

Sur la probabilité d’extinction dans des modèles mathématiques d’épidémies

Nicolas Bacaër * †

Comptes Rendus Biologies 340 (2017) 453-455
doi : 10.1016/j.crvi.2017.09.002

Résumé

On présente une formule pour la probabilité d’extinction d’une épidémie modélisée par un processus de branchement à plusieurs types lorsque ce processus est construit à partir de modèles à compartiments qui sont des systèmes d’équations différentielles.

1 Introduction

De nombreux modèles mathématiques d’épidémies se présentent sous la forme d’un système d’équations différentielles ordinaires non linéaires. Au début de l’épidémie, les personnes infectées, qui peuvent être de plusieurs types, représentent une fraction négligeable de la population, de sorte qu’on peut linéariser le modèle pour obtenir un système linéaire pour les compartiments infectés seuls. Ce système est en général de la forme $\frac{dI}{dt} = (A - B - C)I$, où :

- $I = (I_1, \dots, I_n)$ est un vecteur colonne et $I_i(t)$ est le nombre de personnes infectées de type i ;
- A est une matrice d’infection et $A_{i,j} \geq 0$ est le taux auquel une personne infectée de type j produit de nouvelles personnes infectées de type i ;
- B est une matrice de transfert, $-B_{i,j} \geq 0$ (pour $i \neq j$) est le taux auquel une personne infectée de type j se transforme en une personne infectée de type i , et $B_{j,j} = -\sum_{i \neq j} B_{i,j}$;
- C est une matrice diagonale et $C_{j,j} \geq 0$ est le taux auquel une personne infectée de type j cesse d’être infectée.

À ce système on associe un nombre noté \mathcal{R}_0 à la suite de Lotka et appelé « reproductivité » [1, p. 102]. Ici ce nombre est égal au rayon spectral de la matrice $A(B + C)^{-1}$ [2, §2.1]. Il n’y aura une épidémie que si $\mathcal{R}_0 > 1$. De plus,

*Institut de Recherche pour le Développement, Unité de modélisation mathématique et informatique des systèmes complexes, Bondy, France. Courriel: nicolas.bacaer@ird.fr

†Université de Tokyo, Département de mathématiques, Tokyo, Japon.

la reproductivité \mathcal{R}_0 mesure l'effort nécessaire pour arrêter l'épidémie : il faut diviser la matrice d'infection par \mathcal{R}_0 ou plus.

Au début de l'épidémie, les effets stochastiques sont importants et peuvent conduire à l'extinction de l'épidémie quand bien même $\mathcal{R}_0 > 1$. On peut donc imaginer des modèles qui sont des processus de branchement à plusieurs types en temps continu, construits avec les coefficients des matrices A , B et C [2, 3, 4, 5]. La question était alors de calculer la probabilité d'extinction dans ces modèles si l'on part de (i_1, \dots, i_n) personnes infectées (nombres entiers) à $t = 0$. Cette probabilité a la forme d'un produit $\omega_1^{i_1} \dots \omega_n^{i_n}$. On n'a pas trouvé dans les travaux ci-dessus de formule générale qui relie les probabilités ω_j et les matrices A , B et C . Seuls des exemples particuliers ont été étudiés.

Dans la section 2, on montre ainsi que les probabilités d'extinction ω_j sont, lorsque $\mathcal{R}_0 > 1$, l'unique solution dans le cube $[0, 1]^n$ du problème de point fixe

$$\omega_j = \frac{\sum_i A_{i,j} \omega_i \omega_j - \sum_{i \neq j} B_{i,j} \omega_i + C_{j,j}}{\sum_i A_{i,j} + B_{j,j} + C_{j,j}} \quad (1)$$

avec $1 \leq j \leq n$. Ceci peut aussi s'écrire

$$\sum_i (1 - \omega_i)(A_{i,j} \omega_j - B_{i,j} - C_{i,j}) = 0 \quad (2)$$

avec $1 \leq j \leq n$. Si l'on note $[1 - \omega_i]$ le vecteur ligne $(1 - \omega_1 \dots 1 - \omega_n)$ et $\text{diag}[\omega_i]$ la matrice diagonale avec les ω_i sur la diagonale, le système prend une forme plus compacte :

$$[1 - \omega_i](A \text{diag}[\omega_i] - B - C) = 0. \quad (3)$$

Dans la section 3, on propose des exemples simples d'application de cette formule. Notons que la formule (2) peut se généraliser au cas d'un environnement périodique [6, équation 4].

2 Démonstration

Construisons le modèle stochastique associé naturellement au modèle déterministe. On suppose qu'avec un taux $A_{i,j}$, chaque personne infectée de type j est remplacée en quelque sorte par deux personnes, l'une de type i , l'autre de type j (il y a eu une nouvelle infection). Ceci signifie que la probabilité de cet événement est $A_{i,j} dt$ pendant un intervalle de temps infinitésimal dt . Avec un taux $-B_{i,j}$ pour $i \neq j$, chaque personne infectée de type j se transforme en une personne infectée de type i . Avec un taux $C_{j,j}$, chaque personne infectée de type j cesse d'être infectieuse. Schématiquement,

$$j \xrightarrow[A_{i,j}]{} i + j, \quad j \xrightarrow[-B_{i,j}]{} i \quad (i \neq j), \quad j \xrightarrow[C_{j,j}]{} \emptyset.$$

Comme $-\sum_{i \neq j} B_{i,j} = B_{j,j}$, chaque personne infectée de type j subit un des trois événements ci-dessus avec le taux total

$$\lambda_j = \sum_i A_{i,j} + B_{j,j} + C_{j,j}.$$

Notons $g_j(x_1, \dots, x_n)$ la fonction génératrice du nombre de personnes des différents types $1, 2, \dots, n$, engendrées par une personne de type j selon le schéma ci-dessus. On a

$$g_j(x_1, \dots, x_n) = \frac{1}{\lambda_j} \left(\sum_i A_{i,j} x_i x_j + \sum_{i \neq j} (-B_{i,j}) x_i + C_{j,j} \right).$$

D'après la théorie des processus de branchement à plusieurs types [4, théorème 1.2], on sait que lorsque $\mathcal{R}_0 > 1$, les probabilités $(\omega_1, \dots, \omega_n)$ sont l'unique solution dans le cube $[0, 1]^n$ du problème de point fixe $g_j(\omega_1, \dots, \omega_n) = \omega_j$ pour $1 \leq j \leq n$. Or ceci s'écrit aussi

$$\sum_i A_{i,j} \omega_i \omega_j + \sum_{i \neq j} (-B_{i,j}) \omega_i + C_{j,j} = \omega_j \lambda_j = \omega_j \left(\sum_i A_{i,j} + B_{j,j} + C_{j,j} \right).$$

D'où, en réarrangeant,

$$-\sum_i B_{i,j} \omega_i + C_{j,j}(1 - \omega_j) = \omega_j \sum_i A_{i,j}(1 - \omega_i).$$

Comme $\sum_i B_{i,j} = 0$, ajoutons ce terme dans le membre de gauche :

$$\sum_i B_{i,j}(1 - \omega_i) + C_{j,j}(1 - \omega_j) = \omega_j \sum_i A_{i,j}(1 - \omega_i).$$

Ceci est bien identique à l'équation (2) puisque $C_{i,j} = 0$ pour $i \neq j$.

3 Exemples

Prenons un modèle de type SEIR, qui fait intervenir quatre compartiments : $S(t)$ est le nombre de personnes susceptibles d'attraper la maladie, $E(t)$ le nombre de personnes qui ont été infectées après avoir été exposées lors d'un contact mais qui ne sont pas encore infectieuses, $I(t)$ le nombre de personnes infectieuses et $R(t)$ le nombre de personnes rétablies et immunisées. Notons $N(t) = S(t) + E(t) + I(t) + R(t)$ la population totale, ν le nombre de naissances par unité de temps, a la fréquence des contacts, μ la mortalité naturelle, b le taux auquel les personnes exposées deviennent infectieuses, c le taux auquel les personnes infectieuses guérissent et ε la surmortalité pendant la période infectieuse. Alors on peut proposer le modèle suivant :

$$\begin{aligned} \frac{dS}{dt} &= \nu - a S \frac{I}{N} - \mu S, & \frac{dE}{dt} &= a S \frac{I}{N} - (\mu + b) E, \\ \frac{dI}{dt} &= b E - (\mu + \varepsilon + c) I, & \frac{dR}{dt} &= c I - \mu R. \end{aligned}$$

En l'absence de la maladie, l'état stationnaire est $S = N^* = \nu/\mu$. Au début d'une épidémie, la population est presque entièrement constituée de personnes

susceptibles, de sorte que $S \simeq N \simeq N^*$. On obtient ainsi le modèle linéarisé

$$\frac{dE}{dt} = -(\mu + b)E + aI, \quad \frac{dI}{dt} = bE - (\mu + \varepsilon + c)I.$$

Avec les notations des sections précédentes, on a

$$A = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} b & 0 \\ -b & 0 \end{pmatrix}, \quad C = \begin{pmatrix} \mu & 0 \\ 0 & \mu + \varepsilon + c \end{pmatrix}.$$

Ainsi

$$A(B + C)^{-1} = \begin{pmatrix} \frac{ab}{(b+\mu)(\mu+\varepsilon+c)} & \frac{a}{\mu+\varepsilon+c} \\ 0 & 0 \end{pmatrix}$$

et $\mathcal{R}_0 = ab/[(b + \mu)(\mu + \varepsilon + c)]$. Supposons $\mathcal{R}_0 > 1$. Le système (3) s'écrit

$$(1 - \omega_1 \quad 1 - \omega_2) \left[\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{pmatrix} - \begin{pmatrix} b + \mu & 0 \\ -b & \mu + \varepsilon + c \end{pmatrix} \right] = 0.$$

On trouve donc les deux équations

$$-(b + \mu)(1 - \omega_1) + b(1 - \omega_2) = 0, \quad a\omega_2(1 - \omega_1) - (\mu + \varepsilon + c)(1 - \omega_2) = 0.$$

D'où la solution dans $[0, 1]^2$ qui est

$$\omega_1 = \frac{\mu + b/\mathcal{R}_0}{\mu + b}, \quad \omega_2 = \frac{1}{\mathcal{R}_0}.$$

C'est le même résultat que celui de [4, §3.2], mais nous avons court-circuité avec l'équation (3) la construction des fonctions génératrices.

On calcule de même les probabilités d'extinction pour le modèle de paludisme de [5, §6]. Notons a la fréquence des piqûres infectantes, c_1 la vitesse de guérison des humains, c_2 la mortalité des moustiques, N_1 le nombre d'humains et N_2 le nombre de moustiques. Si I_1 est le nombre d'humains infectés et I_2 le nombre de moustiques infectés, alors le modèle linéarisé est de la forme

$$A = \begin{pmatrix} 0 & a \\ aN_2/N_1 & 0 \end{pmatrix}, \quad B = 0, \quad C = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}.$$

On trouve $\mathcal{R}_0 = a\sqrt{\frac{N_2/N_1}{c_1 c_2}}$. Le système (3) s'écrit

$$(1 - \omega_1 \quad 1 - \omega_2) \left[\begin{pmatrix} 0 & a \\ aN_2/N_1 & 0 \end{pmatrix} \begin{pmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{pmatrix} - \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} \right] = 0.$$

On trouve au bout des calculs

$$\omega_1 = \frac{c_2 + a}{c_2 \mathcal{R}_0^2 + a}, \quad \omega_2 = \frac{c_2 + a/\mathcal{R}_0^2}{c_2 + a}.$$

Néanmoins, même lorsqu'il n'y a que deux types de personnes infectées, le système quadratique de deux équations (2) conduit en général à une équation

polynomiale de degré 4 pour chacune des probabilités ω_j . Comme 1 est toujours une racine, on est ramené à une équation de degré 3 qui ne peut pas en général se réduire plus. C'est le cas par exemple pour un modèle de type SIS ou SIR avec migrations entre deux sites, de sorte que le système linéarisé pour les personnes infectées (I_1, I_2) dans les deux sites est de la forme

$$A = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}, \quad B = \begin{pmatrix} b_1 & -b_2 \\ -b_1 & b_2 \end{pmatrix}, \quad C = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}.$$

Ce n'est que la présence de nombreux zéros dans les matrices A , B et C qui permet des calculs explicites relativement simples dans le cas du modèle SEIR ou du modèle de paludisme. Si l'on se contente de calculs numériques, alors plutôt que d'utiliser le système (3), on obtient le point fixe dans le cube $[0, 1]^n$ du système (1) par de simples itérations en partant de $(x_1, \dots, x_n) = (0, \dots, 0)$.

Remerciements. On remercie le Professeur Hisashi Inaba, de l'Université de Tokyo, pour son invitation à enseigner un module sur la modélisation des épidémies qui a stimulé ces recherches.

Références

- [1] A. Lotka, Théorie analytique des associations biologiques, 2^e partie, Hermann, Paris, 1939.
- [2] L.J.S. Allen, P. van den Driessche, Relations between deterministic and stochastic thresholds for disease extinction in continuous- and discrete-time infectious disease models, *Math. Biosci.* 243 (2013) 99-108.
- [3] G.E.J. Lahodny, L.J.S. Allen, Probability of a disease outbreak in stochastic multipatch epidemic models. *Bull. Math. Biol.* 75 (2013) 1157-1180.
- [4] L.J.S. Allen, *Stochastic Population and Epidemic Models*, Springer, Cham, 2015.
- [5] L.J.S. Allen, A primer on stochastic epidemic models : Formulation, numerical simulation, and analysis, *Infec. Dis. Model.* 2 (2017) 128-142.
- [6] N. Bacaër, E. Ait Dads, Sur la probabilité d'extinction dans un environnement périodique, 2014, <https://hal.archives-ouvertes.fr/hal-01266292>.