

SIR模型中的流行病高峰

Quadrature 117 (2020)

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

Nicolas Bacaër

Institut de recherche pour le développement

复杂系统的数学和计算机建模单元

Les Cordeliers, 法国巴黎

nicolas.bacaer@ird.fr

摘要

我们正在研究以差分SIR型系统为模型的流行病。当人口 N 大时，流行高峰在时间 T ， $T \sim (\ln N)/(a - b)$ 。 a 是有效的接触率。 b 是感染者恢复的速度。

介绍

例如，让我们考虑[1、2、6、7、8]的Kermack和McKendrick流行病的SIR模型

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

$S(t)$ 是健康人的人数。 $I(t)$ 是被感染的人数。 $R(t)$ 是指由于孤立，治愈或死亡而从传播中移出的人数。 $N = S(t) + I(t) + R(t)$ 是总人口。假定该总人口是恒定的且足够大，以使其有理由使用差分系统而非随机过程对流行病进行建模。考虑初始条件

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

对应于一个“零患者”的流行开始。

参数 a （与 $a > 0$ ）实际上是两个数字的乘积：每单位时间的接触数和未感染者与感染者之间接触期间传播的可能性。联系人是随机进行的。 I/N 是一个人在车厢 I 中的概率。这解释了 SI/N 在差速器系统中。 $b > 0$ 是人们离开车厢内的速度。因此，基本假设是，根据参数 b 的指数定律分配在隔室 I 中花费的时间。我们将假设 $a > b$ 表示实际流行。

图1中详细说明了一个示例。我们使用了Scilab软件及其差分系统解析功能。此功能为非陡峭系统实现了Adams预测器-校正器方案。我们选择 $N = 65 \times 10^6$ （法国人口）， $a = 1/2$ 每天和 $b = 1/4$ 每天。因此，在流行开始时，被感染的人平均每两天感染一个人。平均感染时间 $1/b$ 是4天。这对应于可再现性 $\mathcal{R}_0 = a/b = 2$ （该术语和表示法来自[9, p. 102]）。 \mathcal{R}_0 因此，是感染者在流行开始时感染的平均继发病例数。在没有声称是现实的情况下，这些值与在没有干预措施来限制流行病的假设情况下2020年冠状病毒流行病的发现数量级相同。]。

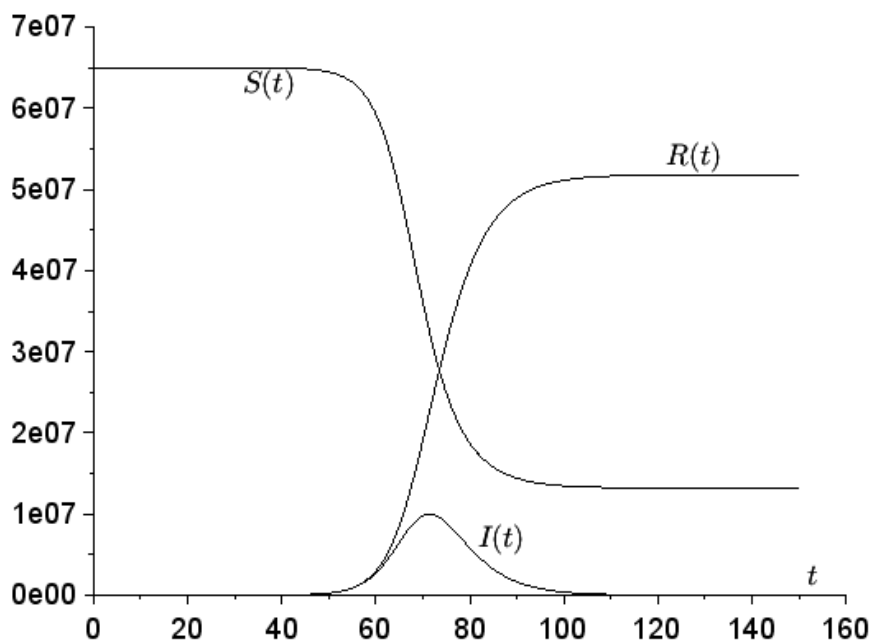


图1. SIR模型的模拟，横坐标为时间 t ，单位为天。

流行高峰

通过组合SIR系统的第一和第三方程，众所周知[8, p. 1]。76]我们找到了第一个积分：

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

所以我们有

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

回到第一个方程，我们发现

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

图1中的流行病高峰对应于 $dI/dt = 0$ ，这是根据第二个公式 $S(t) = Nb/a$ 。 $S(t)$ 是一个递减函数，因为 $dS/dt \leq 0$ 从第一个方程式我们推论如果 T 是流行高峰的时间，即

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

我们有

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

我们定义 $s = S/N$ 。然后我们有

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

我们寻求 T 的上限。为了一切 $b/a < s < 1 - 1/N$, 我们有

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

在附近 $s = 1$, 我们有

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

积分

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

是发散的。用 $N \rightarrow +\infty$, 所以我们有

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

我们寻求 T 的下限。我们有 $\ln s \leq s - 1$ 为了一切 $s > 0$ 。我们还有 $\ln(1 - 1/N) \geq -2/N$ 对所有 $N \geq 2$ 。所以我们有 $s \in]b/a, 1 - 1/N[$ 对于任何 $N \geq 2$ 的整数,

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

所以我们有

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

我们遇到了形式的合理部分 $\frac{1}{s(u-vs)}$ 与

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

它分为简单的元素

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

原语是

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

所以我们有

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

右边的成员是 $\sim (\ln N)/(a - b)$ 如果 $N \rightarrow +\infty$ 。由于它与上面获得的上限相同，因此我们得出以下结论： $T \sim (\ln N)/(a - b)$ 如果 $N \rightarrow +\infty$ 。

在图1的示例中，我们有 $T \simeq 71,4$ 几天 $(\ln N)/(a - b) \simeq 72,0$ 天。近似值似乎很好。在图2中，我们改变了总体 N 的大小以及参数 a ，用 b 固定。我们注意到 $(\ln N)/(a - b)$ 有时高估，有时低估 T 。

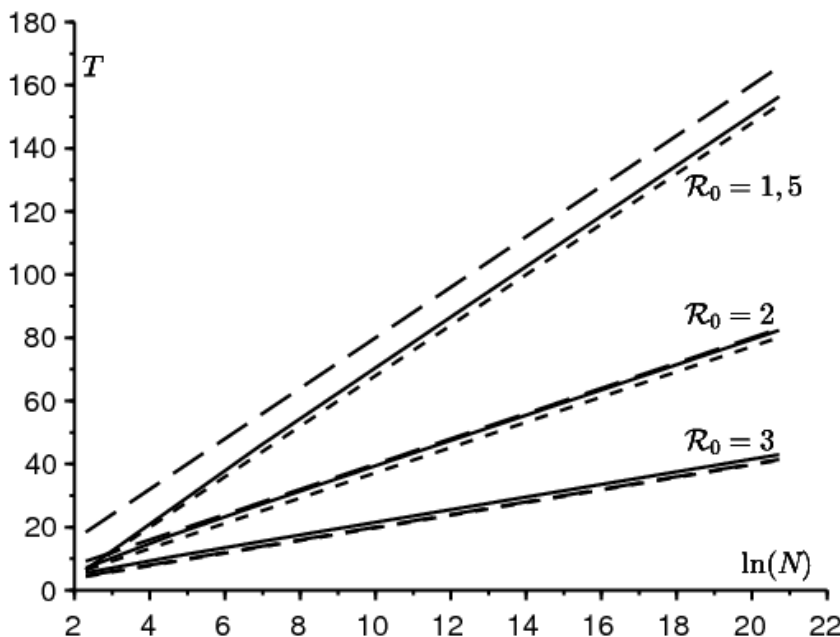


图2.一方面，用公式 (1) 的积分计算出的流行病高峰日期 T (实线) 之间的比较，另一方面 $(\ln N)/(a - b)$ (大虚线) 和近似值 $[\ln(2N(b/a - 1)^2)]/(a - b)$ (小虚线)，具体取决于 $\ln N$ 如果 $10 \leq N \leq 10^8$ 对于三个参数值 a 具有再现性 $R_0 = a/b \in \{1,5; 2; 3\}$ 。

$a - b$ 是该流行病的初始增长率。然后我们有 N 大， $S(t) \simeq N$ 和 $dI/dt \simeq (a - b)I$ 。可以从流行病学数据估计该比率。人口 N 是已知的。因此，我们可以预测流行高峰的日期。实际上，对于整个国家而言，人口同质混合模型所依据的假设仍然值得怀疑，而对于一个城市来说，这一假设就更少了。

从历史上看，Kermack和McKendrick通过假定可重复性接近1来大致解决了SIR系统，这导致了小规模的流行。与近似 $e^x \simeq 1 + x + x^2/2$ ，他们得出了一个明确可解的Riccati方程，因此

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

X , Y 和 Z 是参数 N , a 和 b 的复杂函数。参见例如[3]。流行高峰是 $T \simeq Z/Y$ 。请注意, 据我们所知, 从未严格证明过近似值 (2), 这就是为什么我们在上下使用非正式符号的原因 \simeq 。现在我们假设 $1/N \ll (\mathcal{R}_0 - 1)^2$ 。在这种情况下,

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

(例如参见[4]), 这样

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

图2也显示了这种近似值。它似乎特别接近的真正价值牛逼如果 \mathcal{R}_0 如您所料, 接近1。随着 N 收敛到无穷大, 我们仍然有 $T \simeq \frac{\ln N}{a-b}$ 。我们发现相同的渐近行为。我们基于公式 (1) 的分析对所有 $\mathcal{R}_0 > 1$ 不仅是 \mathcal{R}_0 接近于1。此外, 我们的证明是严格的。

结论

我们发现该流行病的高峰一次发生 $T \sim (\ln N)/(a-b)$ 如果 N 收敛到无穷大。尽管简单, 但此结果似乎是新的。在科学文献中, 我们仅找到有关峰高的信息, 该信息很容易由系统的第一个积分确定。确实, 随着

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

我们获得

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

(例如参见[1, 公式7.26])。在某些参考文献中, 例如[7, p. 238], 存在一些反演, 其中时间 t 是单调变量 R 或 S 的函数, 除了变量的变化外, 它会导致像 (1) 一样的积分。但是SIR微分系统以以下形式写在那里 $dS/dt = -\alpha SI$ 。这个想法可能不是来自采取 $\alpha = a/N$, 这对于较大数量的人口边界产生有趣的结果在某种程度上更加现实和必要。

对于像冠状病毒这样的流行病, SIR模型缺乏真实感。这导致考虑使用具有更多隔室的模型, 例如, 为处于潜伏期但已感染但尚未传染的人提供额外的隔室。参见该主题[5], 该主题还尝试根据流行曲线估算参数。在这种情况下, 对流行病高峰的分析研究无疑更加困难。

谢谢

我们感谢 Hisashi Inaba 他对文本的评论。

书目参考

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.