

Der epidemische Höhepunkt in einem SIR-Modell

Quadrature 117 (2020)

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

Nicolas Bacaër

Institut de recherche pour le développement
Mathematische und Computermodellierungseinheit komplexer Systeme
Les Cordeliers, Paris, Frankreich
nicolas.bacaer@ird.fr

Zusammenfassung

Wir untersuchen eine Epidemie, die durch ein differenzielles SIR-System modelliert wurde. Wenn die Population N groß ist, ist der epidemische Höhepunkt zum Zeitpunkt T mit $T \sim (\ln N)/(a - b)$. a ist die effektive Kontaktrate. b ist die Geschwindigkeit, mit der sich infektiöse Menschen erholen.

Einführung

Betrachten wir wie zum Beispiel in [1, 2, 6, 7, 8] das SIR-Modell von Kermack und McKendrick für eine Epidemie

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

$S(t)$ ist die Anzahl der gesunden Menschen. $I(t)$ ist die Anzahl der infizierten Personen. $R(t)$ ist die Anzahl der Personen, die aufgrund von Isolation, Heilung oder Tod aus der Übertragung entfernt wurden. $N = S(t) + I(t) + R(t)$ ist die Gesamtbevölkerung. Es wird angenommen, dass diese Gesamtpopulation konstant und groß genug ist, um es sinnvoll zu machen, die Epidemie mithilfe eines Differenzialsystems und nicht anhand eines stochastischen Prozesses zu modellieren. Berücksichtigen Sie die Anfangsbedingungen

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

die dem Beginn der Epidemie von einem einzelnen "Null-Patienten" entsprechen.

Der Parameter a (mit $a > 0$) ist eigentlich das Produkt zweier Zahlen: die Anzahl der Kontakte pro Zeiteinheit und die Wahrscheinlichkeit der Übertragung während des Kontakts zwischen einer nicht infizierten Person und einer infizierten Person. Kontakte werden nach dem Zufallsprinzip hergestellt. I/N ist die Wahrscheinlichkeit, dass sich eine Person im Abteil I befindet. Dies erklärt den quadratischen Term in SI/N im Differenzialsystem. $b > 0$ ist die Rate, mit der Personen das Abteil I verlassen. Die zugrunde liegende Annahme ist daher, dass die in Kompartiment I verbrachte Zeit gemäß einem Exponentialgesetz von Parameter b verteilt wird. Wir gehen von $a > b$ für eine echte Epidemie aus.

Ein Beispiel ist in Abbildung 1 dargestellt. Wir haben die Scilab-Software und ihre Funktion zur Auflösung von Differenzialsystemen verwendet. Diese Funktion implementiert ein Adams-Prädiktor-Korrektor-Schema für nicht steile Systeme. Wir wählen $N = 65 \times 10^6$ (die Bevölkerung von Frankreich), $a = 1/2$ pro Tag und $b = 1/4$ pro Tag. Zu Beginn der Epidemie infiziert eine infizierte Person daher durchschnittlich alle zwei Tage eine Person. Durchschnittliche Infektionsdauer $1/b$ ist 4 Tage. Dies entspricht einer Reproduzierbarkeit $\mathcal{R}_0 = a/b = 2$ (Der Begriff und die Notation stammen aus [9, S. 102]). \mathcal{R}_0 ist daher die durchschnittliche Anzahl von Sekundärfällen, die eine infizierte Person zu Beginn der Epidemie infiziert. Ohne zu behaupten, realistisch zu sein, sind dies Werte in der gleichen Größenordnung wie die Werte für die Coronavirus-Epidemie 2020 in dem hypothetischen Fall, in dem keine Intervention zur Begrenzung der Epidemie erfolgt [5].

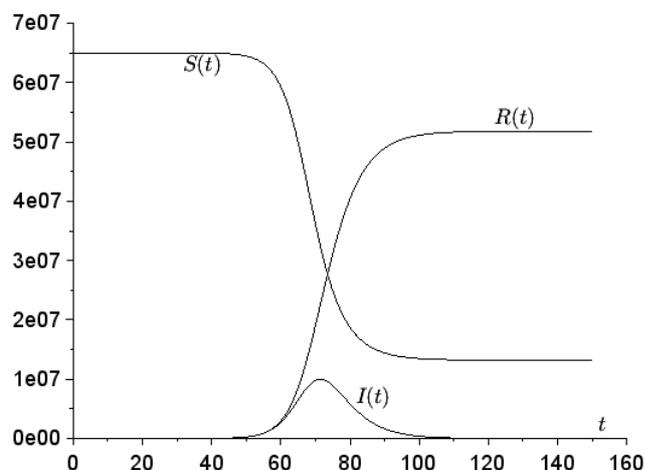


Abbildung 1. Eine Simulation des SIR-Modells mit der Zeit t in Tagen auf der Abszisse.

Der epidemische Höhepunkt

Durch die Kombination der ersten und dritten Gleichung des SIR-Systems ist es bekannt [8, S. 76] dass wir ein erstes Integral finden:

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

Also haben wir

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

Unter Bezugnahme auf die erste Gleichung finden wir

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

Der epidemische Peak in Abbildung 1 entspricht $dI/dt = 0$, was nach der zweiten Gleichung geschieht, wenn $S(t) = Nb/a$. $S(t)$ ist eine abnehmende Funktion, weil $dS/dt \leq 0$ aus der ersten Gleichung. Wir schließen daraus, dass wenn T die Zeit des epidemischen Peaks ist, d. H. Wenn

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

wir haben dann

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

Wir definieren $s = S/N$. Wir haben dann

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

Wir suchen eine obere Grenze von T . Für alles $b/a < s < 1 - 1/N$, wir haben

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

In der Nachbarschaft von $s = 1$, wir haben

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

Das Integral

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

ist divergent. Mit $N \rightarrow +\infty$, Also haben wir

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

Wir suchen eine untere Grenze von T . Wir haben $\ln s \leq s - 1$ für alles $s > 0$. Wir haben auch $\ln(1 - 1/N) \geq -2/N$ für alle $N \geq 2$. Also haben wir für $s \in]b/a, 1 - 1/N[$ und für jede ganze Zahl $N \geq 2$,

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

Also haben wir

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

Wir stoßen auf einen rationalen Bruchteil der Form $\frac{1}{s(u-vs)}$ mit

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

Es zerfällt in einfache Elemente

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

Ein Primitiv ist

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

Also haben wir

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

Das Mitglied auf der rechten Seite ist $\sim (\ln N)/(a - b)$ wenn $N \rightarrow +\infty$. Da es das gleiche Äquivalent wie das der oben erhaltenen Obergrenze ist, schließen wir daraus $T \sim (\ln N)/(a - b)$ wenn $N \rightarrow +\infty$.

Im Beispiel in Abbildung 1 haben wir $T \simeq 71,4$ Tage während $(\ln N)/(a - b) \simeq 72,0$ Tage. Die Annäherung scheint ziemlich gut zu sein. In Abbildung 2 variieren wir die Größe N der Population sowie den Parameter a mit b fest. Wir bemerken $(\ln N)/(a - b)$ manchmal überschätzt, manchmal unterschätzt T .

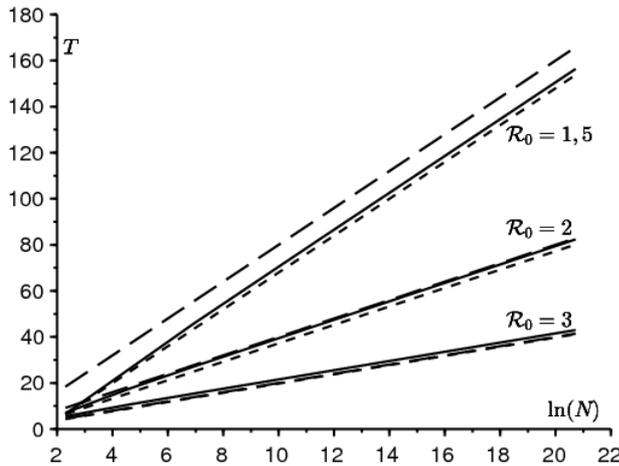


Abbildung 2. Vergleich zwischen dem Datum T des Epidemiepeaks (durchgezogene Linie), berechnet mit dem Integral der Formel (1), und dem Äquivalent $(\ln N)/(a - b)$ (große gepunktete Linien) und die Annäherung $[\ln(2N(b/a - 1)^2)]/(a - b)$ (kleine gepunktete Linien), abhängig von $\ln N$, wenn $10 \leq N \leq 10^8$ für drei Parameterwerte a mit Reproduzierbarkeit $\mathcal{R}_0 = a/b \in \{1,5; 2; 3\}$.

$a - b$ ist die anfängliche Wachstumsrate der Epidemie. Wir haben dann für N groß, $S(t) \simeq N$ und $dI/dt \simeq (a - b)I$. Diese Rate kann aus epidemiologischen Daten geschätzt werden. Bevölkerung N ist bekannt. Wir können daher das Datum des Höhepunkts der Epidemie vorhersagen. In der Praxis bleibt die dem Modell einer homogenen Bevölkerungsmischung zugrunde liegende Hypothese für ein ganzes Land zweifelhaft, für eine Stadt etwas weniger.

In der Vergangenheit hatten Kermack und McKendrick das SIR-System grob gelöst, indem sie angenommen hatten, dass die Reproduzierbarkeit nahe bei 1 lag, was zu einer kleinen Epidemie führte. Mit der Annäherung $e^x \simeq 1 + x + x^2/2$ Sie waren zu einer explizit lösaren Riccati-Gleichung gekommen, so dass

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

X , Y und Z sind komplizierte Funktionen der Parameter N , a und b . Siehe zum Beispiel [3]. Der epidemische Höhepunkt ist $T \simeq Z/Y$. Lassen Sie uns nebenbei bemerken, dass nach unserem Kenntnisstand die Annäherung (2) nie rigoros demonstriert wurde, weshalb wir über und unter dem informellen Zeichen verwenden \simeq . Wir gehen jetzt zusätzlich davon aus $1/N \ll (\mathcal{R}_0 - 1)^2$. In diesem Fall,

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

(siehe zum Beispiel [4]), so dass

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

Diese Annäherung ist auch in 2 gezeigt; es scheint besonders nahe am wahren Wert von T zu liegen, wenn \mathcal{R}_0 liegt nahe bei 1, wie Sie es erwarten würden. Mit N , das gegen unendlich konvergiert, haben wir dennoch $T \simeq \frac{\ln N}{a - b}$. Wir finden das gleiche asymptotische Verhalten. Unsere auf Formel (1) basierende Analyse gilt jedoch für alle $\mathcal{R}_0 > 1$ und nicht nur für \mathcal{R}_0 in der Nähe von 1. Darüber hinaus ist unser Beweis streng.

Fazit

Wir fanden heraus, dass der Höhepunkt der Epidemie zu einem bestimmten Zeitpunkt auftrat $T \sim (\ln N)/(a - b)$ wenn N gegen unendlich konvergiert. Dieses Ergebnis erscheint trotz seiner Einfachheit neu. In der wissenschaftlichen Literatur wurden nur Informationen über die Höhe des Peaks gefunden, die leicht durch das erste Integral des Systems bestimmt werden können. In der Tat mit

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

wir erhalten

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

(siehe zum Beispiel [1, Gleichung 7.26]). In einigen Referenzen wie [7, S. 238] gibt es Inversionen, bei denen die Zeit t eine Funktion der monotonen Variablen R oder S ist, was zu Integralen wie (1) führt, mit Ausnahme einer Änderung der Variablen. Aber das SIR-Differentialsystem wurde dort in der Form geschrieben $dS/dt = -\alpha SI$. Die Idee kam wahrscheinlich nicht von der Aufnahme $\alpha = a/N$, was irgendwie realistischer und notwendiger ist, damit die Grenze einer großen Bevölkerung interessante Ergebnisse liefert.

Für eine Epidemie wie die des Coronavirus fehlt dem SIR-Modell ein wenig Realismus. Dies führt dazu, dass Modelle mit mehr Kompartimenten in Betracht gezogen werden, beispielsweise mit einem zusätzlichen Kompartiment für Personen in der Latenzphase, die infiziert, aber noch nicht ansteckend sind. Siehe zu diesem Thema [5], das auch versucht, die Parameter aus der Epidemiekurve abzuschätzen. Die analytische Untersuchung des Epidemiepeaks in diesem Zusammenhang ist zweifellos schwieriger.

Vielen Dank

Wir danken Hisashi Inaba für seine Kommentare zum Text.

Literaturhinweise

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.