteur par rapport à la répartition théorique.

Mais une deuxième dimension apparaît ici dans la mesure où il s'agit de l'espèce humaine : c'est l'effet *normalisateur* de ces répartitions, c'est-à-dire la nécessité de se conformer à des valeurs de référence. Cet effet normalisateur prenant de plus en plus une valeur *prédictive*.

Décrire, modéliser, expliquer les formes vivantes

On peut *décrire* certains aspects des animaux et végétaux à l'aide des mathématiques, puis se limiter à contempler**,** ou admirer les aspects "géométriques" de la nature, par exemple les symétries, les spirales, les hélices, les suites de Fibonacci. En apparence l'art imite parfois la nature selon la thèse contestable développée autrefois par Aristote. Mais prendre la nature pour modèle c'est le *sens inverse* de la modélisation scientifique (et aussi de la véritable création artistique), mais c'est hélas le *sens commun* du terme. La relation de modèle est effectivement *réversible*, au prix d'une ambiguïté, ou d'une utilisation idéologique, qui s'appuie sur ce double sens pour créer un "contresens" sans le dire. Par exemple, le sens du mot *régulation* dans la bouche du physiologiste ou du technologue est exactement l'inverse du sens que lui donne l'économiste.

Le travail du scientifique consiste à *"transformer pour connaître"*. Toutes les expériences consistent à créer des variations artificielles, à l'aide de techniques variées, et en ce sens tous les verbes d'action formés avec le préfixe "trans" sont parents (sans mauvais jeu de mots), ainsi que ceux créés avec le préfixe "méta" (métamorphose). La forme des êtres vivants se développe dans l'espace mais aussi toujours dans le temps au double sens de développement et d'évolution, et présente donc des transformations à observer.

Or d'une certaine façon une forme ne se décompose pas. C'est un tout, contrairement aux mécanismes biochimiques. En revanche, on peut la transformer, ou observer sa construction progressive pour comprendre sa formation.

Modéliser, c'est rechercher des lois physiques, des réactions chimiques, des équations mathématiques qui permettent d'expliquer les phénomènes biologiques, et non pas seulement de les décrire**.** Regarder le vivant *comme un objet chimique* est assez facilement accepté car il y a des réactions de transformation, une dynamique, une construction (synthèse), une destruction (hydrolyse), donc du temps et des forces à l'œuvre, forces de création et de mort. Il en est de même, mais dans une moindre mesure pour la physique.

Regarder le vivant *comme un objet mathématique* qui explique son fonctionnement laisse plus dubitatif car les mathématisations semblent statiques et seulement descriptives**.** En fait, il faut comparer des états successifs ou des états dans des espèces différentes pour espérer observer des transformations (puisqu'il faut transformer pour connaître). D'où l'importance par exemple du dernier chapitre du livre de D'Arcy Thompson [20] qui propose des grilles de transformation des formes animales. L'opposition trop marquée des attitudes analytiques et des approches globales constitue cependant un obstacle possible à l'acceptation de ce type de travail.

On compare donc la taille, la surface ou le poids d'organes différents à l'intérieur d'un individu au cours de la croissance et/ou chez des espèces différentes actuelles ou fossiles. On compare ces paramètres au cours du temps, mais non pas par rapport au temps. On compare les variations relatives les unes par rapport aux autres (vitesse de croissance relative, croissance dysharmonique, allométrie) en utilisant des échelles logarithmiques si besoin est.

Dans le cadre de l'évolution, deux conceptions de l'évolution des formes qui s'appuient sur les fossiles et la paléontologie et impliquent des mutations et des effets de sélection sont en concurrence sans s'exclure :

– le *gradualisme* : évolution progressive continue, sans rupture, mais éventuellement à des vitesses différentes ;

– les *équilibres*, stables pendant une période**,** ponctués par des ruptures et des évolutions rapides pendant un temps bref (ce qui à l'échelle des temps géologiques peut donner l'impression d'un "saut qualitatif").

Bien évidemment, il faut également faire intervenir le milieu, car tout n'est pas déterminé génétiquement et par ailleurs les formes créées sont, ou non, retenues par la sélection.

Les mathématiques du continu sont les plus habituelles (fut-ce en échelle log, ou log/log) et semblent appuyer l'attitude gradualiste sans le dire, mais il faut également faire usage de mathématiques du discontinu permettant de décrire des seuils, des ruptures, des "sauts".

On retrouve ainsi la génétique du développement car ce petit nombre de transformations mathématiques désigne un petit nombre de gènes en action, à des moments successifs et à des vitesses différentes. Il n'est pas besoin de faire appel à une révolution sous forme d'un très grand nombre de mutations et d'un "saut" difficile à comprendre. Cependant, la difficulté est ici la même que pour la génétique dite formelle ou mendélienne. L'approche mathématique directe des lois de Mendel est possible (le développement du binôme modélise bien la transmission des allèles des gènes) mais ceci n'entraîne aucun développement**,** aucune relance de la recherche tant que les biologistes n'ont pas analysé la méiose et le ballet des chromosomes.

Conclusion

Pour conclure reprenons tous les cas analysés ou simplement évoqués précédemment :

1. Recherche de *facteurs de variation* (et d'imputation causale éventuelle) par la méthode des *corrélations* en épidémiologie, écologie, agronomie (en gardant à l'esprit que corrélation ne signifie pas nécessairement cause, en tous les cas pas cause unique, ni cause directe).

2. Définition *a priori d'un plan* pour réaliser *des expériences ou des observations* (étude statistique sur une population et non pas sur un seul cas, carré latin, tirage au sort…).

3. Proposition d'un *modèle* d'évolution ou de fonctionnement (modéliser la dynamique d'une population, d'un système de régulation, d'un organisme dans lequel on définit des compartiments, etc.), permettant de réaliser *des simulations* (mais il s'agit souvent de fonctions mathématiques inaccessibles aux élèves de première S : équations différentielles du second ordre, etc.).

4. *Comparaison* d'un résultat expérimental à une loi théorique (une probabilité, un pourcentage, une fonction de répartition, la loi binomiale, la loi de Laplace-Gauss, une loi de Poisson) ou à un modèle. *Réfuter ou appuyer* la conformité d'un résultat avec une loi et jouer le rôle de *"preuve"* (risque α et β, tests de signification, intervalle de confiance).

5. *Comparaison* de deux résultats empiriques entre eux (comparaison de deux pourcentages, évolution de la croissance de deux lots de plantes ou d'animaux, comparaison de moyennes et de variances), à l'aide de tests de signification.

6. *Expliquer* en proposant des concepts (la combinatoire permet de prévoir des cas possibles ; le concept d'équiprobabilité permet de prévoir les proportions de garçons et de filles dans la descendance).

7. La statistique permet de *prendre des décisions d'action*, même si l'on n'explique pas le phénomène (cas de l'efficacité des médicaments dont le mécanisme d'action n'est pas connu).

8. Les mathématiques permettent de *réduire la variabilité du vivant* (et donc de dépasser cet obstacle à l'expérimentation constitué par la variabilité) en autorisant l'étude sur une population et non pas sur un seul individu. Mais inversement cette réduction peut être un obstacle dans la mesure ou elle élimine les cas très rares et les oscillations qui ont bien souvent, en biologie, une grande importance.

9. Les mathématiques permettent de réduire la *variabilité due à l'instrument de mesure*.

10. Les mathématiques permettent de critiquer les *fausses représentations* véhiculées par le langage courant sur le concept de hasard par exemple (ou d'autres concepts), ainsi que les fausses mathématisations de doctrines idéologiques [21].

Références bibiographiques

[1] BOURSIN J.-L., *Les structures du hasard*. Paris : Seuil Le rayon de la science. 1966 ; BOREL É. *Probabilités et certitudes* Paris : PUF, Coll. Que sais-je ? n° 445 ; BOLL M. *Les certitudes du hasard*. Paris : PUF, Coll. Que sais-je ? n° 3. 1941

[2] BERNARD C., *Introduction à l'étude de la médecine expérimentale*. Paris : 1865. Édition avec préface de F. Dagognet chez Garnier-Flammarion. 1966

[3] RUMELHARD G., *La génétique et ses représentations* Bern. Peter Lang 1986

[4] BOUTOT A., *Le pouvoir créateur des mathématiques*. La Recherche n°215 Novembre 1989 p.1340-1348.

[5] LECOURT D., *L'épistémologie historique de Gaston Bachelard*. Paris : Vrin 1969

[6] *Décrire dans toutes les disciplines*. Revue *Cahiers pédagogiques* coordonnée par Yves Reuter. n°373 avril 1999

[7] D'ARCY Thompson, *Formes et croissance*. Paris : Seuil coll. Sources. 1994. Préface de Alain Prochiantz ; Prochiantz A., *Machine-esprit*. Paris : Odile Jacob. 2001 p.103-122 sur D'Arcy Thompson.

[8] CANGUILHEM G., *Le normal et le pathologique*. Paris PUF 1966 ; Canguilhem Georges (1968). *Études d'histoire et de philosophie des sciences concernant les vivants et la vie.* Paris : Vrin. (La 7ème édition de 1994 est augmentée de plusieurs articles)

[9] VESSEREAU A., *La statistique* Paris : PUF, Coll Que sais-je ? n°281 1947

[10] GIRAULT Y. et M., *L'aléatoire et le vivant* Paris : Diderot. 1999 (accessible aux élèves)

[11] FORD E., *Génétique écologique*. Paris : Masson. 1973 ; Jacquard Albert (1978) *Éloge de la différence*. *La génétique et les hommes*. Paris : Seuil (accessible aux élèves)

[12] SCHWARTZ D., (1991) *La statistique dans les sciences du vivant.* Dossier documentaire INSERM (épuisé), et une vidéo (accessible aux élèves) ; Schwartz Daniel (1994) *Le jeu de la science et du ha*sard. Paris Flammarion (reprend en partie le dossier INSERM) ; Schwartz Daniel (1963) *Méthodes statistiques à l'usage des médecins et des biologistes*. Paris, Flammarion. 318 pages

[13] RUMELHARD G., (éd) Les formes de la causalité dans les sciences de la vie et de la Terre. Paris INRP. 2000.

[14] LAMOTTE M., (1948) Introduction à la biologie quantitative. Paris : Masson.
[15] BERTRANDIAS J.-P. et F., (1997*) Mathématiques pour les sciences de la vie, de la nature et de la santé*. PUG ; Piquemal Jacques (1993*) Essais et leçons d'histoire de la médecine et de la biologie*. Paris : PUF (articles sur P. Louis et les débuts de l'épidémiologie)
[16] SAFINA C., (1996) *Les excès de la pêche en mer*. Pour la Science n°219 Janvier 1996 pages 28-36. Mot clé : surpêche ; Dajoz Roger (1973). La dynamique des populations exploitées : le cas des poissons marins. *Bull. Écologie* t. IV, 2, pp 121-131 ; Dajoz Roger (1970). *Précis d'écologie*. Dunod ; Lamotte M., Bourlière F. *Problèmes d'écologie : échantillonnage des peuplements animaux des milieux terrestres.* Paris : Masson 1969.
[17] Cité des sciences et de l'industrie : *En période de vache folle* Juin 2001 ; Tangente N° 80 Avril-Mai 2001
[18] TEMAN D., *La croissance de l'espérance de vie*. Tangente n°80 Avril-mai 2001.Consulter les sites Internet de l'INED http://www-census.ined.fr/demogrus/Demographie/Mortalité/ ; Fagot - Largeault Anne (1989) *Les causes de la mort. Histoire naturelle et facteurs de risque*. Paris : Vrin
[19] ARRIGHI M., *La variabilité de la taille et du poids chez l'homme : moyenne statistique, normalité, normativité. Une approche interdisciplinaire au collège*. Aster 30.p. 143-168. 2000
[20] Sur la modélisation de la forme on peut consulter : Collectif (1998) *Les symétries de la nature.* Pour La Science. Juillet. Dossier hors série ; Gayon J. (2000) *De la croissance relative à l'allométrie* (1918-1936). Rev. hist. Sc. 53/3-4, 475-500 ; Douady S., Couder Y., (1993) Les spirales végétales*. La Recherche* Janvier 1993 ; Murray J., (1988) *Les taches du léopard*. Pour La Science n°127 mai ; Reffye Ph de. ; Edelin C. ; Jaeger M. (1989) *La modélisation de la croissance des plantes. La Recherche* n°207 Février 1989 p. 158-168 ; Niklas K., (1986) *L'évolution des plantes simulée par ordinateur* Pour La Science n°103 mai ; Gayon J., Wunenburger J-J. (éd) (1992) *Les figures de la forme.*
[21] LAHANIER-REUTER D., (1999) *Conceptions du hasard et enseignement des probabilités et statistiques*. PUF, Paris.