que sais-je ?

LES THÉORIES MATHÉMATIQUES DES POPULATIONS

ALAIN HILLION

le point des connaissances actuelles

PRESSES UNIVERSITAIRES DE FRANCE

Les théories mathématiques des populations

QUE SAIS-JE ?

Les théories mathématiques des populations

ALAIN HILLION
Ancien Elève de l'Ecole Normale Supérieure
Professeur à l'Ecole Nationale Supérieure
des Télécommunications de Bretagne

ISBN 2 13 039193 1

Dépôt légal — 1re édition : 1986, mai

INTRODUCTION

En 1662, paraît, à Londres, un petit livre au titre très long *Observations naturelles et politiques, répertoriées dans l'index ci-après et faites sur les bulletins de mortalité par John Graunt, citoyen de Londres, en rapport avec le gouvernement, la religion, le commerce, l'accroissement, l'atmosphère, les maladies et les divers changements de ladite cité.* L'auteur, un riche marchand drapier, se propose, à partir des bulletins de mortalité qui donnaient le nombre hebdomadaire de décès dans chaque paroisse de Londres, de répondre à cent six questions couvrant les domaines annoncés dans le titre. En déduisant ses réponses de *calculs* effectués sur les chiffres fournis par les bulletins, J. Graunt a introduit les mathématiques dans les réflexions sur l'état ou l'évolution des populations. Certes, Graunt n'utilise pas plus de mathématique que la simple règle de trois, mais il se sert de toutes les ressources de ce qu'il appelle « son arithmétique de boutiquier », avec une honnêteté scientifique remarquable.

La voie était ouverte à un nouveau champ d'application du raisonnement et des théories mathématiques, restreints jusque-là aux domaines de l'abstraction et de la physique. Au fur et à mesure du développement propre de la mathématique, tantôt en profitant, tantôt le stimulant, *Les théories mathématiques des populations* vont présenter des réponses de plus en plus précises aux questions que Graunt se posait sur l'accroissement des populations ou sur les épidémies et s'étendre à d'autres disciplines : la biologie, la génétique, l'écologie...

Le présent ouvrage se propose de donner un aperçu des mécanismes et des modèles les plus courants des théories mathématiques des populations.

CHAPITRE PREMIER

GÉNÉRALITÉS SUR LES MODÈLES MATHÉMATIQUES DE L'ÉVOLUTION DES POPULATIONS

I. — La dynamique des populations

La dynamique des populations est l'étude mathématique de l'évolution de la taille des populations. On entend le terme *population* au sens le plus large (et le plus vague) : ensemble d'éléments (les individus) dont le nombre (la *taille* ou l'effectif de la population) est susceptible de changer au cours du temps. Les grenouilles d'une mare, les personnes atteintes par une épidémie, mais aussi les livres d'une bibliothèque de prêt ou les pièces détachées en stock dans un atelier sont des exemples de population. Chaque population évolue à un rythme et suivant des mécanismes qui lui sont propres : le nombre d'habitants ne s'accroît pas à la même vitesse dans un pays du Tiers Monde que dans un pays d'Europe, la reproduction et la durée de vie ne sont pas identiques chez l'huître et le mouton, la grippe ne se propage pas comme la poliomyélite.

À partir des observations du spécialiste (démographe, biologiste, médecin...), le mathématicien traduit au mieux en termes mathématiques les mécanismes qui régissent l'évolution de la population. Ces dernières équations sont les axiomes de la théorie mathématique (ou *modèle*) de la population. Oubliant, en quelque sorte, la situation concrète qui l'a conduit à les poser, le mathématicien s'emploie, par des raisonnements mathématiques, à tirer alors de ces équations le maximum de renseignements : quelles seront la taille ou la répartition par âge de la population à un instant donné ? La population pourra-t-elle augmenter indéfiniment, sa croissance sera-t-elle limitée ou finira-t-elle par s'éteindre ?, etc.

II. — Les modèles déterministes et les modèles stochastiques

On distingue classiquement deux types de modèles : les *modèles déterministes* (où le hasard n'intervient pas) et les *modèles stochastiques* (ou aléatoires) (où l'on suppose que les mécanismes d'évolution sont entachés de certaines variations dues au hasard).

Les exemples suivants feront sans doute mieux saisir les différences et le lien entre modèle déterministe et modèle stochastique :

Exemple 1 (1). — Deux amis décident d'acheter systématiquement des livres le premier jour ouvrable de chaque mois. Le premier, M. Bouvard, achète un livre tous les mois. Le second, M. Pécuchet, plus fantaisiste, jette une pièce de monnaie en l'air avant de se décider à tout achat : si la

(1) Cet exemple, fondé sur le jet d'une pièce de monnaie, peut paraître au premier abord artificiel. En fait, on montre qu'il est toujours *théoriquement* possible de simuler tout phénomène, dépendant du hasard, aussi complexe soit-il, en utilisant une simple pièce de monnaie.

pièce tombe sur face, il achète deux livres ; si la pièce tombe sur pile, il n'achète aucun livre ce mois-là. Ils se demandent quelles seront, au bout de n mois, les tailles respectives des deux populations que constituent leurs deux bibliothèques.

Il est facile d'écrire les équations qui régissent l'évolution de la bibliothèque de M. Bouvard. Si on note $x(n)$ la taille de la population au cours du n-ième mois, on aura, puisque la bibliothèque augmente d'une unité chaque mois :

$$x(n+1) = x(n) + 1 \quad \text{et} \quad x(0) = 0 \qquad [1]$$

On en déduit $x(1) = 1$, $x(2) = 2$ et donc $x(n) = n$ pour tout n entier.

Il est par contre impossible d'établir une formule qui déterminerait la taille de la bibliothèque de M. Pécuchet au cours du n-ième mois. Ce nombre (appelons-le $X(n)$) dépend en effet des résultats imprévisibles des n jets de pièce. $X(n)$ peut, *a priori*, prendre n'importe quelle valeur paire comprise entre 0 (la pièce sera toujours tombée du côté pile) et $2n$ (la pièce sera toujours tombée du côté face). Comme il n'est sans doute pas égal à M. Pécuchet d'avoir, au bout de trois mois, 0 ou 6 livres, il doit se demander : « Ai-je peu ou beaucoup de "chances" de trouver les rayons de ma bibliothèque encore vides au bout de trois mois ? »

La théorie des probabilités (2) fournit une réponse mathématique aux questions de ce type :

On définit, pour toute expérience dont les résultats dépendent du hasard, un triplet $(\Omega, \mathscr{B}, P)$ (Ω est un ensemble, $\mathscr{B}$ une collection de sous-ensembles de Ω et P une application de $\mathscr{B}$ dans $[0, 1]$).

(2) Cf., dans la même collection, les ouvrages d'A. Jacquard (Les probabilités) et de P. Deheuvels (La probabilité, le hasard et la certitude).

— Ω est l'ensemble de tous les résultats de l'expérience.

— Les éléments de $\mathscr{B}$ sont appelés événements. (Dans le cas où Ω est fini, on s'arrange souvent pour que tout sous-ensemble de Ω soit un événement : on choisit alors $\mathscr{B} = \mathscr{P}(\Omega)$ = ensemble de tous les sous-ensembles de Ω.)

— P associe à tout événement A un nombre compris entre 0 et 1, appelé *probabilité de A* (noté P(A)), qui mesure, avant que l'expérience ait été effectuée, les « chances » de réalisation de l'événement A. Plus P(A) est proche de 0, moins l'événement A a de « chances » de se réaliser ; plus P(A) est proche de 1, plus l'événement A a de « chances » de se réaliser.

On imposera donc à P de vérifier :

$$P(\varnothing) = 0 \quad (\varnothing \text{ est l'ensemble vide})$$

et $\quad P(\Omega) = 1.$

On demande enfin à P de vérifier l'axiome suivant (dit des « *probabilités totales* ») qui traduit que, si des événements sont disjoints, leurs chances s'ajoutent :

$A_n (n \geqslant 0)$ étant une suite d'événements incompatibles deux à deux (c'est-à-dire tels que $A_n \cap A_m = \varnothing$ pour $n \neq m$)

$$P(\bigcup_{n \geqslant 0} A_n) = P(A_0) + P(A_1) + \ldots = \sum_{n \geqslant 0} P(A_n).$$

Dans l'exemple de M. Pécuchet, Ω est l'ensemble de tous les n jets possibles de pièces, c'est-à-dire l'ensemble des suites $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ où ω_i est égal à P ou F suivant que la pièce est tombée, la n-ième fois, sur le côté pile ou sur le côté face.

On peut maintenant préciser un peu la nature de $X(n)$. Quand les n jets ont été effectués $X(n)$ est parfaitement déterminée : $X(n)$ est donc une application de Ω dans N (ensemble des entiers). Une quantité numérique X dépendant du hasard (comme ici $X(n)$) est une *variable aléatoire*. Les expressions du type « X vaut k », « X est plus petit que k » définissent des événements dont on peut calculer la probabilité ; il suffit, en effet, d'écrire : « X vaut

k » $= \{X = k\}$ = Ensemble des ω tels que $X(\omega) = k$.

Comme l'axiome des probabilités totales entraîne $\Sigma P\{X = k\} = 1$, les quantités $P_X(k) = P\{X = k\}$ définissent une probabilité sur N, appelée *loi de la variable aléatoire* X (et notée P_X).

Notant Y_n la variable aléatoire qui vaut 1 si, au n-ième jet, la pièce tombe sur face et qui vaut 0 si elle tombe sur pile, on peut écrire pour la bibliothèque de M. Pécuchet des relations analogues aux équations [1] qui régissent l'évolution de la bibliothèque de M. Bouvard :

$$X(n + 1) = X(n) + 2Y_n \qquad [2]$$

(qui traduit que la bibliothèque n'augmente de deux unités que si la pièce tombe sur le côté face),

et $X(0) = 0$ (3).

On en déduit donc

$$X(n) = 2(Y_1 + Y_2 + \dots + Y_n)$$

et le problème revient à déterminer la loi de S_n : $S_n = Y_1 + Y_2 + \dots + Y_n$, variable aléatoire égale au nombre de fois que le côté face est apparu au cours des n premiers jets. Mais, pour déterminer cette loi, il faut poser une hypothèse mathématique (qui sera en fait la définition de la probabilité P), traduisant que les n jets de pièce sont sans influence les uns sur les autres, et introduire les notions de probabilité conditionnelle et d'indépendance (stochastique).

(3) $X(0)$ est la variable qui ne prend que la valeur $0 (P\{X(0) = 0\} = 1)$. Un nombre n'étant qu'un cas particulier de variable aléatoire, la classe des modèles stochastiques contient la classe des modèles déterministes.

Si A et B sont deux événements (avec $P(B) > 0$) on définira la *probabilité conditionnelle de* A *sachant* B par :

$$P(A/B) = \frac{P(A \cap B)}{P(B)}.$$

L'application qui à A associe $P(A/B)$ est une nouvelle probabilité qui tient compte de l'information supplémentaire qu'apporte la réalisation de l'événement B.

Quand la réalisation de l'événement B n'apporte aucun renseignement nouveau sur la réalisation de A, on dira que A *et* B *sont indépendants*, ce qui s'écrit mathématiquement :

$$P(A/B) = P(A) \quad \text{ou encore} \quad P(A \cap B) = P(A)\,P(B).$$

On dira de même que *deux variables aléatoires* X *et* Y *sont indépendantes* si, pour tous n et m entiers :

$$P\{X = n, Y = m\} = P\{X = n\}\,P\{Y = m\}.$$

Pour traduire que les n jets de pièce sont sans influence les uns sur les autres, on posera donc que les variables aléatoires Y_i $(1 \leqslant i \leqslant n)$ sont (mutuellement) indépendantes, c'est-à-dire que pour tous $k_1, k_2, \ldots, k_n$ égaux à 0 ou 1.

$$P\{Y_1 = k_1, Y_2 = k_2, \ldots, Y_n = k_n\} = \prod_{i=1}^{n} P\{Y_i = k_i\}.$$

C'est alors un des résultats élémentaires (mais très important) (4) du calcul des probabilités que la somme S_n de n variables aléatoires Y_i $(1 \leqslant i \leqslant n)$, indépendantes, ne prenant que les valeurs 0 ou 1, avec les probabilités

$$P\{Y_i = 1\} = p, \quad P\{Y_i = 0\} = q \quad (p + q = 1),$$

obéit à *la loi binomiale* $B(n, p)$, définie par

$$P\{S_n = k\} = C_n^k\, p^k\, q^{n-k} \quad (0 \leqslant k \leqslant n).$$

(4) Cf. A. Jacquard, op. cit., p. 60.

$\mathrm{C}_n^k = \dfrac{n!}{k!(n-k)!}$ est le nombre de combinaisons de n éléments en groupes de k éléments.

Si la pièce de monnaie est équilibrée $\left(p = q = \dfrac{1}{2}\right)$, la loi de $X(n)$, taille de la bibliothèque de M. Pécuchet au cours du n-ième mois, sera déterminée par

$$P\{X(n) = 2k\} = \frac{\mathrm{C}_n^k}{2^n} \quad (0 \leqslant k \leqslant n).$$

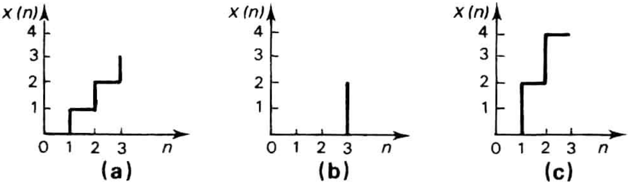

Fig 1. — La courbe (1 a) représente $x(n)$ en fonction du temps n pour $0 \leqslant n \leqslant 3$.
Les courbes (1 b) et (1 c) représentent $X(n)(\omega)$ en fonction du temps n pour les valeurs $\omega = (P, P, F)$ et $\omega = (F, F, P)$.
Les courbes (1 b) et (1 c) sont des trajectoires de $X(n)$.
Il y a autant de trajectoires possibles que de points dans Ω (pour $n = 3$, $2^n = 8$ trajectoires).
Si la pièce est équilibrée, toutes ces trajectoires ont même probabilité $\left(\dfrac{1}{8}\right)$. Une seule trajectoire (celle correspondant à $\omega = (F, F, F)$) aboutit au point 6 pour $n = 3$.

Bien que la seule description complète d'une valeur aléatoire soit la donnée de sa loi (c'est-à-dire des quantités $P\{X = k\}$, il est très utile de pouvoir répondre à des questions du type : autour de quel nombre les valeurs sont-elles le mieux groupées ? Comment s'effectue ce groupement ? On remplacera alors la suite des probabilités $P\{X = k\}$, par des nombres (appelés autrefois, d'une façon très suggestive, « valeurs typiques ») qui donneront une première idée du comportement de la variable aléatoire.

Les deux questions posées plus haut conduisent à introduire les notions très importantes d'espérance mathématique et de variance.

On appelle *espérance mathématique* (ou *moyenne*) d'une variable aléatoire X le nombre (noté E(X)) défini par :

$$E(X) = \Sigma k P\{X = k\}$$

(la sommation étant effectuée sur toutes les valeurs possibles de X).

Plus généralement, l'espérance d'une fonction g de la variable aléatoire X est égale à :

$$E[g(X)] = \Sigma g(k)\, P\{X = k\}.$$

Un cas particulier intéressant est celui des moments de la variable aléatoire X définis par :

$$E(X^n) = \Sigma k^n\, P\{X = n\} \quad (\textit{moment d'ordre } n).$$

On montre alors qu'une variable aléatoire X est la « mieux groupée » (5) autour de sa moyenne $\mu = E(X)$, le groupement étant mesuré par sa *variance* $\sigma^2(X) = E[(X - \mu)^2]$. Plus la variance (qui est un nombre positif) est petite, plus les valeurs de la variable aléatoire sont resserrées autour de la moyenne.

(5) On a, précisément, $\sigma^2(X) = \inf_{a \in R} E[(X - a)^2]$.

Cette propriété de groupement, qui rappelle l'étude du centre de gravité et des moments d'inertie d'un solide, n'est pas la seule qui justifie l'introduction de la notion de moyenne : si on effectue n expériences identiques et indépendantes, la quantité $\dfrac{X_1 + X_2 + \ldots + X_n}{n}$ (moyenne « usuelle » des n résultats X_i, $i \leqslant 1 \leqslant n$, variables aléatoires indépendantes et de même loi que X) se rapproche, quand n devient grand, du nombre $\mu = E(X)$. C'est une des lois des grands nombres (cf. P. Deheuvels, op. cit.).

En particulier, si on note $n(A)$ le nombre de réalisations de l'événement A au cours de n expériences identiques et indépendantes, la fréquence de A $f_n(A) = \dfrac{n(A)}{n}$ se rapproche, quand n devient grand, de P(A). (On peut ainsi expliquer, en étudiant les propriétés de $f_n(A)$, l'axiome des probabilités totales, les notions de probabilité conditionnelle, d'indépendance, d'espérance...)

Comme la moyenne d'une loi binomiale $B(n, p)$ est np, la taille moyenne de la bibliothèque de M. Pécuchet sera égale à :

$$\frac{1}{2} \times 2n = n,$$

c'est-à-dire à la taille de la bibliothèque de M. Bouvard.

On pourra dire que, *en moyenne*, la bibliothèque de M. Pécuchet évolue comme celle de M. Bouvard.

Exemple 2. — Des spécialistes ont résumé leurs longues et sérieuses études sur la durée de vie d'une certaine espèce de papillon en écrivant : « Un papillon sur deux meurt au bout de vingt-quatre heures. »

Essayons de proposer quelques modèles mathématiques qui permettraient, connaissant le nombre de papillons nés à un instant donné (appelons-le $x(0)$) de prévoir le nombre de papillons encore vivants vingt-quatre heures plus tard (appelons-le $x(1)$) :

Modèle 1. — La traduction mathématique immédiate de la phrase des spécialistes est :

$$x(1) = \frac{x(0)}{2},$$

formule qui est tout à fait satisfaisante si $x(0)$ est un nombre pair, mais ennuyeuse à interpréter si $x(0)$ est impair : si $x(0) = 5$, $x(1)$ serait égal à 2,5 (et on aurait deux moitiés de papillon à rassembler pour en faire soit un papillon vivant, soit un papillon mort !). Si $x(0) = 101$, $x(1) = 50{,}5$ (on a encore un papillon dont le destin n'est pas scellé, mais c'est peut-être moins grave pour le spécialiste de ne pas connaître le sort d'un papillon particulier sur une population de 100 que sur une population de 5...).

On est alors tenté d'interpréter de façon un peu plus lâche la phrase des spécialistes et de la lire : « En gros », un papillon sur deux meurt au bout de vingt-quatre heures.

Modèle 2. — La traduction « mathématique » de la nouvelle phrase serait :

$$x(1) \# \frac{x(0)}{2}$$

(où le symbole $\#$, largement utilisé par les physiciens, signifie « peu différent de »).

Les mathématiciens, qui n'aiment pas beaucoup le symbole $\#$, auquel ils n'arrivent pas (et pour cause !) à donner de définition précise, préfèrent tourner un peu la difficulté : comme ils ne peuvent pas bâtir un modèle qui traduise littéralement la phrase des spécialistes et qui décrive exactement l'évolution de la population de papillons (ce serait le modèle 1 qui n'est pas satisfaisant si $x(0)$ est impair), ils vont modifier l'équation :

$$x(1) = \frac{x(0)}{2} \text{ en l'écrivant } x^{\bullet}(1) = \frac{x(0)}{2}.$$

$x^{\bullet}(1)$ n'est qu'une valeur « approchée » de $x(1)$, nombre *entier* de papillons encore vivants à la fin de la journée. On permet donc à $x^{\bullet}(1)$ de prendre des valeurs qui ne sont *pas entières* (rationnelles ici).

Les difficultés rencontrées au cours de l'étude du modèle 1 sont apparemment levées : on calcule $x^{\bullet}(1)$ et on prend pour valeur de $x(1)$ un nombre *entier* proche de $x^{\bullet}(1)$. Mais le choix de $x(1)$ est arbitraire : si $x^{\bullet}(1)$ vaut 50,5 $(x(0) = 101)$, on