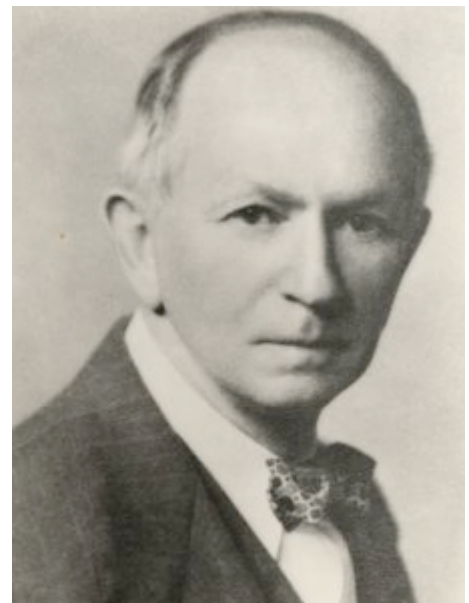


Dynamique des populations

Histoire de R_0 , de Fibonacci à la grippe H1N1

Le 25 novembre 2009, par **Nicolas Bacaër**



En ces temps de grippe, les modélisateurs sont particulièrement sollicités pour essayer de prévoir l'évolution de l'épidémie ou l'effet que pourraient avoir certaines interventions comme la fermeture d'écoles, la restriction des voyages, une campagne de vaccination, etc. Dans le discours de ces apprentis futurologues revient souvent un paramètre connu sous le nom de R_0 (prononcer R-zéro) et qui représente grosso modo le nombre de cas secondaires qu'un malade peut infecter au début de l'épidémie. Cet article essaye de retracer l'histoire de ce paramètre, qui n'est d'ailleurs pas spécifique à l'épidémiologie mais concerne tous les problèmes de dynamique des populations.

AVANT de parler de R_0 , il faut d'abord faire un long détour pour évoquer son proche cousin, le taux de croissance. Commençons avec la suite de Fibonacci. Dans son *Liber Abaci* (1202), Fibonacci considère un problème de lapins, dont la population $p(n)$ le mois n croît suivant l'équation $p(n+1) = p(n) + p(n-1)$, avec $p(0) = 0$ et $p(1) = 1$. Kepler remarquera que $p(n+1)/p(n)$ tend vers le **nombre d'or** $(1 + \sqrt{5})/2$. C'est le taux de croissance mensuel de la population.

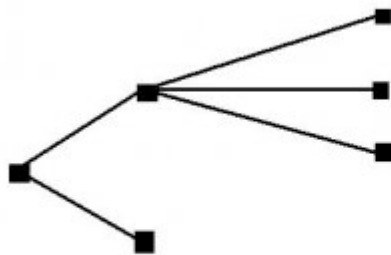
La suite de Fibonacci est évidemment un mauvais modèle pour une population réelle de lapins. Il manque la mortalité. Au XVIII^{ème} siècle, la mortalité commence à être mieux connue, du moins pour les humains. Les premières tables de mortalité sont publiées (Graunt, Halley...), indiquant la probabilité pour un nouveau né d'être encore en vie au bout d'un certain nombre d'années.

Lorsqu'Euler rejoint Berlin en 1741, un de ses collègues à l'académie, Johann Peter Süssmilch, vient de publier le premier traité entièrement consacré à la démographie. La population de Berlin est recensée en 1747. En 1748, Euler publie son « Introduction à l'analyse des infiniment petits ». Le chapitre sur l'exponentielle et le logarithme discute plusieurs questions démographiques, se basant sur l'hypothèse d'une population $p(n)$ croissant d'année en année suivant la formule $p(n+1) = r p(n)$ avec $r > 1$. En 1760, Euler découvre une relation entre le taux de croissance r , le nombre N de naissances par habitant et par année, et la probabilité $q(k)$ d'être en vie k années après sa naissance : $1 = N(1 + q(1)/r + q(2)/r^2 + \dots + q(k)/r^k + \dots)$. En effet, le nombre de naissances croît aussi avec le taux r . La population P est donc composée de $\nu = NP$

nouveaux-nés, des enfants de un an (il y en a $q(1)v/r$), des enfants de deux ans (il y en a $q(2)v/r^2$), etc.

Ce paramètre r va connaître son heure de gloire avec Malthus et son « Essai sur le principe de population » (1798). Les conséquences sociales, qui selon Malthus résultent du déséquilibre entre la croissance exponentielle de la population et la croissance plus lente de la production de nourriture, seront la cible des critiques de Marx. Le livre de Malthus inspirera aussi à Darwin sa théorie de l'évolution par la sélection naturelle (« L'origine des espèces », 1859). Dans les années 1920, alors que se développe une approche mathématique de l'évolution, Ronald Fisher donnera à r le nom de « paramètre malthusien ».

Un passage dans la seconde édition (1803) du livre de Malthus incite également certains à s'intéresser à la question de l'extinction des noms de famille. En 1845, Bienaymé montre que si tout homme a une probabilité $a(k)$ d'avoir k fils, alors un homme verra sa descendance mâle s'éteindre à coup sûr si le nombre moyen de fils $a(1) + 2a(2) + \dots + ka(k) + \dots$ est inférieur ou égal à 1. C'est ce nombre que l'on appellera R_0 . Si au contraire R_0 est strictement supérieur à 1, alors l'extinction se produit avec une certaine probabilité comprise entre 0 et 1, et qui peut se calculer. Le même problème sera étudié en 1873 par Galton et Watson. En 1927, Haldane utilise un modèle analogue pour analyser l'apparition de gènes mutants. Après les travaux de Steffensen puis de Kolmogorov, on parlera de théorie des processus de branchement.



Réalisation d'un processus de branchement avec $R_0 = 2$.

Le premier cas infecte deux personnes. L'un de ces deux cas en infecte trois autres.

En 1911, l'étude du taux de croissance est également reprise par un chimiste américain, Lotka, dans le cadre d'un modèle en temps continu. Si $N(t)$ est le nombre de filles nées par unité de temps à l'instant t , si $q(x)$ est la probabilité pour une femme d'être encore en vie à l'âge x , si a est l'âge minimal et ω l'âge maximal auxquels une femme peut donner naissance à un enfant, et si $f(x)$ est la fertilité (plus exactement la probabilité par unité de temps de donner naissance à une fille) à l'âge x , alors $N(t)$ vérifie l'équation intégrale

$$N(t) = \int_a^\omega f(x)q(x)N(t-x)dx.$$

En effet, il y a $N(t-x)dx$ naissances féminines entre les instants $t-x$ et $t-x+dx$. Une fraction $q(x)$ de ces nouveaux-nés survit jusqu'à l'âge x , donnant naissances à

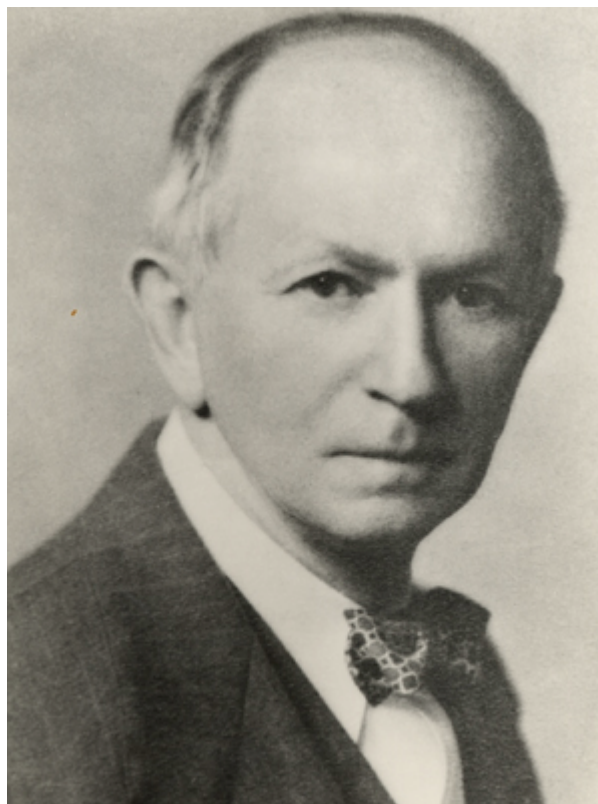
$f(x)q(x)N(t-x)dx$ filles par unité de temps à l'instant t . Lotka devine que $N(t)$ croît asymptotiquement comme $\exp(rt)$, avec un taux instantané de croissance r qui est la solution de l'équation

$$1 = \int_{\alpha}^{\omega} f(x)q(x) \exp(-rx) dx. \quad (1)$$

C'est l'analogie en temps continu de l'équation trouvée par Euler. Dans ce cas, chaque fille peut à sa naissance espérer avoir à son tour

$$R_0 = \int_{\alpha}^{\omega} f(x)q(x)dx$$

filles, et $r > 0$ équivaut à $R_0 > 1$. C'est Lotka qui introduit la notation R_0 , qui est encore utilisée aujourd'hui.



Alfred James Lotka (1880-1949)

En 1927, un médecin et un chimiste britanniques, McKendrick et Kermack, travaillent sur un modèle d'épidémie. Ils se rendent compte que si $I(t)$ est l'incidence (le nombre de nouveaux cas par unité de temps) à l'instant t , si $q(x)$ est la probabilité d'être infectieux x unités de temps après l'infection, si ω est la durée maximale de l'infection, et si $f(x)dx$ est la probabilité d'infecter une autre personne entre x et $x + dx$ unités de temps après l'infection, alors

$$I(t) = \int_0^{\omega} f(x)q(x)I(t-x)dx.$$

C'est formellement la même équation que celle obtenue par Lotka (mais avec $\alpha = 0$). Dans ce cas, r est le taux de croissance de l'épidémie et R_0 le nombre de cas secondaires infectés par chaque

nouveau cas. La différence principale avec la démographie, c'est que l'équation intégrale linéaire pour $I(t)$ n'est valable qu'au tout début de l'épidémie, lorsque le nombre de personnes infectées n'est qu'une petite fraction de la population. Au delà, on n'échappe pas à des équations non linéaires.

En 1941, Feller démontre rigoureusement ce que Lotka avait affirmé concernant son équation intégrale. Dans les années 60, Crump, Mode et Jagers développent avec la même rigueur l'approche probabiliste des modèles de population en temps continu. À ceux-ci peuvent d'ailleurs être associé un processus en temps discret, séparant les générations, comme dans les travaux de Bienaymé, Galton et Watson.

Vers la fin des années 70, les épidémiologistes commencent à s'intéresser de plus près au paramètre R_0 . Il s'agit de l'estimer numériquement pour différentes maladies infectieuses. Si on suppose pour simplifier que la fonction $f(x)$ représentant les contacts ne dépend pas du temps x écoulé depuis l'infection, et si l'on suppose que la guérison a lieu à vitesse constante de sorte que $q(x) = \exp(-ax)$ et $\omega = +\infty$, alors l'équation (1) donne $r = f - a$, tandis que $R_0 = f/a$. On estime d'abord le taux de croissance r en ajustant une exponentielle au début de la courbe épidémique. Sachant à peu près la durée moyenne de l'infection $1/a$, on en déduit f . Puis on calcule R_0 . L'intérêt de R_0 par rapport à r vient du fait que R_0 a une interprétation directe en terme d'intervention : si on divise les contacts par un nombre supérieur à R_0 , alors l'infection ne peut plus se propager (le R_0 après intervention devient inférieur à 1). Cela peut être obtenu en vaccinant une fraction de la population supérieure à $1 - 1/R_0$.

C'est ainsi qu'on lit quelques fois dans la presse des phrases du genre « Il faut savoir qu'un enfant malade est susceptible de contaminer 2,7 autres enfants. » (Le Monde, 15/09/2009, « Deux approches d'experts face à la grippe A »). Ou encore : « Ce taux de 50 % a été calculé en fonction du taux de reproductivité de l'agent infectieux inoculé dans une population saine » (Le Monde, 26/06/2009, « Grippe A : la moitié de la population française est menacée »). Dans la première citation, « l'expert » veut dire que $R_0 = 2.7$. Dans la seconde citation, un autre « expert » se base probablement sur un $R_0 = 2$ (qu'il appelle « taux de reproductivité ») pour trouver $1 - 1/R_0 = 50\%$!

Évidemment, tous ces petits calculs peuvent paraître bien simplistes par rapport à la multitude de facteurs qui peuvent influencer sur l'épidémie. Didier Raoult, spécialiste des maladies infectieuses à l'hôpital de la Timone à Marseille, estime : « Mais depuis l'apparition du virus, rien ne s'est vraiment passé comme prévu, les modélisations mathématiques qui ont été produites ressemblent plus à des prédictions de Nostradamus qu'autre chose » (Le Monde, 15/09/2009, *ibid.*).

Le travail des modélisateurs s'oriente soit vers la simulation de modèles « hyperréalistes » (mais comprenant beaucoup de paramètres inconnus), soit vers l'analyse mathématique de modèles plus simples intégrant pas à pas de nouveaux détails. Pour l'heure, il n'est pas très clair ce qu'il faudrait penser de tous ces efforts. À quel stade de développement se situe la prévision épidémique, comparée par exemple à la prévision en mécanique céleste ou à la prévision météorologique ?

Pour en savoir plus

R.M. Anderson, R.M. May *Infectious Diseases of Humans : Dynamics and Control*. Oxford University Press (1991)

N. Bacaër *Histoires de mathématiques et de populations*. Editions Cassini, Paris (2009)

N. Bacaër *On the final size of epidemics with seasonality.* Bull Math Biol 71 (2009) 1954-66.

O. Diekmann, J.A.P. Heesterbeek *Mathematical Epidemiology of Infectious Diseases.* Wiley, Chichester (2000)

J.A.P. Heesterbeek *A brief history of R_0 and a recipe for its calculation.* Acta Biotheoretica 50 (2002) 189-204.

Organisation Mondiale de la Santé *Modélisation mathématique de la grippe pandémique H1N1 2009.* Revue Epidémiologique Hebdomadaire, 21 Aout 2009 ; vol 84(34):341-8.
<http://www.who.int/wer/2009/wer8434/fr/>

Bacaër Nicolas (2009)

Dynamique des populations : histoire de R_0 , de Fibonacci à la grippe H1N1

Images des Mathématiques - CNRS, 4 p.