

O pico epidêmico em um modelo SIR

Quadrature 117 (2020)

<https://hal.archives-ouvertes.fr/hal-02518993>

Nicolas Bacaër

Institut de recherche pour le développement
Unidade matemática e de modelagem computacional de sistemas complexos
Les Cordeliers, Paris, França
nicolas.bacaer@ird.fr

resumo

Estamos estudando uma epidemia modelada por um sistema diferencial tipo SIR. Quando a população N é grande, o pico da epidemia é no tempo T , com $T \sim (\ln N)/(a - b)$. a é a taxa de contato efetiva. b é a taxa na qual as pessoas infecciosas se recuperam.

Introdução

Vamos considerar, por exemplo, em [1, 2, 6, 7, 8] o modelo SIR de Kermack e McKendrick para uma epidemia

$$\frac{dS}{dt} = -a \frac{SI}{N}, \quad \frac{dI}{dt} = a \frac{SI}{N} - bI, \quad \frac{dR}{dt} = bI.$$

$S(t)$ é o número de pessoas saudáveis. $I(t)$ é o número de pessoas infectadas. $R(t)$ é o número de pessoas removidas da transmissão como resultado de isolamento, cura ou morte. $N = S(t) + I(t) + R(t)$ é a população total. Presume-se que essa população total seja constante e grande o suficiente para tornar razoável modelar a epidemia usando um sistema diferencial em vez de um processo estocástico. Considere as condições iniciais

$$S(0) = N - 1, \quad I(0) = 1, \quad R(0) = 0,$$

que correspondem ao início da epidemia de um único "paciente zero".

O parâmetro a (com $a > 0$) é realmente o produto de dois números: o número de contatos por unidade de tempo e a probabilidade de transmissão durante o contato entre uma pessoa não infectada e uma pessoa infectada. Os contatos são feitos aleatoriamente. I/N é a probabilidade de uma pessoa estar no compartimento *eu*. Isso explica o termo quadrático em SI/N no sistema diferencial. $b > 0$ é a taxa na qual as pessoas saem do compartimento I . A suposição subjacente é, portanto, que o tempo gasto no compartimento I é distribuído de acordo com uma lei exponencial do parâmetro b . Assumiremos $a > b$ para que uma epidemia real ocorra.

Um exemplo é detalhado na figura 1. Usamos o software Scilab e sua função de resolução de sistemas diferenciais. Esta função implementa um esquema de preditor-corretor de Adams para sistemas não íngremes. Nós escolhemos $N = 65 \times 10^6$ (a população da França), $a = 1/2$ por dia e $b = 1/4$ por dia. No início da epidemia, uma pessoa infectada infecta, em média, uma pessoa a cada dois dias. Duração média da infecção $1/b$ são 4 dias. Isso corresponde a uma reprodutibilidade $\mathcal{R}_0 = a/b = 2$ (o termo e a notação vêm de [9, p. 102]). \mathcal{R}_0 é, portanto, o número médio de casos secundários infectados por uma pessoa infectada no início da epidemia. Sem pretender ser realista, são valores da mesma ordem de magnitude que pode ser encontrada para a epidemia de coronavírus em 2020 no caso hipotético em que não há intervenção para limitar a epidemia [5]

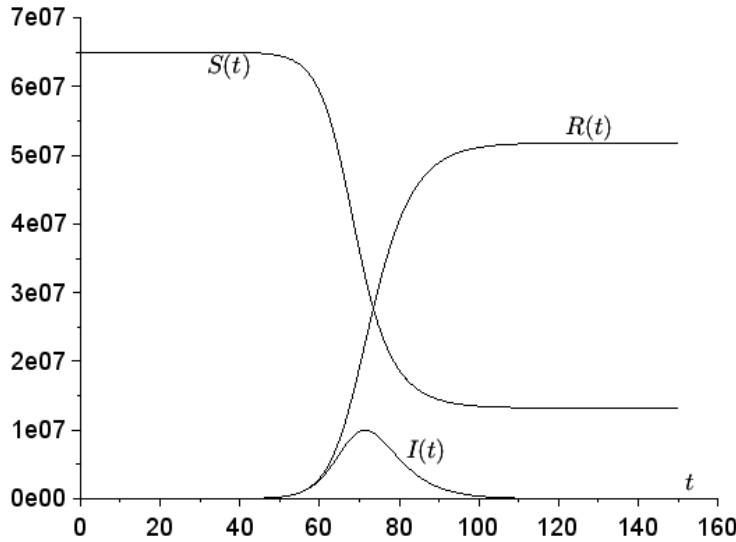


Figura 1. Uma simulação do modelo SIR com o tempo t em dias na abcissa .

O pico da epidemia

Combinando a primeira e a terceira equações do sistema SIR, é bem conhecido [8, p. 76] que encontramos uma primeira integral:

$$\frac{1}{S} \frac{dS}{dt} = -a \frac{I}{N} = -\frac{a}{bN} \frac{dR}{dt}.$$

Então nós temos

$$\ln[S(t)/S(0)] = -\frac{a}{bN} R(t).$$

Voltando à primeira equação, encontramos

$$\frac{dS}{dt} = -a \frac{S}{N} (N - S - R) = -a \frac{S}{N} \left(N - S + \frac{bN}{a} \ln[S(t)/S(0)] \right).$$

O pico da epidemia na Figura 1 corresponde a $dI/dt = 0$, o que acontece de acordo com a segunda equação quando $S(t) = Nb/a$. $S(t)$ é uma função decrescente porque $dS/dt \leq 0$ da primeira equação. Deduzimos que, se T é a hora do pico da epidemia, ou seja, se

$$I(T) = \max_{t>0} I(t),$$

nós temos então

$$T = \int_0^T dt = \int_{S(0)}^{bN/a} \frac{dS}{-a \frac{S}{N} \left(N - S + \frac{bN}{a} \ln[S/S(0)] \right)}.$$

Nós definimos $s = S/N$. Nós temos então

$$T = \frac{1}{a} \int_{b/a}^{1-1/N} \frac{ds}{s \left(1 - s + \frac{b}{a} \ln[s/(1-1/N)] \right)}. \quad (1)$$

Procuramos um limite superior de T . Para tudo $b/a < s < 1 - 1/N$, temos

$$\frac{1}{s \left(1 - s + \frac{b}{a} \ln[s/(1-1/N)] \right)} \leq \frac{1}{s \left(1 - s + \frac{b}{a} \ln s \right)}.$$

No bairro de $s = 1$, temos

$$\frac{1}{s \left(1 - s + \frac{b}{a} \ln s \right)} \sim \frac{1}{(1-s)(1-b/a)}.$$

A integral

$$\int_{b/a}^1 \frac{ds}{1-s}$$

é divergente. Com $N \rightarrow +\infty$, Então nós temos

$$T \leq \frac{1}{a} \int_{b/a}^{1-1/N} \frac{ds}{s(1-s + \frac{b}{a} \ln s)} \sim \frac{1}{a} \int_{b/a}^{1-1/N} \frac{ds}{(1-s)(1-b/a)} \sim \frac{\ln N}{a-b}.$$

Procuramos um limite inferior da T . Nós temos $\ln s \leq s-1$ para tudo $s > 0$. Nos também temos $\ln(1-1/N) \geq -2/N$ para todos $N \geq 2$. Então nós temos para $s \in]b/a, 1-1/N[$ e para qualquer número inteiro $N \geq 2$,

$$\begin{aligned} 1-s + \frac{b}{a} \ln[s/(1-1/N)] &= 1-s + \frac{b}{a} \ln s - \frac{b}{a} \ln(1-1/N) \\ &\leq 1-s + \frac{b}{a}(s-1) + \frac{2b}{aN}. \end{aligned}$$

Então nós temos

$$T \geq \frac{1}{a} \int_{b/a}^{1-1/N} \frac{ds}{s[1 - \frac{b}{a} + \frac{2b}{aN} - (1-b/a)s]}.$$

Nos deparamos com uma fração racional da forma $\frac{1}{s(u-vs)}$ com

$$u = 1 - \frac{b}{a} + \frac{2b}{aN}, \quad v = 1 - b/a.$$

Divide-se em elementos simples

$$\frac{1}{s(u-vs)} = \frac{1}{u} \left(\frac{1}{s} + \frac{v}{u-vs} \right).$$

Um primitivo é

$$\frac{1}{u} [\ln s - \ln(u-vs)] = \frac{1}{u} \ln \frac{s}{u-vs}.$$

Então nós temos

$$\begin{aligned} T &\geq \frac{1}{a(1 - \frac{b}{a} + \frac{2b}{aN})} \ln \left\{ \frac{1-1/N}{b/a} \frac{1 - \frac{b}{a} + \frac{2b}{aN} - (1-b/a)b/a}{1 - \frac{b}{a} + \frac{2b}{aN} - (1-b/a)(1-1/N)} \right\} \\ &= \frac{1}{a-b+2b/N} \ln \left\{ \frac{1-1/N}{b/a} \frac{(1-b/a)^2 + \frac{2b}{aN}}{(1+b/a)/N} \right\} \\ &= \frac{1}{a-b+2b/N} \left[\ln N + \ln \left\{ \frac{1-1/N}{b/a} \frac{(1-b/a)^2 + \frac{2b}{aN}}{(1+b/a)} \right\} \right]. \end{aligned}$$

O membro à direita é $\sim (\ln N)/(a-b)$ E se $N \rightarrow +\infty$. Como é o mesmo equivalente ao do limite superior obtido acima, concluímos que $T \sim (\ln N)/(a-b)$ E se $N \rightarrow +\infty$.

No exemplo da Figura 1, temos $T \simeq 71,4$ dias enquanto $(\ln N)/(a-b) \simeq 72,0$ dias. A aproximação parece bastante boa. Na Figura 2, variamos o tamanho N da população e o parâmetro a , com b fixo. Nós notamos $(\ln N)/(a-b)$ às vezes superestimada, por vezes subestimada T .

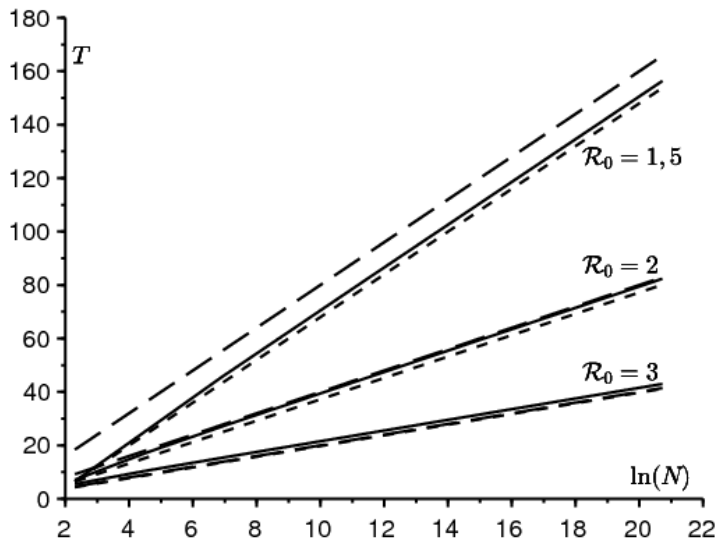


Figura 2. Comparação entre, por um lado, a data T do pico da epidemia (linha contínua), calculada com a integral da fórmula (1), por outro, o equivalente $(\ln N)/(a - b)$ (linhas pontilhadas grandes) e a aproximação $[\ln(2N(b/a - 1)^2)]/(a - b)$ (pequenas linhas pontilhadas), dependendo da $\ln N$, E se $10 \leq N \leq 10^8$ para três valores de parâmetro a com reprodutibilidade $\mathcal{R}_0 = a/b \in \{1,5 ; 2 ; 3\}$.

$a - b$ é a taxa de crescimento inicial da epidemia. Temos então para N grande, $S(t) \simeq N$ e $dI/dt \simeq (a - b)I$. Essa taxa pode ser estimada a partir de dados epidemiológicos. A população N é conhecida. Portanto, podemos prever a data do pico da epidemia. Na prática, a hipótese subjacente ao modelo de uma mistura homogênea da população permanece duvidosa para um país inteiro, um pouco menos para uma cidade.

Historicamente, Kermack e McKendrick haviam resolvido o sistema SIR, assumindo que a reprodutibilidade era próxima de 1, o que levou a uma pequena epidemia. Com a aproximação $e^x \simeq 1 + x + x^2/2$, eles chegaram a uma equação de Riccati explicitamente solucionável, de modo que

$$I(t) \simeq \frac{X}{\text{ch}^2(Yt - Z)}. \quad (2)$$

X , Y e Z são funções complicadas dos parâmetros N , a e b . Veja por exemplo [3]. O pico da epidemia é $T \simeq Z/Y$. Observe de passagem que, até onde sabemos, a aproximação (2) nunca foi rigorosamente demonstrada, razão pela qual usamos o sinal informal acima e abaixo \simeq . Agora assumimos, além $1/N \ll (\mathcal{R}_0 - 1)^2$. Nesse caso,

$$Y \simeq \frac{a - b}{2}, \quad Z \simeq \frac{1}{2} \ln[2N(b/a - 1)^2]$$

(veja por exemplo [4]), para que

$$T \simeq \frac{\ln[2N(b/a - 1)^2]}{a - b}.$$

Essa aproximação também é mostrada na Figura 2; parece particularmente próximo do verdadeiro valor de T se \mathcal{R}_0 está perto de 1, como seria de esperar. Com N que converge para o infinito, ainda assim temos $T \simeq \frac{\ln N}{a-b}$. Encontramos o mesmo comportamento assintótico. Nossa análise baseada na fórmula (1) é válida, no entanto, para todos $\mathcal{R}_0 > 1$ e não apenas para \mathcal{R}_0 perto de 1. Além disso, nossa prova é rigorosa.

Conclusão

Descobrimos que o pico da epidemia ocorreu em um momento $T \sim (\ln N)/(a - b)$ se N converge para o infinito. Este resultado parece novo, apesar de sua simplicidade. Na literatura científica, encontramos apenas informações sobre a altura do pico, que é facilmente determinada pela primeira integral do sistema. De fato, com

$$S(T) = \frac{Nb}{a}, \quad \ln \frac{S(T)}{S(0)} = -\frac{a}{bN}R(T), \quad N = S(T) + I(T) + R(T),$$

nós obtemos

$$\frac{I(T)}{N} = 1 - \frac{b}{a} + \frac{b}{a} \ln \left(\frac{N}{N-1} \frac{b}{a} \right) \simeq 1 - \frac{b}{a} + \frac{b}{a} \ln \frac{b}{a}$$

(veja por exemplo [1, equação 7.26]). Em algumas referências como [7, p. 238], existem inversões onde o tempo t é uma função das variáveis monotônicas R ou S , o que leva a integrais como (1), exceto por uma mudança de variável. Mas o sistema diferencial SIR foi escrito lá na forma $dS/dt = -\alpha SI$. A ideia provavelmente não veio de tomar $\alpha = a/N$, que é de alguma forma mais realista e necessário para que os limites de uma grande população produzam resultados interessantes.

Para uma epidemia como a do coronavírus, o modelo SIR carece de um pouco de realismo. Isso leva a considerar modelos com mais compartimentos, por exemplo, com um compartimento adicional para pessoas na fase latente infectadas, mas ainda não contagiosas. Veja sobre este assunto [5], que também tenta estimar os parâmetros da curva epidêmica. O estudo analítico do pico epidêmico nesse contexto é, sem dúvida, mais difícil.

obrigado

Nós agradecemos Hisashi Inaba por seus comentários sobre o texto.

Referências bibliográficas

1. B. Anselme, Biomathématiques, Dunod, Paris, 2015.
2. N. Bacaër, Histoires de mathématiques et de populations, Cassini, Paris, 2009.
3. N. Bacaër (2012) Le modèle de Kermack et McKendrick pour la peste à Bombay et la reproductivité nette d'un type avec saisonnalité. *Journal of Mathematical Biology* 64 (2012) 403 – 422. <https://hal.archives-ouvertes.fr/hal-01340008>
4. N. Bacaër (2019) Sur la taille finale des épidémies dans un environnement périodique. *Comptes Rendus Biologies* 342 (2019) 119 – 123. <https://doi.org/10.1016/j.crvi.2019.07.001>
5. N. Bacaër (2020) Un modèle mathématique de l'épidémie de coronavirus en France. <https://hal.archives-ouvertes.fr/hal-02509142>
6. M. Corlosquet – Habart, J. Janssen, R. Manca, Modélisation stochastique du risque de pandémie : stratégies de couverture et d'assurance, Lavoisier, Cachan, 2012.
7. J. Gani, Processus stochastiques de population, P. –L. Hennequin (éd.), École d'été de probabilités de Saint – Flour IV, Springer, Berlin, 1975, p. 188 – 293.
8. A. Hillion (1986) Les Théories mathématiques des populations. Presses Universitaires de France, Paris.
9. A. J. Lotka (1939) Théorie analytique des associations biologiques, 2e partie. Hermann, Paris.