

# Un modello matematico dell'epidemia di coronavirus in Francia

Nicolas Bacaër

Institut de Recherche pour le Développement  
Unità di modellazione matematica e computerizzata di sistemi complessi  
Les Cordeliers, Parigi, Francia  
nicolas.bacaer@ird.fr

<https://hal.archives-ouvertes.fr/hal-02509142>

## riassunto

Adattiamo un modello epidemico di tipo SEIR ai dati epidemici del nuovo coronavirus in Francia nel 2020. Se i contatti vengono ridotti a zero da una certa data vicino all'inizio dell'epidemia, le dimensioni finali dell'epidemia sono vicine di quello ottenuto moltiplicando il numero cumulativo di casi in questa data per la riproducibilità  $\mathcal{R}_0$  dell'epidemia. Con le nostre stime,  $\mathcal{R}_0 \simeq 2,3$  in Francia.

## Il modello

Figura 1<sup>a</sup> mostra il numero cumulativo di casi rilevati di coronavirus in Francia tra il 25 febbraio e il 15 marzo 2020. Tra queste due date, il numero cumulativo è passato da 13 a 5427. Osserviamo una crescita che sembra esponenziale. Figura 1<sup>b</sup> mostra come una linea di regressione si regola nelle coordinate logaritmiche. Scopriamo che la curva aumenta come  $e^{\lambda t}$  con un tasso di crescita  $\lambda \simeq 0,31$  al giorno (rosso giusto). Il tempo del raddoppio è  $\log(2)/\lambda \simeq 2,2$  giorni. Se, d'altra parte, ci limitiamo agli ultimi 9 giorni, che sono particolarmente ben allineati su una scala logaritmica, otteniamo  $\lambda \simeq 0,225$  al giorno e un tempo di raddoppio di 3,1 giorni (blu a destra). Poiché i dati all'inizio dell'epidemia sono disturbati da una grande percentuale di nuovi casi importati, è probabilmente la seconda stima a essere più affidabile; questo è quello che useremo di seguito.

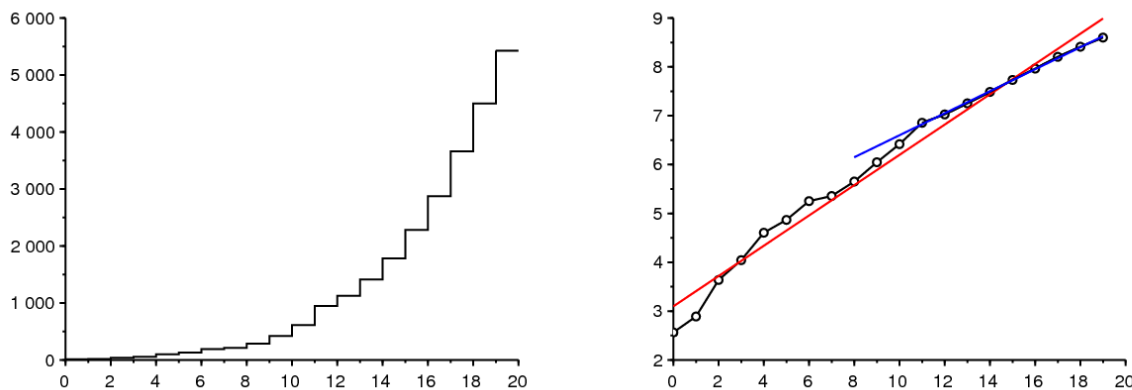


Figura 1. a) Numero cumulativo di casi rilevati in Francia tra il 25 febbraio e il 15 marzo 2020, secondo Santé publique France. b) Logaritmo naturale del numero cumulativo di casi e linee di regressione lineare.

Proponiamo un modello matematico per questa epidemia. Dividiamo la popolazione francese in quattro scomparti secondo il classico modello SEIR (vedi ad esempio [2, p. 61]): probabile che sia infetto ( $S$ ), infetto ma non infettivo ( $E$ ), infettivo senza protezione ( $I$ ) e non è probabile che trasmetta l'infezione (cioè contagiosa ma confinata, guarita o deceduta,  $R$ ). Possiamo ovviamente affinare questo modello all'infinito per renderlo più realistico (vedi ad esempio tra molti altri riferimenti [3,7]) ma abbiamo cercato di limitare il numero di parametri sconosciuti il più possibile; miriamo anche a ottenere un risultato più teorico che pratico.

nota  $N$  la popolazione totale,  $a$  il tasso di contatto effettivo,  $b$  la velocità con cui le persone infette diventano infettive e  $c$  la velocità con cui le persone infette sono isolate:

$$\frac{dS}{dt} = -a S \frac{I}{N} \quad (1)$$

$$\frac{dE}{dt} = a S \frac{I}{N} - b E \quad (2)$$

$$\frac{dI}{dt} = b E - c I \quad (3)$$

$$\frac{dR}{dt} = c I, \quad (4)$$

con  $N = S(t) + E(t) + I(t) + R(t)$ . Per creare il collegamento con i dati, possiamo pensarci  $R(t)$  corrisponde al numero cumulativo di casi noti al momento  $t$ .

All'inizio dell'epidemia, il numero di casi rimane molto piccolo rispetto alla popolazione totale, quindi  $S(t) \simeq N$ , che porta alla linearizzazione

$$\frac{dE}{dt} \simeq a I - b E, \quad \frac{dI}{dt} \simeq b E - c I.$$

L'epidemia quindi tende a crescere esponenzialmente come  $e^{\lambda t}$ , dove  $\lambda$  è il più grande autovalore della matrice

$$\begin{pmatrix} -b & a \\ b & -c \end{pmatrix}. \quad (5)$$

Il polinomio caratteristico è

$$\lambda^2 + (b+c)\lambda + b(c-a) = 0. \quad (6)$$

pertanto

$$\lambda = \frac{-(b+c) + \sqrt{(b+c)^2 - 4b(c-a)}}{2} = \frac{-(b+c) + \sqrt{(b-c)^2 + 4ab}}{2}. \quad (7)$$

Durata media nel compartimento  $E$ , che vale  $1/be$  che assimiliamo al periodo di incubazione, sarebbe dell'ordine di 4 giorni [4]. Quindi prendi  $b = 0,25$  al giorno. Durata media nel compartimento  $I$  prima dell'isolamento, che vale la pena  $1/c$ , è più difficile da stimare perché dipende non solo dalle caratteristiche biologiche del virus, ma anche dalla tempestività con cui i casi vengono isolati, che varia da paese a paese. Supponiamo che sia trascorso circa 1 giorno dall'epidemia in Francia, quando gli abitanti sono già ben consapevoli dell'esistenza della pandemia; i pazienti non tardano ad essere isolati. In tal modo  $c = 1$  al giorno. Ne deduciamo dalla formula (7) che

$$a = \frac{(2\lambda + b + c)^2 - (b - c)^2}{4b} = (\lambda + c) \left(1 + \frac{\lambda}{b}\right). \quad (8)$$

che consentirebbe di calcolare numericamente la velocità di contatto effettiva  $a$  dal tasso di crescita osservato  $\lambda$ .

Immagina che le misure di sanità pubblica potrebbero dividere il tasso di contatto effettivo per un numero  $k$  che è maggiore di 1. Quanto deve valere almeno  $k$  per fermare l'epidemia? Questo valore di  $k$ , notato tradizionalmente  $\mathcal{R}_0$  seguendo Lotka e da lui chiamato "riproducibilità" [8, p. 102], si ottiene semplicemente osservando che quando  $a$  è sostituito da  $a' = a/\mathcal{R}_0$ , il nuovo tasso di crescita dell'epidemia  $\lambda'$  deve essere zero, che secondo l'equazione (6) porta a  $c - a/\mathcal{R}_0 = 0$  e a

$$\mathcal{R}_0 = \frac{a}{c} = \left(1 + \frac{\lambda}{b}\right) \left(1 + \frac{\lambda}{c}\right) \simeq 2,33$$

se utilizziamo il valore numerico ( $\lambda \simeq 0,225$  al giorno) suggerito dalla curva epidemica della figura 1. Visualizza le incertezze sui parametri  $b$  e  $c$ , questo può essere solo un valore approssimativo.

In un modo più tecnico (vedi ad esempio [9]), si sarebbe potuto notare  $\mathcal{R}_0$  era anche il raggio spettrale della matrice

$$\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b & 0 \\ -b & c \end{pmatrix}^{-1}.$$

Torniamo al modello SEIR (1) - (4). Ricordiamo come determinare la dimensione finale dell'epidemia in completa assenza di intervento; è un facile adattamento del metodo usato per il modello SIR (vedi ad esempio [6, p. 76]). L'equazione (1) lo mostra

$$\frac{d}{dt} \log S = -\frac{a}{N} I(t),$$

così integrando  $t = 0$  a  $t = +\infty$ :

$$\log S(\infty) - \log S(0) = -\frac{a}{N} \int_0^{\infty} I(t) dt. \quad (9)$$

All'inizio dell'epidemia, nessuno era ancora nel compartimento  $R$ , pertanto  $R(0) = 0$ . L'equazione (4) lo mostra

$$R(\infty) = c \int_0^{\infty} I(t) dt, \quad (10)$$

dove  $R(\infty)$  indica il limite quando  $t \rightarrow +\infty$  della funzione  $R(t)$ , che aumenta e aumenta di  $N$ . Inoltre, abbiamo sempre  $S(t) + E(t) + I(t) + R(t) = N$ . quando  $t \rightarrow +\infty$ , l'epidemia alla fine si ferma così  $E(t)$  e  $I(t)$  tendono verso 0. Al limite, quindi, solo le persone che sono fuggite dall'epidemia e quelle che sono state infettate ma che sono passate nel compartimento  $R$ :

$$S(\infty) + R(\infty) = N. \quad (11)$$

Combinando (9), (10) e (11), vediamo che

$$N - R(\infty) = S(0) \exp\left(-\frac{a}{c} \frac{R(\infty)}{N}\right).$$

All'inizio dell'epidemia, c'erano solo poche persone infette nella popolazione, quindi  $S(0) \simeq N$ . L'equazione per la dimensione finale dell'epidemia può essere scritta come

$$1 - \frac{R(\infty)}{N} \simeq \exp\left(-\mathcal{R}_0 \frac{R(\infty)}{N}\right),$$

che ha la stessa forma del modello SIR [6]. con  $\mathcal{R}_0 \simeq 2,33$ , troviamo digitalmente  $R(\infty)/N \simeq 87\%$ .

## Seconda fase con intervento drastico

Immagina che ad una certa data  $\tau$ , vengono prese misure drastiche in modo che la nuova velocità di contatto effettiva sia ridotta a 0 mentre è presente  $R(\tau)$  casi cumulativi. Ad esempio, ci sono stati 5.427 casi cumulativi a partire dal 15 marzo, data in cui sono entrate in vigore le misure relative a scuole e luoghi pubblici [5]. Possiamo quindi prevedere quale sarà la nuova dimensione finale dell'epidemia?

Verso la fine della fase esponenziale dove  $t \leq \tau$  e dove il numero totale di casi rappresenta ancora una piccola parte della popolazione totale, abbiamo

$$E(t) \simeq u e^{\lambda t}, \quad I(t) \simeq v e^{\lambda t}, \quad R(t) \simeq w e^{\lambda t},$$

dove  $(u, v)$  è un autovettore associato al più grande autovalore  $\lambda$  della matrice (5). In tal modo,  $-bu + av = \lambda u$ . Con l'equazione (8), lo troviamo

$$u = \frac{av}{\lambda + b} = \frac{\lambda + c}{b} v.$$

come  $dR/dt \simeq \lambda R$  per  $t < \tau$  quando  $t$  non è troppo vicino a 0, abbiamo

$$I(\tau) = \frac{1}{c} \frac{dR}{dt}(\tau) \simeq \frac{\lambda}{c} R(\tau).$$

ma  $I(\tau) \simeq v e^{\lambda \tau}$ , pertanto

$$E(\tau) \simeq u e^{\lambda \tau} = \frac{\lambda + c}{b} v e^{\lambda \tau} \simeq \frac{\lambda + c}{b} I(\tau) \simeq \left( \frac{\lambda^2}{bc} + \frac{\lambda}{b} \right) R(\tau).$$

Abbiamo dovuto ridurre i contatti a zero  $t > \tau$

$$\frac{dS}{dt} = 0 \tag{12}$$

$$\frac{dE}{dt} = -bE \tag{13}$$

$$\frac{dI}{dt} = bE - cI \tag{14}$$

$$\frac{dR}{dt} = cI. \tag{15}$$

Senza dover risolvere questo sistema, è chiaro che sarà la dimensione finale dell'epidemia  $R(\infty) = R(\tau) + E(\tau) + I(\tau)$ , poiché c'è  $E(\tau) + I(\tau)$  individui infetti che non sono ancora nel compartimento  $R$  al momento  $\tau$ . In tal modo

$$R(\infty) \simeq R(\tau) \left( 1 + \frac{\lambda^2}{bc} + \frac{\lambda}{b} + \frac{\lambda}{c} \right) = R(\tau) \left( 1 + \frac{\lambda}{b} \right) \left( 1 + \frac{\lambda}{c} \right) = \mathcal{R}_0 R(\tau).$$

Pertanto, se i contatti si riducono a zero da una certa data vicino all'inizio dell'epidemia (abbastanza vicini da rendere valida l'approssimazione lineare ma non troppo vicini perché il sistema linearizzato abbia avuto il tempo di convergere verso l'autovettore associato al primo autovalore), quindi la dimensione finale dell'epidemia è vicina a quella ottenuta moltiplicando il numero cumulativo di casi in questa data per la riproducibilità  $\mathcal{R}_0$  dell'epidemia. Un risultato simile si ottiene facilmente allo stesso modo per un modello SIR. Ma il grado di generalità di questa osservazione resta da determinare.

Notiamo nel passare l'analogia con il concetto di "potenziale di crescita" delle popolazioni in demografia [10, p. 176]. È il rapporto tra la popolazione stazionaria finale e la popolazione in un determinato istante se la fertilità è improvvisamente divisa in questo istante dalla riproducibilità  $\mathcal{R}_0$ , in modo che la popolazione rimanga con un tasso asintotico di crescita zero. Come nel nostro calcolo, è supponendo che la popolazione iniziale sia "stabile" nel senso di Lotka (vale a dire dato dal primo autovettore) che Keyfitz ha ottenuto una formula relativamente semplice per il potenziale di aumentare, una formula che coinvolge anche  $\mathcal{R}_0$  sebbene in modo più complicato rispetto al nostro modello SEIR [10, p. 179].

Si noti inoltre che la stima di  $R(\tau) + E(\tau) + I(\tau)$  dai dati  $R(\tau)$  solo uno è analogo al problema sorto all'inizio dell'epidemia di HIV per stimare il numero di persone sieropositive dal numero di casi di AIDS segnalati.

con  $R(\tau) = 5427$  e  $\mathcal{R}_0 \simeq 2,33$ , dà  $R(\infty) \simeq 12\,600$ . Ancora una volta, evidenzia l'incertezza attorno ai parametri  $b$  e  $c$ , che si riflette nel valore di  $R(\infty)$ .

Figura 2<sup>a</sup> illustra questo modello a due fasi. Abbiamo preso  $N = 65 \times 10^6$ ,  $S(0) = N - 1$ ,  $E(0) = 1$ ,  $I(0) = 0$  e  $R(0) = 0$ . Il parametro  $a$  è dato dalla formula (8) con  $\lambda = 0,225$  al giorno. Abbiamo preso  $\tau = 40,1$  giorni così  $R(\tau) \simeq 5415$  è molto vicino ai dati 5427 del 15 marzo. Continuando la simulazione un po' più a lungo rispetto alla figura, troviamo numericamente quello  $R(\infty)/R(\tau) \simeq 2,33 \simeq \mathcal{R}_0$ .

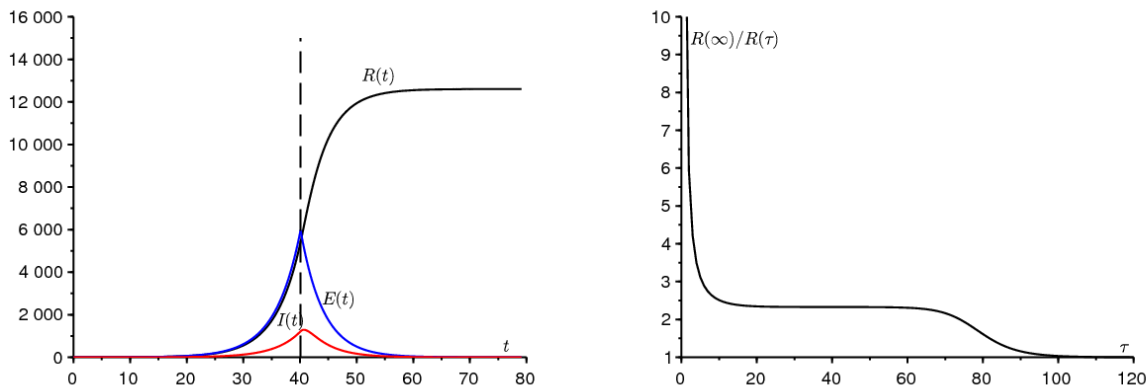


Figura 2. a) Esempio di simulazione del modello bifase. b) Il rapporto  $R(\infty)/R(\tau)$  a seconda di  $\tau$ .

Figura 2<sup>b</sup> mostra come il rapporto  $R(\infty)/R(\tau)$  varia a seconda del tempo  $\tau$  dove la velocità di contatto è ridotta a zero. Possiamo osservare chiaramente un plateau in cui questo rapporto è vicino  $\mathcal{R}_0$ . quando  $\tau \rightarrow 0$ , abbiamo  $R(\tau) \rightarrow 0$  e  $R(\infty) \rightarrow E(0) + I(0) = 1$ , quindi il rapporto  $R(\infty)/R(\tau)$  tende all'infinito. Il rapporto si sta avvicinando  $\mathcal{R}_0$  quando  $\tau$  è dell'ordine dell'inverso della differenza tra i due autovalori della matrice (5). Quando, al contrario  $\tau \rightarrow \infty$ , quindi l'intervento arriva troppo tardi, l'epidemia è già passata e  $R(\infty)/R(\tau) \rightarrow 1$ . Si prevede che la larghezza della mensola dove  $R(\infty)/R(\tau)$  è vicino a  $\mathcal{R}_0$  cresce come una costante moltiplicata per  $\log(N)$  quando  $N \rightarrow +\infty$ , poiché questo è ad esempio il comportamento del tempo fino al picco epidemico in un modello SIR con coefficienti costanti.

In realtà, la velocità di contatto effettiva non sarà certamente del tutto nulla per  $t > \tau$ . Il valore ottenuto per  $R(\infty)$  può comunque essere considerato come un limite inferiore del valore reale poiché è certo che la dimensione finale dell'epidemia sarà maggiore con contatti diversi da zero rispetto a contatti zero per  $t > \tau$ . A questo proposito, va ricordato che i modelli epidemici di tipo SIR o SEIR con una velocità di contatto variabile non sono "monotoni", nel senso che una riduzione della frequenza di contatto a volte può portare a una dimensione finale più grande dell'epidemia. [1].

Se vogliamo sapere quando si verificherà il picco dell'arrivo di nuovi casi, che può essere importante per i problemi di congestione negli ospedali, dobbiamo risolvere il sistema (12) - (15). Noi troviamo  $E(t) = E(\tau) e^{-b(t-\tau)}$  e

$$\begin{aligned} I(t) &= I(\tau) e^{-c(t-\tau)} + b \int_{\tau}^t e^{-c(t-s)} E(s) ds \\ &= I(\tau) e^{-c(t-\tau)} + b E(\tau) \frac{e^{-b(t-\tau)} - e^{-c(t-\tau)}}{c-b}. \end{aligned}$$

Con i valori numerici di figura 2<sup>a</sup>, abbiamo  $E(\tau) \simeq 5970$ ,  $I(\tau) \simeq 1218$  e il derivato di  $I(t)$  annulla 0,6 giorni dopo che la velocità di contatto è stata ridotta a 0. Per una velocità di contatto diversa da zero, il ritardo sarebbe ovviamente più lungo.

In conclusione, abbiamo esplorato lo scenario più favorevole, quello in cui la velocità di contatto è ridotta a zero da una certa data. Abbiamo trovato una semplice formula approssimativa per la dimensione finale dell'epidemia in base al numero di casi rilevati al momento della riduzione. Tuttavia, resta da dichiarare e dimostrare questo risultato in modo più rigoroso, senza dubbio aumentando la popolazione totale  $N$  a  $+\infty$ .

**grazie**

Ringraziamo Hisashi Inaba e Ali Moussaoui per i loro commenti sul manoscritto.

**riferimenti**

1. Bacaër N., Gomes M. G. M. (2009) On the final size of epidemics with seasonality. *Bull. Math. Biol.* 71 : 1954 – 1966.
2. Corlosquet – Habart M., Janssen J., Manca R. (2012) Modélisation stochastique du risque de pandémie : stratégies de couverture et d'assurance. Lavoisier, Cachan.
3. Danchin A., Ng P. T. W., Turinici G. (2020) A new transmission route for the propagation of the SARS – CoV – 2 coronavirus. <https://doi.org/10.1101/2020.02.14.20022939>
4. Guan W. J. et coll. (2020) Clinical characteristics of 2019 novel coronavirus infection in China. *New England Journal of Medicine*, doi : 10.1056/NEJMoa2002032
5. Heckeitsweiler C., Pietralunga C. (2020) Coronavirus : les simulations alarmantes des épidémiologistes pour la France. *Le Monde*, publié en ligne le 15/03/2020.
6. A. Hillion (1986) *Les Théories mathématiques des populations*. Presses Universitaires de France, Paris.
7. Lin Q. Y., Zhao S., Gao D. Z., et coll. (2020) A conceptual model for the coronavirus disease 2019 (COVID – 19) outbreak in Wuhan, China with individual reaction and governmental action. *Int. J. Infect. Dis.* 93 : 211 – 216.
8. Lotka A. J. (1939) *Théorie analytique des associations biologiques*, 2e partie. Hermann, Paris.
9. Nkague Nkamba L. (2012) Robustesse des seuils en épidémiologie et stabilité asymptotique d'un modèle à infectivité et susceptibilité différentielle. Thèse, Université de Lorraine et Université Gaston Berger.
10. Pressat R. (1995) *Éléments de démographie mathématique*. AIDELF, Paris.