

La matematica delle epidemie

Edoardo Milotti^[1,2]

Il fenomeno delle epidemie può essere descritto matematicamente in diversi modi, tutti molto interessanti. Uno di questi – il modello di Kermack e McKendrick, noto anche come modello SIR (Susceptible-Infected-Removed)¹ – utilizza alcune semplici equazioni differenziali, che contengono gli elementi essenziali per valutare l'impatto di un'epidemia.

Gli oggetti descritti dalle equazioni sono popolazioni di individui, che nel caso più semplice sono tutti i "sani", tutti i "malati" e tutti gli "altri" (individui che hanno superato la malattia e sono diventati resistenti, e quelli che sono morti).

Indichiamo dunque con $N(t)$ la popolazione sana al tempo t . Analogamente, indichiamo con $N_{i,0}$ la popolazione di malati al tempo $t = 0$ e con $N_i(t)$ la popolazione di malati al tempo t . La variazione del numero di malati al termine dell'intervallo di tempo $(t, t + dt)$ è data dall'equazione

$$dN_i = \alpha N(t)N_i(t)dt - \frac{1}{\tau}N_i(t)dt.$$

In questa equazione il termine $\alpha N(t)N_i(t)$ corrisponde al numero di individui infettati per unità di tempo, α è un coefficiente che è proporzionale alla velocità di trasmissione dell'infezione (è il numero medio di infettati per unità di tempo e per individuo infetto), e infine τ è la durata media dell'infezione. Il termine $\alpha N(t)N_i(t)$ descrive un'interazione binaria tra la popolazione sana e quella malata, afferma che il numero di nuovi infetti è proporzionale sia al numero attuale di individui sani, sia a quello di individui malati. L'ultimo termine dell'equazione, $N_i(t)/\tau$ indica la diminuzione del numero di individui infetti, sia perché guariti (e quindi sani, non più infettabili), sia perché deceduti; in entrambi i casi questi individui non rientrano più nella dinamica della diffusione della malattia.

L'equazione logistica

Se si ignora il secondo termine dell'equazione scritta sopra – che vuol dire ignorare guarigioni e morti – la si può riscrivere nella forma

$$dN_i = \alpha[N_0 - N_i(t)]N_i(t)dt,$$

che non è altro che una forma particolare di equazione logistica, un'equazione nonlineare facilmente risolvibile analiticamente, dove il parametro $N_0 = N(t) + N_i(t)$ rappresenta il numero totale di individui. Poiché vengono ignorate guarigioni e morti, in questo caso tutti gli individui della popolazione iniziale vengono infettati, e il valore finale di N_i è N_0 .

Per trovare la soluzione si utilizza il metodo della separazione delle variabili, riscrivendo l'equazione nella forma

$$\alpha dt = \frac{dN_i}{[N_0 - N_i]N_i}$$

oppure anche nella forma

$$\alpha dt = \frac{1}{N_0} \left(\frac{1}{N_i} + \frac{1}{N_0 - N_i} \right) dN_i.$$

Integrando ed utilizzando le condizioni iniziali, si trova immediatamente

$$\alpha t = \frac{1}{N_0} \left(\ln \frac{N_i}{N_{i,0}} - \ln \frac{N_0 - N_i}{N_0 - N_{i,0}} \right).$$

Con semplici manipolazioni algebriche si trova la seguente espressione che descrive il numero di malati in funzione del tempo

$$N_i(t) = \frac{N_0 N_{i,0} e^{\alpha N_0 t}}{(N_0 - N_{i,0}) + N_{i,0} e^{\alpha N_0 t}}$$

La derivata della soluzione è rilevante nella gestione pratica di un'epidemia, perché rappresenta la variazione del numero di malati per unità di tempo. Anche se la soluzione $N_i(t)$ è stata ottenuta in un caso irrealistico ($\lim \tau \rightarrow \infty$) è interessante notare che la derivata raggiunge il suo massimo in corrispondenza all'istante

$$t_{max} = \frac{1}{\alpha N_0} \ln \frac{N_0 - N_{i,0}}{N_{i,0}}$$

e in corrispondenza a questo tempo il massimo della derivata vale

$$\left. \frac{dN_i}{dt} \right|_{t_{max}} = \frac{\alpha N_0^2}{4}$$

La conclusione è che quando il parametro α è piccolo, a parità degli altri parametri, il massimo

¹ W.O. Kermack and A.G. McKendrick, *A contribution to the mathematical theory of epidemics*, Proc. R. Soc. Lond., 115(1927), 700-721

aumento del numero di contagiati si sposta in avanti nel tempo ed il suo valore è ridotto.

Passaggio al sistema di equazioni completo

Per produrre una descrizione completa bisogna introdurre anche la popolazione degli "altri", e in questo caso il sistema di equazioni che descrive il sistema è il seguente

$$\begin{cases} dN = -\alpha N(t)N_i(t)dt \\ dN_i = \alpha N(t)N_i(t)dt - \frac{1}{\tau}N_i(t)dt \\ dN_a = \frac{1}{\tau}N_i(t)dt \end{cases}$$

Osservando questo sistema differenziale si nota che la combinazione di parametri

$$\frac{1}{\alpha\tau},$$

corrisponde dimensionalmente ad un "numero di individui" e si può associare quindi alla dimensione delle popolazioni nei punti critici del sistema. In particolare, questo parametro è associato alla diminuzione della densità degli individui sani nel corso del tempo ed impedisce che il termine di interazione $\alpha N(t)N_i(t)$ porti all'infezione dell'intera popolazione. In effetti, tornando all'equazione centrale del sistema

$$dN_i = \left[\alpha N(t) - \frac{1}{\tau} \right] N_i(t)dt,$$

si vede che il "coefficiente istantaneo"

$$\left[\alpha N(t) - \frac{1}{\tau} \right]$$

corrisponde all'esponente di una crescita esponenziale, e che la crescita si arresta quando

$$N(t) = \frac{1}{\alpha\tau}$$

che è il nuovamente il numero di individui sani restanti alla fine dell'epidemia.

In base alle stesse considerazioni, si vede anche che l'epidemia non può avere inizio se

$$\alpha N(0) < \frac{1}{\tau}$$

o, in altri termini, esiste una popolazione critica iniziale

$$N(0) = \frac{1}{\alpha\tau}$$

superata la quale può iniziare l'epidemia. Questa considerazione può venire anche riformulata in termini di densità di popolazione $n(t) = N(t)/A$, considerando l'area A della regione in cui è compresa la popolazione sana:

$$\alpha n(0) = \frac{1}{A\tau}$$

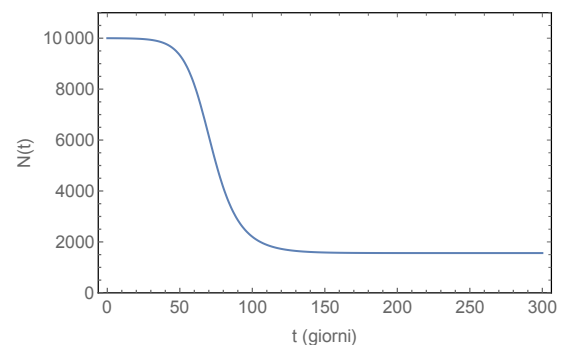
che significa che l'epidemia può partire solo se il prodotto a sinistra (numero di contagi per unità di tempo per unità di area per persona infetta) è maggiore del termine a destra.

Soluzione numerica

Per concludere questa breve serie di considerazioni vale la pena di soffermarsi su alcuni grafici ottenuti integrando numericamente il sistema.

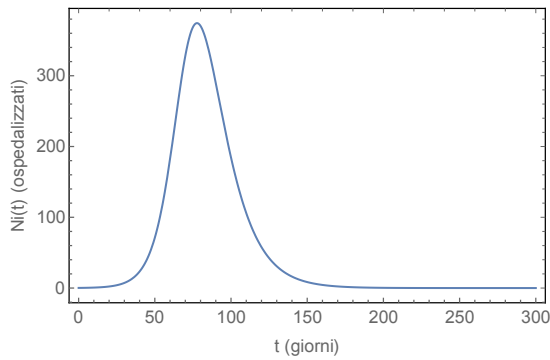
Consideriamo una malattia infettiva caratterizzata dai seguenti parametri $\tau = 10$ giorni; $\alpha\tau = 2.2 \times 10^{-4}$, e che si diffonde in una popolazione iniziale di 10000 individui, a partire da un solo individuo infetto. Il valore di $\alpha\tau$ è abbastanza alto e corrisponde ad una contagiosità elevata. Supponiamo anche che la mortalità sia del 2.5% e il tasso di ospedalizzazione del 20%.

La figura seguente mostra l'andamento del numero totale di individui sani (non contagiati):

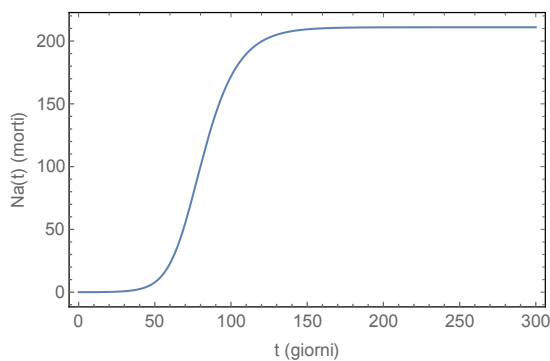


da cui si vede che una frazione elevata (circa l'80% della popolazione) viene contagiata prima che l'epidemia si spenga spontaneamente.

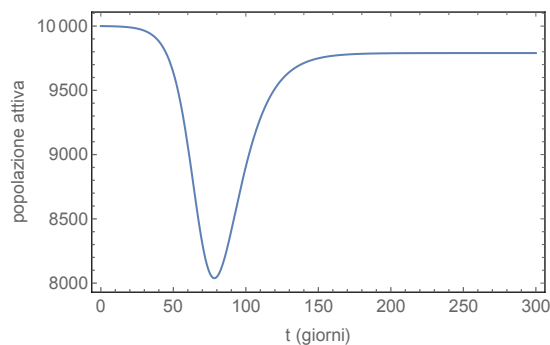
Il numero di infetti ospedalizzati è importante, ed è mostrato nella figura che segue



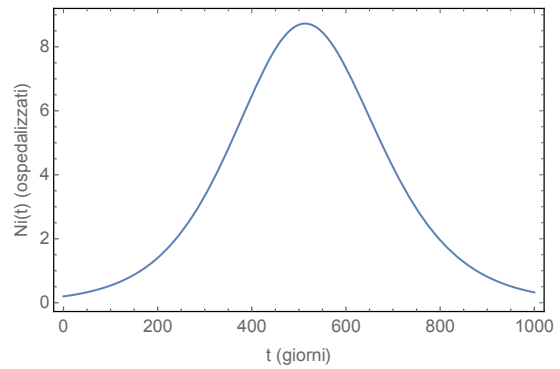
e come si vede ha un picco che viene raggiunto circa 3 mesi dopo l'infezione iniziale. Il numero totale di morti segue l'andamento del numero di infetti, ed è mostrato nella figura successiva



L'ultima figura di questa serie mostra il numero di individui attivi (sani + guariti)

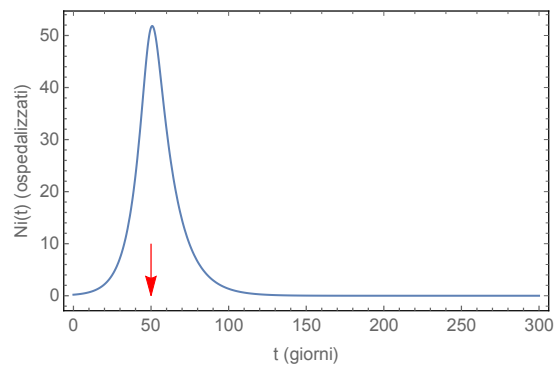


Il picco degli ospedalizzati può essere particolarmente rilevante perché mette sotto stress le strutture sanitarie. Come fare per ridurlo? Per rispondere alla domanda, possiamo ricorrere alle considerazioni fatte prima riguardo alla soluzione con l'equazione logistica, e notare che il picco si sposta in avanti nel tempo e si riduce in ampiezza se si riduce α . Questo termine di contagiosità si può ridurre con misure di isolamento della popolazione. Se si riesce a dimezzarne il valore con un isolamento parziale passando ad $\alpha\tau = 1.1 \times 10^{-4}$ fin dall'inizio dell'epidemia, si trova il seguente grafico per il numero di ospedalizzati



In questo caso la malattia si propaga molto più lentamente nella popolazione, il picco di ospedalizzazione cala di un fattore ~ 50 , e si sposta ad un anno e mezzo dall'inizio della diffusione del contagio.

Il modello ci permette anche di dire cosa succede in un caso intermedio, in cui si parte con la stessa contagiosità elevata $\alpha\tau = 2.2 \times 10^{-4}$, che viene poi abbassata di un fattore 10 in seguito a misure molto restrittive che vengono prese con gradualità tra il 45esimo e il 55esimo giorno. Il centro di questo intervallo è indicato dalla freccia rossa nella figura che segue, che mostra l'effetto delle misure sul numero di ospedalizzati, e indica che in tal modo non solo si abbassa il valore di picco, ma si contiene anche la durata dell'epidemia



Necessità di modelli più complessi

A questo punto ci si può chiedere, come sia stata fatta la stima del parametro α che riguarda la contagiosità. Di solito tra i dati dell'epidemia viene riportato il numero medio di nuovi contagi prodotti da ciascun infetto e se riconsideriamo la prima e la seconda equazione del sistema differenziale vediamo che il numero medio di contagi per unità di tempo prodotti da ciascun infetto è semplicemente il fattore $\alpha N(t)$. Perciò se assumiamo che la misura sia fatta all'inizio dell'epidemia, con una popolazione sana di N_0 individui, e che il tempo di decorso della malattia τ corrisponda anche al tempo medio di infettività, troviamo che il numero medio di contagi indotti da ciascun infetto è $r = \alpha\tau N_0$. Se

si fa riferimento alla sezione precedente si vede che nel nostro esempio questo corrisponde a $r = 2.2$ senza misure di contenimento dell'epidemia.

Naturalmente ci si potrebbe anche chiedere perché il valore N_0 rientri in questa definizione di α . Il fatto è che α deve tenere conto sia della contagiosità intrinseca della malattia, sia del numero medio di contatti tenuti da ciascun infetto, in modo tale che il prodotto αN_0 si riferisca solo ad una frazione della popolazione sana, infatti *un infetto non può avere contatti con l'intera popolazione distribuita su un vasto territorio*. In altre parole, α deve compensare una forte limitazione del modello, il fatto di non tenere conto della distribuzione spaziale della popolazione.

Queste ultime considerazioni indicano che *modelli più avanzati delle epidemie devono dipendere non solo dal tempo ma anche dallo spazio*, in altri termini devono contenere dei termini di diffusione. Inoltre, i termini diffusivi devono essere anisotropi, per tenere conto della diversa velocità di propagazione in direzioni diverse, e lo spazio non può essere semplicemente connesso, ma deve tenere conto della possibilità di avere collegamenti a lunga distanza.

Infine, un'ovvia critica che si può fare al modello è che i parametri devono variare a seconda della classe d'età degli individui, e quindi in un modello più preciso gli individui devono essere suddivisi in classi di età, così che alla fine le equazioni devono incorporare termini del tipo

$$\sum_{m,n} \alpha_{mn}(\mathbf{x}, t) N_m(\mathbf{x}, t) N_{n,i}(\mathbf{x}, t)$$

dove $N_m(\mathbf{x}, t)$ rappresenta la densità di popolazione sana di classe d'età m nella posizione \mathbf{x} al tempo t , e $N_{n,i}(\mathbf{x}, t)$ rappresenta la corrispondente densità di infetti di classe d'età n , e infine $\alpha_{mn}(\mathbf{x}, t)$ è la "contagiosità" della classe d'età n nei confronti della classe d'età m .

Nonostante tutte queste limitazioni, il modello di Kermack e McKendrick mostra degli andamenti medi che rispecchiano in modo assai realistico quelli osservati nelle epidemie, e non si può non concludere che anche equazioni e modelli semplici possono aiutare in modo significativo nella comprensione della diffusione delle malattie infettive e suggerire misure per il contenimento del contagio.

[1] Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, [2] Dipartimento di Fisica, Università di Trieste. E-mail: edoardo.milotti@ts.infn.it