



L'indice che decide i *lockdown*. Come funziona e quali sono i problemi di R_t ?

di Carlo Cottarelli e Giulio Gottardo
18 novembre 2020

L'indice R_t corrisponde al numero medio di persone al tempo t a cui un infetto trasmetterà una malattia. In linea di principio, è possibile stimare R_t a partire dalla curva dei nuovi contagi e da semplici dati epidemiologici. In pratica le difficoltà sono rilevanti. Cerchiamo di capire quali sono.

* * *

Cosa significa R_t ?

Il significato del famigerato R_t è intuitivo. Rappresenta il numero medio di persone a cui un infetto al tempo t trasmetterà una malattia in tutti i periodi successivi. È chiaro perché la soglia "critica" dell' R_t è 1: se l'indice è minore di 1, in media, ogni positivo contagerà meno di una persona e quindi il numero dei contagiati si ridurrà nel tempo. Di contro, un R_t maggiore di 1 si tradurrà in una crescita dei nuovi casi che le guarigioni non potranno compensare.

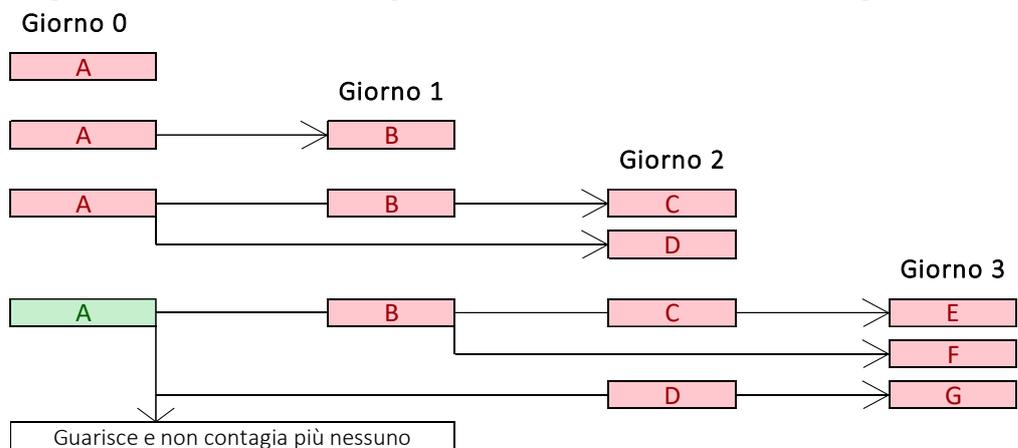
Come si calcola l' R_t ?

Una premessa: il "vero" R_t (solitamente contrassegnato con R_t^c) è calcolabile soltanto *ex post*, ovvero quando si è in possesso dei dati sui contagi avvenuti dopo il tempo t . Tuttavia, è possibile stimare R_t al tempo t o poco dopo. Per farlo, invece di attendere i dati sui contagi che avverranno, si utilizzano i dati sulle infezioni registrate finora, assumendo che quelle dopo il tempo t si evolveranno in modo simile.

Con questa premessa, si può intuire quali informazioni servano per calcolare R_t attraverso un esempio. Supponiamo che il primo giorno in cui una persona è diventata contagiosa infetti un'altra persona, quindi un'altra persona il secondo giorno, e che in seguito (terzo giorno) guarisca. Sappiamo quindi che l' R_t è 2. Ma supponiamo di non saperlo. Conosciamo però il numero di nuovi contagiati giorno per giorno:

- All'inizio ce n'era solo 1 (il paziente A, il nostro paziente "zero");
- il primo giorno sempre 1 (il paziente B contagiato da A);
- il secondo giorno ci sono 2 nuovi contagiati (quello infettato da A e quello infettato da B; chiamiamoli C e D);
- il terzo giorno A è guarito (o per lo meno non è più contagioso), ma B, C e D ne infettano altri 3;
- con la stessa logica si capisce che il quarto giorno ce ne sono 5 e così via (vedi Figura 1).

Fig. 1: Rappresentazione grafica dell'evoluzione dei contagi



Sapendo quanti contagiati ci sono ogni giorno e qual è la distribuzione dei contagi nel tempo (cioè che un contagiato ne contagia 1 il giorno dopo l'infezione e un altro due giorni dopo, prima di guarire) si può dedurre che l'andamento dei contagi osservato giorno per giorno richiede che l' R_t sia 2.

C'è una formula che riassume tutto questo ed è:

$$I_t = R_t \sum_{s=1}^t w_s I_{t-s}$$

dove I_t sono i nuovi contagiati al tempo t (e quindi I_{t-s} gli infetti al tempo $t-s$) e w_s è l'infettività relativa al tempo t di un individuo che è stato infettato al tempo $t-s$.¹ Nel breve esempio appena illustrato avremmo $w_1 = w_2 = 0,5$ e $w_{s>2} = 0$. Per il Covid-19, w_s è massimo quando s è tra 6 e 7 giorni, ovvero, l'infettività di un positivo è maggiore a circa 6 giorni dal momento del contagio.

R_t si può quindi calcolare a partire da una "curva dei contagi", ossia dal numero dei nuovi contagi ogni giorno, e dall'insieme dei coefficienti di infettività

¹ Per coerenza matematica con la definizione di R_t , si impone $\sum_{s=1}^t w_s = 1$, in modo che w_s rappresenti solo l'infettività relativa di un infetto in ciascun momento senza influenzare la rapidità complessiva della diffusione del contagio, che è espressa da R_t .

relativa (ossia dai “pesi” w_s). Basta risolvere l’equazione soprastante per R_t . Da dove si prendono i pesi, w_s ? Di solito, dagli studi epidemiologici passati.

Una precisazione: i “nuovi infetti” rilevanti per la formula non sono quelli che ci dicono alle 17 di ogni giorno, che sono i contagiati osservati attraverso i tamponi. Occorre sapere quando si sono contagiati ed è per questo che, per valutare la qualità di dati forniti dalle regioni, si va a vedere se le regioni raccolgono i dati sul giorno in cui i primi sintomi sono stati osservati. Si utilizza quindi la data di inizio dei sintomi come la data del contagio, anche se sintomi e contagio non sono la stessa cosa (vedi sotto).

Come opera l’ISS per stimare R_t ?

La metodologia seguita dall’Istituto Superiore di Sanità (ISS) per stimare R_t è una versione (molto più) sofisticata a livello statistico di quanto mostrato finora.² In particolare, la metodologia dell’ISS consente anche di ottenere intervalli di confidenza (cioè un intervallo in cui R_t si colloca con maggiore probabilità), in modo da avere un quadro più completo della variabilità delle stime (per esempio si dice che al 95 per cento R_t sta tra 1,6 e 2, con una stima centrale di 1,8, che è quella che viene riportata).

In ogni caso, come abbiamo visto, la stima di R_t richiede (1) una curva dei contagi e (2) una valutazione dell’infettività dei contagiati nel tempo (nell’esposizione semplificata, la serie dei pesi w_s). Per quanto riguarda quest’ultima, l’ISS si basa su uno studio epidemiologico condotto all’inizio della pandemia in Lombardia.³

I problemi principali riguardano i dati sui contagi. **L’ISS calcola R_t utilizzando soltanto i casi sintomatici**, in quanto “l’individuazione delle infezioni asintomatiche dipende molto dalla capacità di effettuare screening da parte dei dipartimenti di prevenzione e questa può variare molto nel tempo”. In altre parole, il dato sui sintomatici è l’unico abbastanza affidabile nel tempo da poter essere utilizzato per stimare R_t . Questo potrebbe portare ad una sottostima,

² Si veda https://www.iss.it/primo-piano/-/asset_publisher/o4oGR9qmvUz9/content/faq-sul-calcolo-del-rt per una rapida illustrazione delle caratteristiche del metodo seguito dall’ISS. Si veda invece Anne Cori, Neil M. Ferguson, Christophe Fraser, Simon Cauchemez, 2013, *A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics*, *American Journal of Epidemiology*, **178(9)**, p. 1505–1512 (<https://academic.oup.com/aje/article/178/9/1505/89262#86224209>) per l’esposizione puntuale della metodologia adottata dall’ISS.

³ Più precisamente, la metodologia dell’ISS necessita dei due parametri di una funzione Gamma $\Gamma(k, \theta)$, che si ricavano appunto dal seguente studio sulla Lombardia all’inizio dell’epidemia: D Cereda, M Tirani, F Rovida, V Demicheli, M Ajelli, P Poletti, F Trentini, G Guzzetta, V Marziano, A Barone, M Magoni, S Deandrea, G Diurno, M Lombardo, M Faccini, A Pan, R Bruno, E Pariani, G Grasselli, A Piatti, M Gramegna, F Baldanti, A Melegaro, S Merler, 2020, “The early phase of the COVID-19 outbreak in Lombardy, Italy”, <https://arxiv.org/abs/2003.09320>.

soprattutto alla luce dell'elevata incidenza di asintomatici, che, anche se in misura minore, possono essere contagiosi e quindi provocare successive infezioni, anche sintomatiche. A riprova di questo, a inizio settembre l'**Istituto Nazionale di Fisica Nucleare ha provato a stimare R_t includendo gli asintomatici, e ha ottenuto un valore quasi triplo rispetto all'ISS** (circa 3,0 vs 1,2), anche se molto più variabile, a causa dell'instabilità del dato degli asintomatici.⁴

C'è un problema ulteriore che riguarda "il ritardo" delle stime dell' R_t . Ragionevolmente, l'ISS utilizza dati "consolidati" sul numero di contagiati. Il consolidamento consiste nel correggere i numeri dei nuovi casi (sintomatici) per tenere conto di tutte le particolarità nel tempismo della loro registrazione e comunicazione. Inoltre, da questo totale si sottraggono anche i cosiddetti "casi importati", per evitare di gonfiare erroneamente l' R_t di una regione (o del paese intero). Queste operazioni di consolidamento ed esclusione richiedono diversi giorni, facendo sì che **l' R_t disponibile sia al massimo quello della settimana precedente.**

Infine, recentemente è emerso come i dati sui contagi riportati da alcune regioni – e poi utilizzati per stimare R_t dall'ISS – fossero incompleti, portando a una sottostima ulteriore dell'indice.⁵

⁴ Si veda:

https://www.ansa.it/canale_scienza_tecnica/notizie/fisica_matematica/2020/09/09/covid-dalle-stime-con-asintomatici-lindice-rt-vicino-a-3- fc5b91c3-82ce-4a2b-9d43-21cb989100b6.html.

⁵ Di questo è stata data notizia attraverso un tweet. Vedi

<https://twitter.com/OpenCovidM/status/1325964881967931394?s=20>.