

De epidemische piek in een SIR-model

Quadrature 117 (2020)

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

Nicolas Bacaër

Institut de recherche pour le développement
Wiskundige en computermodellering van complexe systemen
Les Cordeliers, Parijs, Frankrijk
nicolas.bacaer@ird.fr

Overzicht

We bestuderen een epidemie gemodelleerd door een differentieel SIR-type systeem. Wanneer populatie N groot is, is de epidemische piek op tijdstip T , met $T \sim (\ln N)/(a - b)$. a is het effectieve contactpercentage. b is de snelheid waarmee besmettelijke mensen herstellen.

Invoering

Laten we bijvoorbeeld eens kijken naar [1, 2, 6, 7, 8] het SIR-model van Kermack en McKendrick voor een epidemie

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

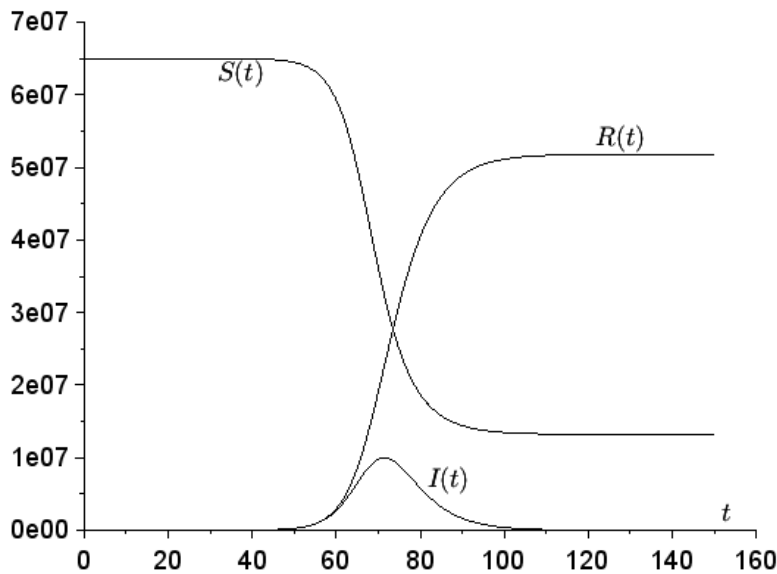
$S(t)$ is het aantal gezonde mensen. $I(t)$ is het aantal geïnfecteerde mensen. $R(t)$ is het aantal mensen dat uit de overdracht is verwijderd als gevolg van isolatie, genezing of overlijden. $N = S(t) + I(t) + R(t)$ is de totale bevolking. Aangenomen wordt dat deze totale populatie constant en groot genoeg is om het redelijk te maken de epidemie te modelleren met behulp van een differentieel systeem in plaats van een stochastisch proces. Overweeg de beginvoorwaarden

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

die overeenkomen met het begin van de epidemie van een enkele "nul patiënt".

De parameter a (met $a > 0$) is eigenlijk het product van twee cijfers: het aantal contacten per tijdseenheid en de kans op overdracht tijdens contact tussen een niet-geïnfecteerde persoon en een geïnfecteerde persoon. Contacten worden willekeurig gemaakt. I/N is de kans dat een persoon in compartiment I zit. Dit verklaart de kwadratische term in SI/N in het differentieel systeem. $b > 0$ is de snelheid waarmee mensen het compartiment I verlaten. De onderliggende aanname is daarom dat de tijd doorgebracht in compartiment I wordt verdeeld volgens een exponentiële wet van parameter b . We gaan uit van $a > b$ om een echte epidemie te laten plaatsvinden.

Een voorbeeld is gedetailleerd in figuur 1. We gebruikten de Scilab-software en de functie voor differentieële systeemresolutie. Deze functie implementeert een Adams predictor-corrector-schema voor niet-steile systemen. We kiezen $N = 65 \times 10^6$ (de bevolking van Frankrijk), $a = 1/2$ per dag en $b = 1/4$ per dag. Aan het begin van de epidemie infecteert een besmette persoon daarom gemiddeld elke twee dagen één persoon. Gemiddelde duur van infectie $1/b$ is 4 dagen. Dit komt overeen met een reproduceerbaarheid $\mathcal{R}_0 = a/b = 2$ (de term en de notatie komen uit [9, p. 102]). \mathcal{R}_0 is dus het gemiddelde aantal secundaire gevallen dat een geïnfecteerde persoon aan het begin van de epidemie besmet. Zonder te beweren realistisch te zijn, zijn dit waarden van dezelfde orde van grootte als wat kan worden gevonden voor de coronavirusepidemie in 2020 in het hypothetische geval waarin er niet wordt ingegrepen om de epidemie te beperken [5].



Figuur 1. Een simulatie van het SIR-model met de tijd t in dagen op de abscis .

De epidemische piek

Door de eerste en derde vergelijking van het SIR-systeem te combineren, is het bekend [8, p. 76] dat we een eerste integraal vinden:

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

Dus we hebben

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

Terugverwijzend naar de eerste vergelijking vinden we

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

De epidemische piek in figuur 1 komt overeen met $dI/dt = 0$, wat gebeurt volgens de tweede vergelijking wanneer $S(t) = Nb/a$. $S(t)$ is een afnemende functie omdat $dS/dt \leq 0$ uit de eerste vergelijking. We leiden hieruit af dat als T de tijd van de epidemische piek is, d.w.z. als

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

we hebben dan

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

We definiëren $s = S/N$. We hebben dan

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

Wij zoeken een bovengrens van T . Voor alles $b/a < s < 1 - 1/N$, men heeft

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

In de buurt van $s = 1$, men heeft

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

De integraal

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

is verschillend. Met $N \rightarrow +\infty$, Dus we hebben

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

Wij streven naar een ondergrens van T . Men heeft $\ln s \leq s-1$ voor alles $s > 0$. We hebben ook $\ln(1-1/N) \geq -2/N$ voor iedereen $N \geq 2$. Dus we hebben voor $s \in]b/a, 1-1/N[$ en voor elk geheel getal $N \geq 2$,

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

Dus we hebben

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

We komen een rationeel deel van het formulier tegen $\frac{1}{s(u-vs)}$ met

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

Het valt uiteen in eenvoudige elementen

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

Een primitief is

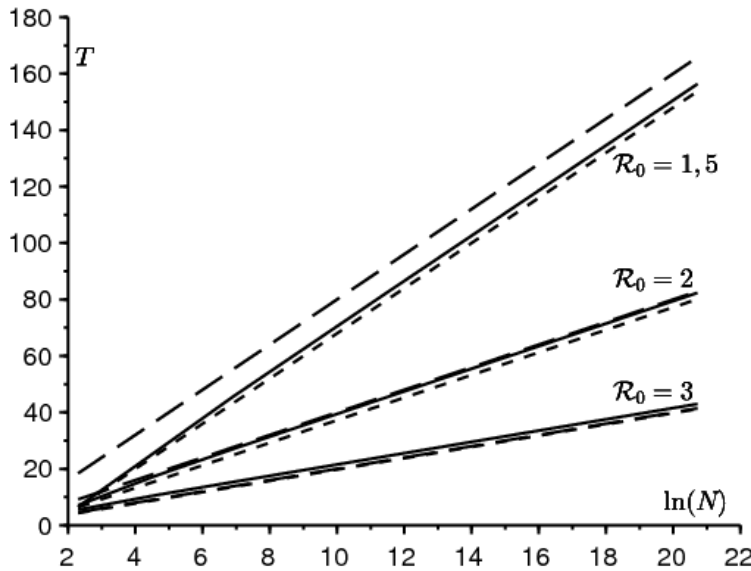
$$\frac{1}{u} [\ln s - \ln(u-vs)] = \frac{1}{u} \ln \frac{s}{u-vs}.$$

Dus we hebben

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

Het lid aan de rechterkant is $\sim (\ln N)/(a-b)$ als $N \rightarrow +\infty$. Aangezien het hetzelfde equivalent is als dat van de hierboven verkregen bovengrens, concluderen we dat $T \sim (\ln N)/(a-b)$ als $N \rightarrow +\infty$.

In het voorbeeld in figuur 1 hebben we $T \simeq 71,4$ dagen terwijl $(\ln N)/(a-b) \simeq 72,0$ dagen. De benadering lijkt redelijk goed. In Figuur 2 variëren we de grootte N van de populatie en de parameter a , met b vast. We merken $(\ln N)/(a-b)$ soms overschat, soms onderschat T .



Figuur 2. Vergelijking tussen enerzijds de datum T van de piek van de epidemie (ononderbroken lijn), berekend met de integraal van de formule (1), anderzijds het equivalent $(\ln N)/(a - b)$ (grote stippellijnen) en de benadering $[\ln(2N(b/a - 1)^2)]/(a - b)$ (kleine stippellijnen), afhankelijk van $\ln N$, als $10 \leq N \leq 10^8$ voor drie parameterwaarden a met reproduceerbaarheid $\mathcal{R}_0 = a/b \in \{1,5 ; 2 ; 3\}$.

$a - b$ is de initiële groeisnelheid van de epidemie. We hebben dan voor N groot, $S(t) \simeq N$ en $dI/dt \simeq (a - b)I$. Dit percentage kan worden geschat op basis van epidemiologische gegevens. Populatie N is bekend. We kunnen dus de datum van de piek van de epidemie voorspellen. In de praktijk blijft de hypothese die ten grondslag ligt aan het model van een homogene mix van de bevolking twijfelachtig voor een heel land, iets minder voor een stad.

Historisch gezien hadden Kermack en McKendrick het SIR-systeem ruwweg opgelost door aan te nemen dat de reproduceerbaarheid dicht bij 1 lag, wat leidde tot een kleine epidemie. Met de benadering $e^x \simeq 1 + x + x^2/2$, waren ze tot een expliciet oplosbare Riccati-vergelijking gekomen, dus dat

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

X , Y en Z zijn ingewikkelde functies van de parameters N , a en b . Zie bijvoorbeeld [3]. De epidemische piek is $T \simeq Z/Y$. Merk op dat voor zover wij weten de benadering (2) nooit rigoureus is aangetoond, daarom gebruiken we het informele teken hierboven en hieronder \simeq . We gaan er nu van uit $1/N \ll (\mathcal{R}_0 - 1)^2$. In dat geval,

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

(zie bijvoorbeeld [4]), zodat

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

Deze benadering wordt ook getoond in Figuur 2; het lijkt bijzonder dicht bij de werkelijke waarde van T als \mathcal{R}_0 ligt dicht bij 1, zoals je zou verwachten. Met N die naar het oneindige convergeert, hebben we toch $T \simeq \frac{\ln N}{a - b}$. We vinden hetzelfde asymptotische gedrag. Onze analyse op basis van formule (1) is echter geldig voor iedereen $\mathcal{R}_0 > 1$ en niet alleen voor \mathcal{R}_0 bijna 1. Bovendien is ons bewijs rigoureus.

Conclusie

We ontdekten dat de piek van de epidemie zich tegelijkertijd voordeed $T \sim (\ln N)/(a - b)$ als N convergeert naar oneindig. Dit resultaat lijkt ondanks zijn eenvoud nieuw. In de wetenschappelijke literatuur hebben we alleen informatie gevonden over de hoogte van de piek, die gemakkelijk bepaald kan worden door de eerste integraal van het systeem. Inderdaad met

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

we verkrijgen

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

(zie bijvoorbeeld [1, vergelijking 7.26]). In sommige referenties zoals [7, p. 238], zijn er inversies waarbij tijd t een functie is van de monotone variabelen R of S , wat leidt tot integralen zoals (1) behalve een verandering van variabele. Maar het SIR-differentiaalsysteem was daar geschreven in de vorm $dS/dt = -\alpha SI$. Het idee kwam waarschijnlijk niet van het nemen $\alpha = a/N$, wat op de een of andere manier realistischer en noodzakelijk is voor de grens van een grote populatie om interessante resultaten op te leveren.

Voor een epidemie als die van het coronavirus mist het SIR-model een beetje realisme. Dit leidt tot het overwegen van modellen met meer compartimenten, bijvoorbeeld met een extra compartiment voor mensen in de latente fase die besmet zijn maar nog niet besmettelijk. Zie hierover [5], dat ook probeert de parameters van de epidemische curve te schatten. De analytische studie van de epidemiepiek in deze context is ongetwijfeld moeilijker.

Bedankt

Wij danken Hisashi Inaba voor zijn opmerkingen over de tekst.

Bibliografische referenties

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.