

El pico epidémico en un modelo SIR

Quadrature 117 (2020)

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

Nicolas Bacaër

Institut de recherche pour le développement
Unidad de modelado matemático y computacional de sistemas complejos
Les Cordeliers, París, Francia
nicolas.bacaer@ird.fr

resumen

Estamos estudiando una epidemia modelada por un sistema diferencial tipo SIR. Cuando la población N es grande, el pico epidémico es en el momento T , con $T \sim (\ln N)/(a - b)$. a es la tasa de contacto efectiva b es la velocidad a la que se recuperan las personas infecciosas.

Introducción

Consideremos, por ejemplo, en [1, 2, 6, 7, 8] el modelo SIR de Kermack y McKendrick para una epidemia

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

$S(t)$ es el número de personas sanas $I(t)$ es la cantidad de personas infectadas. $R(t)$ es el número de personas eliminadas de la transmisión como resultado del aislamiento, la curación o la muerte. $N = S(t) + I(t) + R(t)$ es la población total Se supone que esta población total es constante y lo suficientemente grande como para que sea razonable modelar la epidemia utilizando un sistema diferencial en lugar de un proceso estocástico. Considere las condiciones iniciales.

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

que corresponden al inicio de la epidemia de un solo "paciente cero".

El parámetro a (con $a > 0$) es en realidad el producto de dos números: el número de contactos por unidad de tiempo y la probabilidad de transmisión durante el contacto entre una persona no infectada y una persona infectada. Los contactos se hacen al azar. I/N es la probabilidad de que una persona sea en el compartimento I . Esto explica el término cuadrático en SI/N en el sistema diferencial. $b > 0$ es la velocidad a la que las personas dejan el compartimento I . Por lo tanto, el supuesto subyacente es que el tiempo que se pasa en el compartimento I se distribuye de acuerdo con una ley exponencial del parámetro b . Asumiremos $a > b$ para que ocurra una verdadera epidemia.

Un ejemplo se detalla en la figura 1. Utilizamos el software Scilab y su función de resolución de sistemas diferenciales. Esta función implementa un esquema predictor-corrector Adams para sistemas no empinados. Nosotros elegimos $N = 65 \times 10^6$ (la población de Francia), $a = 1/2$ por día y $b = 1/4$ por día. Al comienzo de la epidemia, una persona infectada infecta a un promedio de una persona cada dos días. Duración promedio de la infección. $1/b$ Son 4 días. Esto corresponde a una reproducibilidad $\mathcal{R}_0 = a/b = 2$ (el término y la notación provienen de [9, p. 102]). \mathcal{R}_0 es, por lo tanto, el número promedio de casos secundarios que una persona infectada infecta al comienzo de la epidemia. Sin pretender ser realistas, estos son valores del mismo orden de magnitud que los que se pueden encontrar para la epidemia de coronavirus de 2020 en el caso hipotético donde no hay intervención para limitar la epidemia [5]

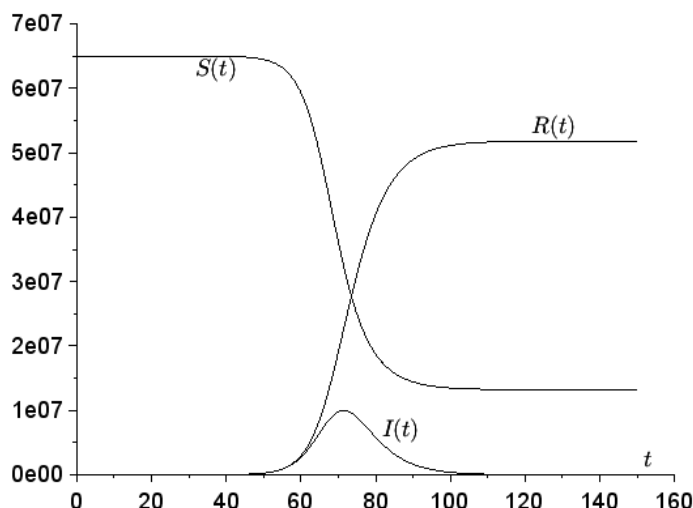


Figura 1. Una simulación del modelo SIR con el tiempo t en días en la abscisa .

El pico epidémico

Al combinar la primera y la tercera ecuaciones del sistema SIR, es bien conocido [8, p. 76] que encontramos una primera integral:

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

Entonces tenemos

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

Volviendo a la primera ecuación, encontramos

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

El pico epidémico en la Figura 1 corresponde a $dI/dt = 0$, que sucede de acuerdo con la segunda ecuación cuando $S(t) = Nb/a$. $S(t)$ es una función decreciente porque $dS/dt \leq 0$ de la primera ecuación. Deducimos que si T es el momento del pico epidémico, es decir, si

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

tenemos entonces

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

Definimos $s = S/N$. Tenemos entonces

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

Buscamos un límite superior de la T . Por todo $b/a < s < 1 - 1/N$, se tiene

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

En el barrio de $s = 1$, se tiene

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

La integral

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

es divergente Con $N \rightarrow +\infty$, Entonces tenemos

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

Buscamos un límite inferior de la *camiseta*. Se tiene $\ln s \leq s - 1$ por todo $s > 0$. También tenemos $\ln(1-1/N) \geq -2/N$ para todos $N \geq 2$. Entonces tenemos para $s \in]b/a, 1 - 1/N[$ y para cualquier número entero $N \geq 2$,

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

Entonces tenemos

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

Nos encontramos con una fracción racional de la forma $\frac{1}{s(u-vs)}$ con

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

Se descompone en elementos simples.

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

Un primitivo es

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

Entonces tenemos

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

El miembro de la derecha es $\sim (\ln N)/(a-b)$ Si $N \rightarrow +\infty$. Como es el mismo equivalente que el del límite superior obtenido anteriormente, concluimos que $T \sim (\ln N)/(a-b)$ Si $N \rightarrow +\infty$.

En el ejemplo de la Figura 1, tenemos $T \simeq 71,4$ días mientras $(\ln N)/(a-b) \simeq 72,0$ días. La aproximación parece bastante buena. En la Figura 2, variamos el tamaño N de la población, así como el parámetro a , con b fijo. Observamos $(\ln N)/(a-b)$ sobreestimado veces, a veces subestimado T .

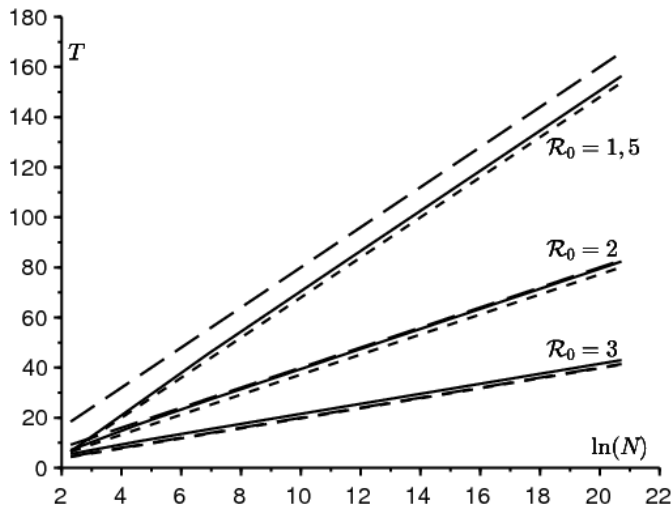


Figura 2. Comparación entre, por un lado, la fecha T del pico de la epidemia (línea continua), calculada con la integral de la fórmula (1), por otro lado, el equivalente $(\ln N)/(a-b)$ (grandes líneas de puntos) y la aproximación $[\ln(2N(b/a-1)^2)]/(a-b)$ (pequeñas líneas punteadas), dependiendo de $\ln N$, Si $10 \leq N \leq 10^8$ para tres valores de parámetros a con reproducibilidad $\mathcal{R}_0 = a/b \in \{1,5; 2; 3\}$.

$a-b$ es la tasa de crecimiento inicial de la epidemia. Entonces tenemos para N grande, $S(t) \simeq N$ y $dI/dt \simeq (a-b)I$. Esta tasa puede estimarse a partir de datos epidemiológicos. La población N es conocida. Por lo tanto, podemos predecir la fecha del pico de la epidemia. En la práctica, la hipótesis subyacente al modelo de una mezcla homogénea de la población sigue siendo dudosa para todo un país, un poco menos para una ciudad.

Históricamente, Kermack y McKendrick habían resuelto aproximadamente el sistema SIR suponiendo que la reproducibilidad era cercana a 1, lo que condujo a una pequeña epidemia. Con la aproximación $e^x \simeq 1 + x + x^2/2$, habían llegado a una ecuación de Riccati explícitamente solucionable, de modo que

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

X , Y y Z son funciones complicadas de los parámetros N , a y b . Ver por ejemplo [3]. El pico epidémico es $T \simeq Z/Y$. Tenga en cuenta de paso que, a nuestro entender, la aproximación (2) nunca se ha demostrado rigurosamente, por lo que utilizamos el signo informal arriba y abajo \simeq . Ahora asumimos además $1/N \ll (\mathcal{R}_0 - 1)^2$. En este caso,

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

(ver por ejemplo [4]), para que

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

Esta aproximación también se muestra en la Figura 2; parece particularmente cercano al verdadero valor de T si \mathcal{R}_0 está cerca de 1, como era de esperar. Con N que converge al infinito, sin embargo tenemos $T \simeq \frac{\ln N}{a-b}$. Encontramos el mismo comportamiento asintótico. Nuestro análisis basado en la fórmula (1) es válido, sin embargo, para todos $\mathcal{R}_0 > 1$ y no solo para \mathcal{R}_0 cerca de 1. Además, nuestra prueba es rigurosa.

Conclusión

Descubrimos que el pico de la epidemia ocurrió en un momento $T \sim (\ln N)/(a - b)$ si N converge al infinito. Este resultado parece nuevo a pesar de su simplicidad. En la literatura científica, solo hemos encontrado información sobre la altura del pico, que se determina fácilmente por la primera integral del sistema. De hecho, con

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

obtenemos

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

(ver por ejemplo [1, ecuación 7.26]). En algunas referencias como [7, p. 238], hay inversiones donde el tiempo t es una función de las variables monotónicas R o S , lo que lleva a integrales como (1) excepto por un cambio de variable. Pero el sistema diferencial SIR fue escrito allí en la forma $dS/dt = -\alpha SI$. La idea probablemente no vino de tomar $\alpha = a/N$, que de alguna manera es más realista y necesario para que el límite de una gran población produzca resultados interesantes.

Para una epidemia como la del coronavirus, el modelo SIR carece de un poco de realismo. Esto lleva a considerar modelos con más compartimentos, por ejemplo, con un compartimento adicional para personas en la fase latente que están infectadas pero aún no son contagiosas. Ver sobre este tema [5], que también trata de estimar los parámetros a partir de la curva epidémica. El estudio analítico del pico epidémico en este contexto es indudablemente más difícil.

Gracias

Agradecemos Hisashi Inaba por sus comentarios sobre el texto.

Referencias 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.crv.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.