

## LA MATEMATICA DELLE EPIDEMIE: IL MODELLO SIR

di Stefano Daniele Sarti

### INTRODUZIONE

Negli ultimi mesi del 2019 inizia a diffondersi nella popolazione umana un virus, denominato COVID-19, e dà origine a una pandemia che ha investito l'intero pianeta e che ancora oggi, 1° luglio 2020, non si è arrestata.

Il 23 febbraio 2020, per contenere i contagi, il governo italiano chiude le scuole e le università di alcune regioni del nord; chiusura che nel giro di breve tempo (9 marzo) si estende a tutto il territorio nazionale. Si inizia anche a parlare quotidianamente di epidemiologia e dei modelli matematici per lo studio della diffusione delle malattie infettive che ne stanno alla base. Tali modelli permettono di fare previsioni e di disegnare i diversi scenari dell'evoluzione di un'epidemia e pertanto guidano le scelte di coloro che hanno responsabilità politiche e di governo della società.

Da quel momento la scuola italiana prosegue con la didattica a distanza e inevitabilmente, visto che insegno matematica e fisica in un liceo scientifico, gli studenti cominciano a pormi domande su quei modelli e in particolare sul parametro  $R_0$ , il *numero di riproduzione di base dell'epidemia*, che a seconda che sia maggiore o minore di 1 decide delle nostre sorti.

Gli interrogativi degli studenti e la speranza di dare un contributo alla comprensione del periodo che stiamo vivendo mi hanno spinto a scrivere queste note.

L'obiettivo è quello di delineare un approccio semplice ad uno dei modelli di base per lo studio della diffusione delle malattie infettive, il modello SIR. In questo caso *semplice* significa comprensibile per uno studente di scuola secondaria di secondo grado che abbia affrontato nel suo percorso di studi il tema delle funzioni e delle successioni. Vedremo come si costruisce il modello, quali sono le sue implicazioni, qual è il significato di  $R_0$ , come fare simulazioni con un foglio elettronico per osservare cosa succede al variare dei parametri. Dal punto di vista matematico il modello porta ad un sistema di equazioni alle differenze (o a un sistema di equazioni differenziali), che non è necessario risolvere per portare a termine una semplice analisi teorica.

Questo lavoro è basato su un approfondimento del prof. Andrea Pugliese [1], professore di analisi matematica presso l'Università di Trento e tra i massimi studiosi in Italia dei modelli matematici per le epidemie.

Devo inoltre ringraziare un giovane collega, Andrea Zanellati [2], che con un seminario presso il nostro liceo ha acceso il mio interesse per la biomatematica.

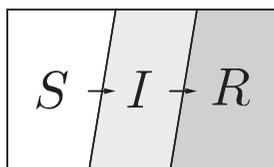
Per recenti sviluppi a carattere matematico che riguardano il modello si può consultare [3] e la relativa bibliografia.

Naturalmente esistono molti testi che trattano del modello SIR, delle sue estensioni e delle sue applicazioni, ma questo articolo è leggibile senza prerequisiti particolari poiché tutti i concetti necessari sono richiamati.

## IL MODELLO SIR

Il modello che descriveremo è stato proposto nel 1927 da Kermack e McKendrick [4]. Pur avendo quasi un secolo di vita il modello introduce molti concetti importanti per comprendere la dinamica delle epidemie. Tale modello suddivide la popolazione in tre insiemi: i suscettibili, gli infetti e i rimossi. I suscettibili sono gli individui sani che possono essere infettati; gli infetti sono coloro che, avendo la malattia, possono trasmettere l'infezione; infine i rimossi sono gli individui guariti e quindi immuni, oppure sono i morti a causa del contagio. Si suppone che l'immunità conferita dalla malattia sia permanente.

I suscettibili che vengono infettati passano nella categoria degli infetti; gli infetti, quando guariscono o muoiono, passano nella categoria dei rimossi. In questo modello non si può «tornare indietro». Possiamo schematizzare i processi con una figura.



Indichiamo con  $S$  il numero dei suscettibili, con  $I$  il numero degli infetti e con  $R$  il numero dei rimossi. Le variabili  $S$ ,  $I$  e  $R$  sono funzioni del tempo  $t$ , tempo che per ora sarà una variabile di tipo discreto (tipicamente, misurata in giorni). Lo scopo del modello SIR è prevedere l'evoluzione nel futuro delle tre variabili

$$S = S(t) \quad I = I(t) \quad R = R(t)$$

a partire dal tempo presente e con certe condizioni iniziali.

## LE IPOTESI E I PARAMETRI DEL MODELLO

Per costruire il modello matematico assumiamo le seguenti ipotesi e introduciamo alcuni parametri.

- Il numero di individui della popolazione è costante, cioè non ci sono fenomeni di emigrazione o immigrazione.
- Sia  $N$  il numero totale di individui della popolazione.

- Ogni individuo ha la stessa probabilità di entrare in contatto con qualunque altro individuo della popolazione.
- Sia  $\beta$  il numero (medio) di contatti nell'unità di tempo di un individuo all'interno dell'intera popolazione.
- Il contagio avviene per contatto diretto e la malattia non ha tempo d'incubazione, cioè ogni infetto è immediatamente infettivo.
- Sia  $\phi$  la probabilità che un suscettibile entrato in contatto con un infetto contragga la malattia e diventi un infetto.
- Tutti gli infetti diventano rimossi dopo un periodo  $T$ , che può essere pensato come periodo medio di durata (e infettività) della malattia o, nel peggiore dei casi, come periodo dopo il quale un infetto muore.

A queste ipotesi aggiungeremo sempre l'ipotesi che la popolazione studiata sia molto grande, cosicché si possano ignorare le fluttuazioni casuali. In sostanza eguaglieremo il numero di individui con una certa caratteristica con il loro numero medio.

## LA COSTRUZIONE DEL MODELLO

Poiché il numero di individui della popolazione rimane costante, per ogni  $t$

$$S(t) + I(t) + R(t) = N. \quad (1)$$

Nella schematizzazione dell'epidemia ci sono due fasi principali: il processo di contagio che porta gli individui suscettibili a diventare infetti e il processo di guarigione (o morte<sup>(1)</sup>) che sposta gli infetti nella categoria dei rimossi.

L'analisi del primo processo ci porta a capire come cambia il numero di suscettibili passando da un istante  $t$  ad un istante successivo  $t + \Delta t$ . Infatti il numero di suscettibili  $S(t + \Delta t)$  si ottiene da  $S(t)$  togliendo le persone che nel frattempo si sono ammalate. Per calcolare quest'ultimo numero facciamo riferimento alle nostre ipotesi.

Ogni infetto ha  $\beta$  contatti nell'unità di tempo all'interno dell'intera popolazione, quindi il numero di contatti nell'intervallo di tempo  $\Delta t$  è  $\beta \cdot \Delta t$ . Questo è il numero di contatti nell'intera popolazione, ma ad infettarsi possono essere solo i suscettibili, pertanto in proporzione il numero di contatti di ogni infetto all'interno dei suscettibili è

$$\beta \Delta t \frac{S(t)}{N}.$$

Tenendo conto che la probabilità che un contatto tra un infetto e un suscettibile dia luogo alla malattia è pari a  $\phi$ , e che ci sono  $I(t)$  infetti, il numero totale di individui che si ammalano in  $\Delta t$  è

<sup>(1)</sup> D'ora in poi non distingueremo i due diversi esiti, in verità molto diversi dal punto di vista umano.

$$\phi\beta\Delta t\frac{S(t)}{N}I(t). \quad (2)$$

Questo è il decremento che subisce l'insieme dei suscettibili in  $\Delta t$ , perciò

$$S(t + \Delta t) = S(t) - \frac{\phi\beta}{N}S(t)I(t)\Delta t. \quad (3)$$

Per fissare le idee facciamo un esempio. Consideriamo una popolazione di  $N = 50.000$  individui in cui il numero medio di contatti giornalieri di una persona sia  $\beta = 15$  individui/giorno. Supponiamo che una certa malattia infettiva nella fase iniziale (in cui non ci sono ancora rimossi) si stia diffondendo nella comunità con una probabilità di contagio pari a  $\phi = 20\%$ . Se ci sono 10 infetti (quindi 49.990 suscettibili), allora dopo  $\Delta t = 1$  giorno possiamo calcolare con (2) un incremento di infetti pari a 30, per un totale di 40 infetti. Dopo un altro giorno il numero degli infetti arriva a 160.

Analizziamo ora il processo che porta gli infetti a diventare rimossi in un intervallo di tempo  $[t, t + \Delta t]$ . Questo ci permetterà di scrivere l'equazione dei rimossi, infatti  $R(t + \Delta t)$  si ottiene aggiungendo a  $R(t)$  le persone che sono guarite nell'intervallo di tempo considerato.

Se  $T$  è la durata della malattia, le persone che guariscono in questo intervallo di tempo sono gli infetti che hanno contratto la malattia tra  $t - T$  e  $t - T + \Delta t$ .



Se supponiamo che all'interno degli infetti ci sia una distribuzione uniforme dell'età d'infezione, cioè del tempo trascorso dal momento dell'infezione, allora il numero di persone che passano da infetti a rimossi è

$$\frac{\Delta t}{T}I(t).$$

Questo è anche l'incremento che subisce l'insieme dei rimossi, pertanto

$$R(t + \Delta t) = R(t) + \frac{\Delta t}{T}I(t). \quad (4)$$

Per fare un esempio immaginiamo che la durata media della malattia sia  $T = 10$  giorni. Allora da qui a 2 giorni guarirà  $1/5$  degli infetti, cioè la frazione degli infetti che ha contratto la malattia nell'intervallo di tempo compreso tra 10 e 8 giorni fa.

Infine possiamo determinare come variano gli infetti in un intervallo di tempo  $[t, t + \Delta t]$ . Potremmo utilizzare direttamente l'equazione (1), ma è chiaro che il

numero degli infetti  $I(t + \Delta t)$  si ottiene da  $I(t)$  aggiungendo i suscettibili che si sono ammalati e togliendo gli infetti che sono guariti, pertanto

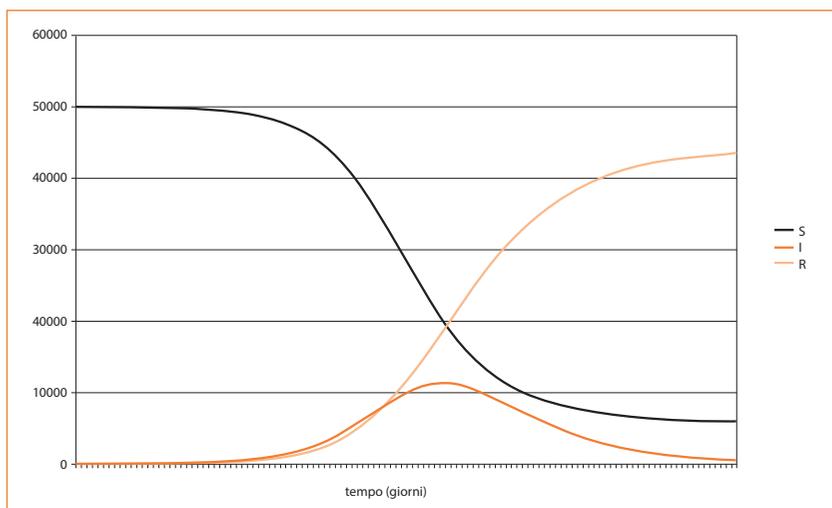
$$I(t + \Delta t) = I(t) + \frac{\phi\beta}{N} S(t) I(t) \Delta t - \frac{\Delta t}{T} I(t). \quad (5)$$

Le equazioni precedenti, nel loro insieme, costituiscono il modello SIR: permettono di fare delle previsioni sull'evoluzione delle epidemie, possono essere analizzate al fine di scoprire strategie per contenerne la diffusione, consentono di prefigurare i vari scenari al variare dei parametri, ecc. Possiamo scrivere il modello in varie forme a seconda dell'utilizzo che ne vogliamo fare.

Per fare delle simulazioni con un foglio elettronico consideriamo le equazioni (3), (4) e (5) scegliendo un intervallo di tempo unitario, per esempio  $\Delta t = 1$  giorno, e scriviamo le equazioni nel linguaggio delle successioni.

$$\begin{cases} S_{t+1} = S_t - \frac{\phi\beta}{N} S_t I_t \\ I_{t+1} = I_t + \frac{\phi\beta}{N} S_t I_t - \frac{1}{T} I_t \\ R_{t+1} = R_t + \frac{1}{T} I_t \end{cases} \quad t = 0, 1, 2, \dots \quad (6)$$

Di seguito mostriamo l'andamento tipico dell'evoluzione di un'epidemia in un periodo di quattro mesi realizzata implementando le equazioni precedenti in un foglio elettronico con i parametri  $\beta = 8$  contatti/giorno,  $\phi = 0,03$  e  $T = 10$  giorni. Naturalmente per avviare una simulazione sono necessarie delle *condizioni iniziali*; nell'esempio abbiamo posto  $S(0) = 49.990$ ,  $I(0) = 10$  e  $R(0) = 0$ .



Simulazione del modello con  $\beta = 8$  contatti/giorno,  $\phi = 0,03$  e  $T = 10$  giorni

Le tre equazioni di (6) possono essere riscritte come equazioni alle differenze ricavando da ciascuna delle equazioni il corrispondente incremento rispetto al tempo. Per esempio dalla (3) possiamo dedurre

$$\frac{\Delta S}{\Delta t} = \frac{S(t + \Delta t) - S(t)}{\Delta t} = -\frac{\phi\beta}{N} S(t) I(t).$$

Il modello SIR assume la forma

$$\begin{cases} \frac{\Delta S}{\Delta t} = -\frac{\phi\beta}{N} S I \\ \frac{\Delta I}{\Delta t} = \frac{\phi\beta}{N} S I - \frac{1}{T} I \\ \frac{\Delta R}{\Delta t} = \frac{1}{T} I \end{cases} \quad (7)$$

Se poi studiamo il fenomeno in un'opportuna scala temporale<sup>(2)</sup>, in cui  $\Delta t \rightarrow 0$ , allora i rapporti incrementali tendono alle corrispondenti derivate e il sistema precedente diventa un sistema di equazioni differenziali ordinarie (non lineare).

$$\begin{cases} S'(t) = -\frac{\phi\beta}{N} S(t) I(t) \\ I'(t) = \frac{\phi\beta}{N} S(t) I(t) - \frac{1}{T} I(t) \\ R'(t) = \frac{1}{T} I(t) \end{cases} \quad (8)$$

Osserviamo esplicitamente che a causa dell'ipotesi di popolazione costante, espressa dall'equazione (1), nei sistemi che abbiamo scritto le equazioni indipendenti sono solo due.

### IL NUMERO DI RIPRODUZIONE DI BASE $R_0$

L'analisi del sistema (8) permette di trarre conclusioni significative sull'andamento dell'epidemia. Cominciamo con l'osservare che le derivate  $S'(t)$  e  $R'(t)$  sono rispettivamente negativa e positiva per ogni  $t$ , perciò  $S(t)$  è una funzione decrescente e  $R(t)$  è una funzione crescente rispetto al tempo<sup>(3)</sup>.

<sup>(2)</sup> Possiamo immaginare di fare uno zoom all'indietro e di osservare i grafici da lontano.

<sup>(3)</sup> Queste proprietà, del resto, erano già implicite nelle ipotesi del modello.

Molto più interessante è invece l'andamento di  $I(t)$  che è determinato dal segno della derivata  $I'(t)$ . Quando l'infezione è attiva, cioè quando  $I(t) > 0$ , la seconda equazione del sistema implica che

$$I'(t) < 0 \quad \Leftrightarrow \quad \phi\beta T \frac{S(t)}{N} < 1 \quad (9)$$

Pertanto emerge in modo naturale la quantità

$$R_0 = \phi\beta T, \quad (10)$$

in funzione della quale possiamo descrivere l'evoluzione dell'epidemia:

- (i) se  $R_0 \leq 1$  allora l'epidemia non si propaga e il numero di infetti diminuisce costantemente;
- (ii) se  $R_0 > 1$  allora, supponendo che in un certo istante iniziale  $t_i$  si abbia

$$\frac{S(t_i)}{N} > \frac{1}{R_0},$$

il numero di infetti aumenta fino ad un istante  $t_{\max}$  dopo il quale comincia a diminuire senza inversioni di tendenza.

Per provare (i), osserviamo che l'insieme dei suscettibili è una parte propria dell'intera popolazione, pertanto per ogni  $t$  abbiamo

$$\frac{S(t)}{N} < 1 \leq \frac{1}{R_0}.$$

Quindi per la (9) la derivata  $I'$  è costantemente negativa e  $I$  è decrescente.

Per quanto riguarda (ii), osserviamo che poiché  $S(t)$  è strettamente decrescente, anche  $S(t)/N$  è strettamente decrescente. Se indichiamo con  $t_{\max}$  l'istante in cui <sup>(4)</sup>

$$\frac{S(t_{\max})}{N} = \frac{1}{R_0},$$

per la (9) abbiamo che  $I'(t) > 0$  per  $t_i \leq t < t_{\max}$  e  $I'(t) < 0$  per  $t > t_{\max}$ . Questo prova che l'andamento della funzione  $I$  è quello descritto.

La quantità  $R_0$  è chiamata *numero di riproduzione di base dell'epidemia* e rappresenta il numero medio di individui contagiati da un infetto nel corso del

<sup>(4)</sup> Si può dimostrare che tale istante esiste sempre.

suo periodo infettivo, supponendo che ogni altro individuo della popolazione sia suscettibile<sup>(5)</sup>.

Per esempio, se il numero di contatti è  $\beta = 8$  individui/giorno, con una probabilità di contagio  $\phi = 3\%$  e un periodo di contagiosità  $T = 10$  giorni, allora nella fase iniziale dell'epidemia ogni infetto contagherà mediamente  $R_0 = 0,03 \cdot 8 \cdot 10 = 2,4$  persone.

Nella fase iniziale dell'epidemia, in cui l'insieme dei suscettibili coincide praticamente con l'intera popolazione e  $S(t)/N \approx 1$ , è facile rendersi conto che un numero di riproduzione  $R_0 > 1$  provoca un processo di diffusione della malattia di tipo esponenziale. Infatti, se per esempio  $R_0 = 2,4$ , allora un infetto contagia 2,4 persone, che a loro volta contagieranno 2,4 persone, e così via, innescando una reazione a catena. Il numero di infetti aumenta in modo esponenziale fino a quando ci sono abbastanza persone suscettibili al contagio. A questo punto il numero di infetti raggiunge il picco e inizia la fase decrescente con l'epidemia che nel tempo si estingue<sup>(6)</sup>.

Viceversa, se il numero di riproduzione di base  $R_0 < 1$ , allora ogni infetto contagia mediamente meno di una persona e l'epidemia si estinguerà rapidamente.

La definizione (10) di  $R_0$  mostra quale possa essere una strategia di contrasto di un'epidemia, nel caso in cui non si disponga di un vaccino o di una terapia efficace. Infatti se vogliamo abbassare il valore di  $R_0$  in modo che il suo valore diventi minore di 1, non potendo agire sul tempo  $T$  di contagiosità della malattia, possiamo: o ridurre il numero medio  $\beta$  di contatti, per esempio con il distanziamento sociale, o far diminuire la probabilità di contagio  $\phi$  con mezzi di protezione individuale.

La quantità  $R_0$  è significativa anche in termini di prevenzione per progettare le campagne vaccinali. Consideriamo un'infezione con  $R_0 > 1$ . Quale frazione di popolazione dovremmo vaccinare affinché, nel momento in cui si manifestasse l'infezione, non esploda un'epidemia?

Se nell'istante  $t = 0$  avessimo

$$\frac{S(0)}{N} < \frac{1}{R_0}, \quad (11)$$

allora, poiché  $S(t)$  è decrescente, per la (9) l'epidemia andrebbe verso una rapida estinzione. Nel nostro contesto, vaccinare una persona significa toglierlo dall'insieme dei suscettibili e collocarlo nell'insieme dei rimossi, quindi il numero di persone da vaccinare è il numero iniziale di rimossi  $R(0)$  che rende vera la condizione (11). Esprimiamo il rapporto  $S(0)/N$  in termini della percentuale richiesta:

$$\frac{S(0)}{N} = \frac{N - R(0) - I(0)}{N} < \frac{N - R(0)}{N} = 1 - \frac{R(0)}{N}.$$

<sup>(5)</sup> Quest'ultima precisazione viene aggiunta poiché  $\beta$  è il numero di contatti all'interno dell'intera popolazione.

<sup>(6)</sup> La scelta di non contrastare la diffusione dell'epidemia e quindi di farla estinguere in modo *naturale*, può chiaramente avere effetti drammatici dal punto di vista umano e sociale.

La condizione (11) è certamente verificata se

$$1 - \frac{R(0)}{N} = \frac{1}{R_0},$$

dunque la percentuale critica di persone da vaccinare per evitare un'epidemia è

$$1 - \frac{1}{R_0}. \quad (12)$$

In questo caso si dice che la popolazione ha raggiunto l'*immunità di gregge*.

Vista l'importanza del numero di riproduzione di base per le strategie di prevenzione e di controllo delle infezioni sono state ottenute delle stime di  $R_0$ . Per esempio per il morbillo vengono riportati valori di  $R_0$  compresi tra 15 e 17, per la polio valori tra 5 e 6. Le percentuali di persone da vaccinare per eradicare le infezioni sono pertanto 95% e 84%, rispettivamente.

Da quando l'epidemia del nuovo coronavirus (2019-nCoV) emerso in Cina ha cominciato a diffondersi, l'Organizzazione Mondiale della Sanità e numerosi istituti di ricerca hanno diffuso stime di  $R_0$  comprese tra 1,4 e 3,8 nelle aree colpite in questa prima fase di diffusione.

Vista l'elevata letalità dell'infezione, non disponendo ancora di terapie o di vaccini, l'unica possibilità per arginare l'epidemia è, come abbiamo mostrato, far scendere  $R_0$  sotto 1 con il distanziamento sociale e con i dispositivi di protezione personale.

---

### Stefano Daniele Sarti

Liceo Rambaldi-Valeriani  
Alessandro da Imola – Imola  
sarti@imolalicei.istruzioneer.it

---

### Bibliografia

- [1] <http://maddmaths.simai.eu/divulgazione/focus/epidemie-matematica/>
- [2] A. ZANELLATI, *Il modello di Fister-Panetta per la crescita di cellule tumorali: analisi qualitativa e aspetti didattici*. Tesi di laurea, Università di Bologna, 2011.
- [3] T. HARKO – F. S. N. LOBO – M. K. MAK, *Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates*. «Applied Mathematics and Computation», 236 (2014), 184-194.
- [4] W. O. KERMACK – A. G. MCKENDRICK, *A contribution to the mathematical theory of epidemics*. «Proceedings of the Royal Society A», 115 (1927), 700-721.