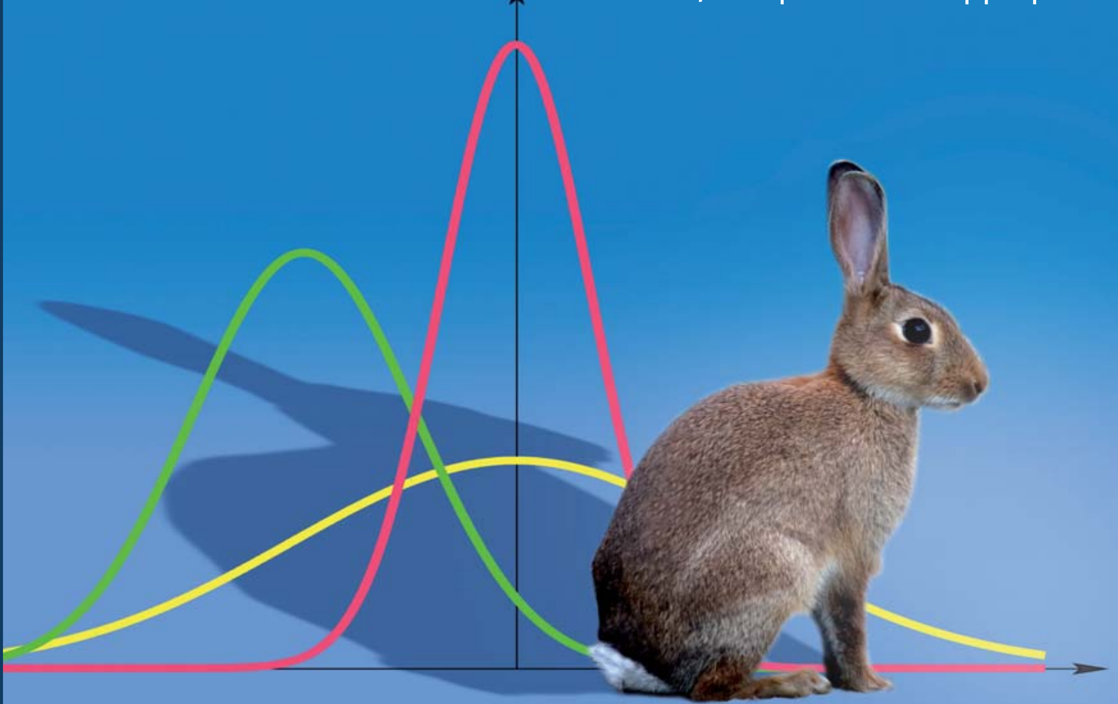
COLLECTION ENSEIGNEMENT SUP //// Mathématiques

Gérard Biau, Jérôme Droniou et Marc Herzlich

L1-L3

Mathématiques et statistique pour les sciences de la nature

Modéliser, comprendre et appliquer



EDP SCIENCES

MATHÉMATIQUES ET STATISTIQUE POUR LES SCIENCES DE LA NATURE

Modéliser, comprendre et appliquer

Gérard Biau, Jérôme Droniou et Marc Herzlich

Collection dirigée par Daniel Guin

EDP SCIENCES

17, avenue du Hoggar
Parc d'activités de Courtabœuf, BP 112
91944 Les Ulis Cedex A, France

Illustration de couverture : Antoine Fournier (antoine@chimachima.net)

Imprimé en France

ISBN : 978-2-7598-0481-8

Tous droits de traduction, d'adaptation et de reproduction par tous procédés réservés pour tous pays. Toute reproduction ou représentation intégrale ou partielle, par quelque procédé que ce soit, des pages publiées dans le présent ouvrage, faite sans l'autorisation de l'éditeur est illicite et constitue une contrefaçon. Seules sont autorisées, d'une part, les reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective, et d'autre part, les courtes citations justifiées par le caractère scientifique ou d'information de l'œuvre dans laquelle elles sont incorporées (art. L. 122-4, L. 122-5 et L. 335-2 du Code de la propriété intellectuelle). Des photocopies payantes peuvent être réalisées avec l'accord de l'éditeur. S'adresser au : Centre français d'exploitation du droit de copie, 3, rue Hautefeuille, 75006 Paris. Tél. : 01 43 26 95 35.

© **2010, EDP Sciences**, 17, avenue du Hoggar, BP 112, Parc d'activités de Courtabœuf,
91944 Les Ulis Cedex A

TABLE DES MATIÈRES

This page intentionally left blank

AVANT-PROPOS

Pourquoi ce livre ?

Ce livre présente, en un seul volume, un choix de concepts et d'outils pouvant constituer le programme de mathématiques des trois premières années d'études universitaires en sciences de la nature ou de la vie.

Il est né de l'expérience que nous avons acquise dans l'enseignement des mathématiques à l'Université Montpellier 2 devant des étudiants de licences de biologie, chimie et sciences de la Terre. Dans le droit fil de cette expérience, nous avons souhaité écrire un ouvrage *de mathématiques* destiné en priorité à *des étudiants en sciences de la nature et de la vie* et, plus généralement, à tout lecteur curieux de découvrir une présentation moins abstraite, mais pas pour autant imprécise, des concepts mathématiques indispensables à la modélisation des phénomènes naturels.

L'ambition que nous nous sommes fixée est donc double :

– Que cet ouvrage ne soit pas un traité abstrait de mathématiques.

– Qu'il ne se résume pas à un recueil de techniques, où tout souci de compréhension profonde serait évacué au profit de la seule pratique.

Nous ne voulions pas non plus d'un intermédiaire maladroit qui sacrifierait tour à tour, selon les chapitres et les notions abordés, l'exigence de rigueur ou les objectifs pédagogiques. Nous avons donc fait le pari qu'il était possible d'écrire un livre d'un niveau mathématique homogène, nourri des applications et destiné à l'utilisateur (plutôt qu'au concepteur) des mathématiques.

Les trois maîtres mots de cet ouvrage sont : *modéliser*, *comprendre* et *appliquer*.

La *modélisation* est un élément essentiel de la démarche scientifique, qui permet de passer des résultats d'expériences ou d'un recueil d'observations à une description organisée du monde. La nature est complexe et parvenir à un bon modèle est un travail délicat. Ce travail impose une réflexion profonde sur les phénomènes étudiés, sur les résultats souhaités et sur les hypothèses simplificatrices que le scientifique sera prêt à accepter. Même dans un ouvrage comme celui-ci, que nous avons voulu accessible aux non spécialistes, il nous a semblé essentiel de ne pas masquer ces difficultés. En tant que mathématiciens, nous considérons en effet que la démarche de modélisation fait partie intégrante de l'enseignement scientifique universitaire, y compris et surtout dans un cours de mathématiques ayant les ambitions décrites plus haut.

Le deuxième objectif de ce livre consiste à *faire comprendre* les concepts et les outils mathématiques introduits dans la première étape de modélisation. Il n'était pas question de produire un traité de mathématiques abstraites, mais nous pensons néanmoins qu'un usage efficace des mathématiques n'est possible que s'il s'accompagne d'une compréhension suffisamment profonde des concepts qui les sous-tendent. Sans cette dernière, l'esprit reste prisonnier de la technique et est incapable de s'en libérer lorsque le domaine d'application ou les circonstances l'exigent.

L'objectif final de l'ouvrage, enfin, réside dans l'*application* des concepts et des outils ou, en d'autres termes, dans la mise en action du modèle mathématique. Ce retour à l'origine des problèmes est évidemment indispensable. Il permet d'abord de tester la précision du modèle, d'en déterminer le domaine de validité et, une fois ces points établis, d'agir, de prévoir ou de prendre des décisions, justifiant ainsi tout le travail accompli.

Quel est son contenu ?

Les thèmes abordés dans le livre recouvrent l'essentiel des mathématiques enseignées aux étudiants de sciences de la nature et de la vie lors des trois premières années des études universitaires : bases de l'analyse des fonctions de une et plusieurs variables, probabilités élémentaires, concepts et outils statistiques, et introduction aux systèmes dynamiques. Ce contenu n'a évidemment rien de nouveau, l'originalité que nous réclamons résidant dans la volonté de rassembler *en un seul ouvrage* l'ensemble de ces notions mathématiques. Lorsqu'il était dispensé à Montpellier, cet enseignement était divisé en quatre cours répartis sur les cinq premiers semestres de la licence, chaque cours comptant une cinquantaine d'heures composées d'enseignement magistraux et de travaux dirigés.

La première partie du livre[1], intitulée « Bases », peut être pensée comme le bagage mathématique minimal que devraient posséder les étudiants à l'issue d'une première année d'études universitaires. Elle s'ouvre par un chapitre consacré à l'étude des fonctions d'une variable réelle, où l'on revient sur des connaissances déjà bien balisées au lycée. Lui succède un chapitre plus court dévolu aux fonctions de plusieurs variables, qui débute sur des considérations élémentaires (surfaces-graphe, lignes de niveau, dérivées partielles...) et se conclut par une introduction à des notions mathématiques plus élaborées utilisées en modélisation thermodynamique. Le calcul des probabilités et ses applications constituent le corps du troisième chapitre. Cette première partie se referme par un chapitre d'une nature plus descriptive, dédié aux relations délicates – et malheureusement pas toujours bien comprises – entre probabilités et statistique.

La « Statistique » est au cœur de la deuxième partie de l'ouvrage, qui porte d'ailleurs ce titre. Cette partie expose la démarche, les principaux concepts et les outils essentiels de l'approche inférentielle : estimation ponctuelle et par intervalle, tests paramétriques et non paramétriques, et enfin corrélation, régression et introduction à l'analyse de variance. La statistique inférentielle n'a d'autre objectif que de transporter des résultats numériques obtenus sur un échantillon à la population entière dont ce dernier est issu. Il s'agit d'un domaine essentiel des mathématiques appliquées, et nous avons souhaité en faire une présentation rigoureuse mais sans formalisme excessif, s'appuyant sur de nombreuses applications.

La troisième partie est consacrée aux « Systèmes Dynamiques », appellation que nous avons préférée à celle, plus réductrice, d'équations différentielles. Le fil directeur consiste ici à modéliser des phénomènes dépendant du temps de manière déterministe, c'est-à-dire ne laissant pas de place à l'aléatoire. Un premier chapitre présente quelques éléments d'étude des équations différentielles, pour la plupart déjà abordés dans le secondaire mais qu'il est bon de se remémorer (avec peut-être un éclairage nouveau), tant ces équations sont indispensables pour modéliser de nombreux phénomènes naturels. Le chapitre suivant introduit les bases du calcul matriciel, partant de considérations assez simples sur les systèmes linéaires pour aboutir au concept de diagonalisation, indispensable à l'analyse de systèmes dynamiques en temps discret (décrivant par exemple des phénomènes dont l'évolution est annuelle ou saisonnière). Le dernier chapitre de cette partie est un mariage entre équations différentielles et calcul matriciel, thèmes des deux développements précédents ; il a pour objet les systèmes d'équations différentielles, qui jouent eux aussi un rôle important dans la modélisation de nombreux phénomènes dynamiques complexes.

[1] À l'exception des paragraphes 2.4 à 2.6 du chapitre consacré aux fonctions de plusieurs variables.

Enfin, la quatrième et dernière partie de l'ouvrage regroupe les solutions aux exercices proposés à la fin de chaque chapitre.

Comment le lire ?

La démarche pédagogique que nous avons choisie consiste à ouvrir chaque chapitre par la présentation d'une ou plusieurs situations concrètes se prêtant à une modélisation. Les concepts mathématiques pertinents se trouvant ainsi naturellement motivés, le reste du texte est alors consacré à leur étude. Le chapitre se referme enfin par un retour à la problématique ayant servi de motivation initiale.

De nombreux exercices permettent de compléter l'exposé et d'ouvrir vers davantage d'applications[(2)].

Nous nous sommes efforcés d'adopter un mode de présentation adapté à notre public, remplaçant autant que possible la litanie classique « Définition – Proposition – Théorème » bien connue des mathématiciens par un style plus discursif. En particulier,

Lorsqu'un concept ou outil nouveau est introduit pour la première fois, sa définition apparaît dans un cadre bleu, le ***mot nouveau*** *étant mis en gras et bleu.*

Lorsque son importance ne justifie pas une telle mise en exergue, le concept est placé dans une phrase imprimée en bleu, mais il est toujours mis en gras lors de sa première apparition.

Les propriétés les plus importantes, énoncés de théorèmes ou de techniques à retenir, sont mises en valeur par un cadre gris.

Enfin, un index terminal reprend les termes les plus significatifs. Profitons de l'occasion pour rappeler qu'un texte mathématique doit toujours être lu plusieurs fois, et crayon en main ! À ce titre, les exercices font partie intégrante du texte, étant entendu que l'on n'apprend les mathématiques qu'en les pratiquant.

[(2)] Soulignons néanmoins un point important : cet ouvrage ne prétend aucunement présenter des descriptions réalistes de phénomènes naturels. Les modèles utilisés doivent être envisagés comme autant d'exemples *fictifs*, mais néanmoins bien souvent classiques et *similaires dans leur esprit* à ceux utilisés par les véritables spécialistes.

Remerciements

C'est Daniel Guin qui nous a poussés (non sans mal !) à écrire ce livre. Il aura finalement eu gain de cause, et qu'il soit donc remercié pour avoir su nous encourager sans nous décourager. Nos collègues Philippe Castillon, Thomas Hausberger, Pierre-Louis Montagard et Nicolas Saby (du côté mathématique), Jean-Baptiste Ferdy, Bernard Godelle et Agnès Mignot (du côté biologie) ont participé à Montpellier à la réflexion sur les contenus et à la mise en place des enseignements qui ont donné naissance à ce livre. Même s'ils n'ont pas contribué à son écriture, les nombreuses discussions que nous avons eues avec eux ont été très utiles dans la réalisation du projet. Pierre Jacob, qui a gentiment accepté de nous communiquer l'ensemble des notes qu'il avait patiemment écrites à l'occasion d'enseignements de même nature, fait l'objet de toute notre gratitude.

De nombreux autres collègues ont bien voulu prendre sur leur temps afin de nous faire partager leurs avis, remarques ou commentaires sur tout ou partie de l'ouvrage. Citons ainsi Olivier Bouaziz, Claire Coiffard, Robert Eymard, Aurélie Fischer, Alain Prignet, Philippe Saint Pierre et Clara Zelli. À nouveau, une mention spéciale doit être accordée à Jean-Baptiste Ferdy, avec qui nous avons eu de multiples occasions de discuter du contenu du livre, des meilleures manières de l'aborder, et qui nous a entretenu nombre de fois de son expérience d'enseignant et de chercheur en biologie intéressé par les méthodes mathématiques. Les étudiants montpelliérains de licences de biologie, de sciences de la Terre et de chimie biomoléculaire des années 2005 à 2009 ont vécu avec nous l'enseignement de la plus grande partie de ce manuel, nous permettant de corriger de nombreux défauts. Un grand merci en particulier à Sylvain Desruelles pour sa relecture des chapitres 8 à 10.

Enfin, le relecteur nous a fait nombre de critiques constructives, pour lesquelles nous lui sommes infiniment redevables.

This page intentionally left blank

Première partie

Bases

This page intentionally left blank

1

FONCTIONS D'UNE VARIABLE

1.1. Problème : évolution d'un pathogène

On connaît, dans le monde des micro-pathogènes, quelques virus et bactéries extrêmement virulents, qui conduisent généralement à la mort de leur hôte. C'est le cas, par exemple, de la fièvre hémorragique Ebola, du virus de la variole ou encore de certaines formes d'anthrax. Ces pathologies, quoique fortement médiatisées, n'en demeurent pas moins exceptionnelles, et la grande majorité des agents infectieux induisent fort heureusement des maladies aux conséquences moins dramatiques.

La théorie de la sélection naturelle, énoncée par Darwin au XIXe siècle, prédit que l'évolution d'une espèce se fait dans le sens d'une plus grande compétitivité. En d'autres termes, un individu sera d'autant plus favorisé qu'il engendrera davantage de descendants que ses compétiteurs. Transposée au cas des micro-pathogènes, cette « règle » affirme que si plusieurs souches d'une même bactérie ou d'un même virus sont en compétition, la souche favorisée par la sélection naturelle sera la plus apte à se développer et à diffuser au sein de la population. Mais dès lors, comment expliquer que la plupart des microbes soient aussi peu agressifs et virulents ? Interrogé, le biologiste peut fournir un modèle exprimant (avec une formule mathématique) l'efficacité de transmission d'un agent pathogène en fonction de sa virulence intrinsèque. La sélection naturelle tend alors à maximiser cette efficacité, et là réside l'explication de la grande diversité du monde des micro-pathogènes.

L'*étude de fonctions* consiste précisément à mettre en place des techniques permettant d'étudier efficacement le comportement d'une quantité (ici, l'efficacité de transmission) en fonction d'une autre quantité (ici, la virulence du pathogène),

Exercice 2

1. Le composé A subit deux phénomènes : l'ajout en continu (à un débit $d(t)$) et la réaction transformant A (et B) en C. D'après les hypothèses, entre deux instants t et $t+\delta t$ proches, et en notant $q_A(t)$ la quantité de A présente à l'instant t,

$$q_A(t+\delta t) - q_A(t) = d(t)\delta t - k_1 q_A(t)\delta t.$$

En divisant par δt et en le faisant tendre vers 0, on parvient à

$$q_A'(t) = d(t) - k_1 q_A(t).$$

On traite de même les deux autres produits intéressants C et E. Avec des notations évidentes,

$$\begin{aligned} q_C(t+\delta t) - q_C(t) &= k_1 q_A(t)\delta t - k_2 q_C(t)\delta t + k_3 q_E(t)\delta t \\ q_E(t+\delta t) - q_E(t) &= k_2 q_C(t)\delta t - k_3 q_E(t)\delta t, \end{aligned}$$

et finalement

$$q_C'(t) = -k_2 q_C(t) + k_3 q_E(t) + k_1 q_A(t), \quad q_E'(t) = k_2 q_C(t) - k_3 q_E(t).$$

On a donc une équation indépendante (celle sur q_A), qui peut être résolue directement, et un système de deux équations couplées portant sur q_C et q_E, dans lequel q_A doit être traité comme un second membre.

Le modèle en temps continu peut se justifier par le fait que les évolutions de A, C, etc. sont le résultat d'un cumul de très nombreux phénomènes rapides (les réactions à l'échelle moléculaire).

2. Avec le choix de valeurs numériques de l'énoncé, la première équation devient $q_A'(t) = -\frac{1}{10}q_A(t) + \frac{1}{10}$, qui se résout facilement en

$$q_A(t) = 1 + C_A e^{-\frac{1}{10}t}$$

(avec C_A une constante). Le système couplé est alors

$$q_C'(t) = -\tfrac{1}{10}q_C(t) + \tfrac{2}{10}q_E(t) + \tfrac{1}{10} + \tfrac{C_A}{10}e^{-\frac{1}{10}t}, \quad q_E'(t) = \tfrac{1}{10}q_C(t) - \tfrac{2}{10}q_E(t).$$

La matrice carrée pertinente est $\begin{pmatrix} -\frac{1}{10} & \frac{2}{10} \\ \frac{1}{10} & -\frac{2}{10} \end{pmatrix}$, de valeurs propres 0 et $-\frac{3}{10}$, donc diagonalisable, par exemple dans le repère formé des vecteurs propres $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Si l'on note U la matrice de passage vers ce repère (dont les colonnes sont données par ces deux vecteurs), les coefficients du vecteur $\begin{pmatrix} r_0(t) \\ r_1(t) \end{pmatrix} = U^{-1} \begin{pmatrix} q_C(t) \\ q_E(t) \end{pmatrix}$ sont solutions de

$$\begin{aligned} r_0'(t) &= \tfrac{1}{30} + \tfrac{C_A}{30} e^{-\frac{1}{10}t} \\ r_1'(t) &= -\tfrac{3}{10} r_1(t) - \tfrac{1}{30} - \tfrac{C_A}{30} e^{-\frac{1}{10}t}, \end{aligned}$$

et on trouve

$$\begin{aligned} r_0(t) &= C_0 + \tfrac{1}{30}t - \tfrac{C_A}{3} e^{-\frac{1}{10}t} \\ r_1(t) &= C_1 e^{-\frac{3}{10}t} - \tfrac{1}{9} - \tfrac{C_A}{6} e^{-\frac{1}{10}t} \end{aligned}$$

(chercher par exemple une solution particulière de la seconde équation de la forme $t \longmapsto E + F e^{-\frac{1}{10}t}$). Les constantes C_A, C_0 et C_1 se calculent en fonction des conditions initiales : on a $0 = q_A(0) = 1 + C_A$, donc $C_A = -1$ et

$$q_A(t) = 1 - e^{-\frac{1}{10}t}.$$

Par ailleurs, $q_C(0) = q_E(0) = 0$, ce qui implique $r_0(0) = r_1(0) = 0$. En conséquence, $C_0 - \frac{C_A}{3} = 0$ et $C_1 - \frac{1}{9} - \frac{C_A}{6} = 0$, soit finalement $C_0 = -\frac{1}{3}$ et $C_1 = -\frac{1}{18}$. Un calcul direct de $\begin{pmatrix} q_C(t) \\ q_E(t) \end{pmatrix} = U \begin{pmatrix} r_0(t) \\ r_1(t) \end{pmatrix}$ mène alors à

$$\begin{aligned} q_C(t) &= \tfrac{1}{15}t - \tfrac{5}{9} + \tfrac{1}{2} e^{-\frac{1}{10}t} + \tfrac{1}{18} e^{-\frac{3}{10}t} \\ q_E(t) &= \tfrac{1}{30}t - \tfrac{4}{9} + \tfrac{1}{2} e^{-\frac{1}{10}t} - \tfrac{1}{18} e^{-\frac{3}{10}t}. \end{aligned}$$

3. La quantité de composé A reste toujours inférieure à 1 (le débit constant, ici), ce qui est souhaitable si A est toxique. En revanche, les quantités de C et E deviennent aussi grandes que l'on veut en attendant assez longtemps. Une méthode de ce type permet donc