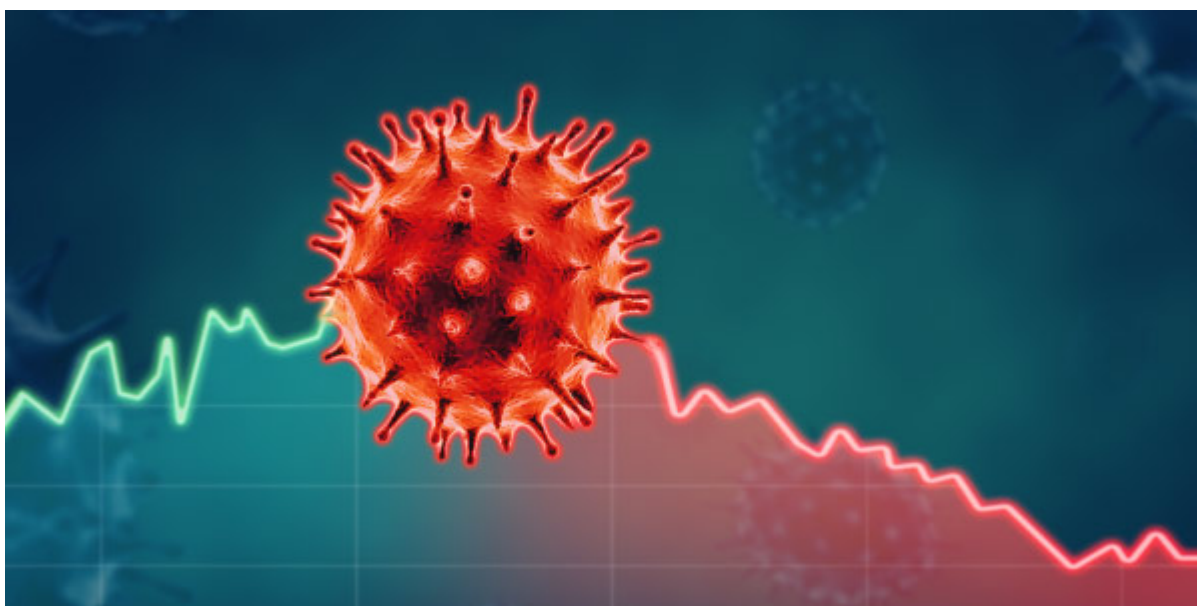


Covid-19: Proviamo a prevedere cosa succederà



Stiamo vivendo giorni terribili, e ancora non sembra di vedere la fine del tunnel. A questo proposito, Mauro Maltagliati, dopo aver riflettuto sulla scelta della miglior variabile da analizzare per capire come sta evolvendo in Italia la pandemia da Coronavirus, ne segue l'evoluzione temporale e azzarda una previsione: l'emergenza finirà il... (leggere per sapere)

Tutta Italia segue ormai con interesse quasi spasmodico le statistiche sul Coronavirus, con il comprensibilissimo obiettivo di capire come evolverà il fenomeno e, soprattutto, quando finirà la pandemia, almeno nel nostro paese.

Quando ci si avventura in previsioni, le strade che si possono seguire sono essenzialmente due. La prima è forse logicamente un po' più difendibile: si usa un modello "causale", e cioè si prevede l'evoluzione del fenomeno di interesse (la diffusione del virus) in funzione di altri, quelli che lo favoriscono o lo frenano, come ad esempio le politiche di contenimento, il meteo (con l'arrivo della stagione calda: servirà?), i progressi della medicina, ...

Però la relazione tra questi fattori "causali" e la diffusione del virus non è chiara, e ancor meno lo è la loro evoluzione nei prossimi giorni e nelle prossime settimane. Insomma: una strada difficile.

Un approccio di tipo “serie storiche”

Proviamo allora a seguire una pista diversa, del resto già intrapresa da molti, anche su Neodemos. Si osserva l'evoluzione del fenomeno nel recente passato (diciamo 10-15 gg), si ipotizza che la regolarità già evidenziata venga mantenuta anche in seguito e si proiettano i valori così ottenuti. E' un approccio di tipo “black-box”, o “scatola nera”: non si indaga sui meccanismi causali di cui si diceva prima, ma ci si limita a seguire il fenomeno, e a estrapolarlo al futuro. Ovviamente, questo approccio è difendibile solo fino a che non si manifestano “shock” di qualsiasi natura, che nel caso specifico si spera siano in meglio (si esita a dire “positivi”), come ad esempio un insperato successo della politica di confinamento degli italiani, o la scoperta di una cura.

Tanti numeri, poca chiarezza

Dobbiamo però scegliere una variabile da analizzare. In questi giorni se ne nominano tante (contagiati, positivi, morti, tasso di mortalità, tasso di letalità, ...), ma poche dicono quel che comunemente si pensa.

Prendiamo il numero di infetti, o contagiati, ad esempio. Noi conosciamo in realtà i “positivi”, ovvero il numero di infetti tra coloro che sono stati sottoposti al tampone. Però al test viene sottoposto solo chi è fortemente sospettato di essere infettato, e questo distorce la visione, sia del livello che della tendenza del fenomeno. Per conoscere davvero i contagiati, o meglio per stimarli correttamente, si dovrebbero fare tamponi su un campione casuale della popolazione - operazione evidentemente improponibile oggi (e, beninteso, non la sto qui proponendo).

Persino i “guariti” sono difficilmente misurabili, in quanto sono in realtà una frazione dei “positivi”, e non degli infetti. Paradossalmente, l'aumento dei guariti può addirittura fornire un'immagine fuorviante: non è necessariamente una buona notizia, perché, a parità di successo della terapia, i guariti aumentano all'aumentare dei malati.

Quanto poi ai vari tassi, i problemi di misura sono anche maggiori. Il tasso di mortalità ad esempio si conoscerà solo alla fine dell'epidemia (morti per Covid-19 / residenti). Nel frattempo, si potrebbe pensare al tasso di letalità (morti per Covid-19 / infetti), che però non è facile da calcolare correttamente per due ragioni:

1) come detto, non conosciamo il numero dei contagiati, ma solo dei positivi, che più o meno corrispondono ai contagiati conclamati, e quindi più gravi. Se per il tasso di letalità usassimo il denominatore corretto (gli infetti, e non i positivi), la letalità sarebbe probabilmente molto bassa.

2) c'è un lag temporale tra quando si viene infettati e quando disgraziatamente si muore (se ciò avviene), e per calcolare la vera letalità dovremmo sapere "che fine faranno gli infettati" - cioè, se muoiono o invece guariscono. Ma poiché non conosciamo l'evoluzione del numero di infetti, né il momento in cui si sono infettati, né i tempi (diversi tra loro) in cui ciascun infetto si trasforma in un guarito o in un morto, il quadro si complica considerevolmente

Facciamo parlare i morti

Scartate le altre variabili, rimane la triste, ma necessaria (e in questo caso anche utile) conta dei morti. Un dato meno incerto di tanti altri, ma che ha anch'esso bisogno di qualche precisazione. Non in tutte le nazioni, infatti, ci si riferisce, come avviene in Italia, al numero di morti *con* Covid-19 (contrapposti ai morti *per* Covid-19). Eppure si tratta della definizione più affidabile, dato che non è banale accertare se alcuni individui, magari con altre gravi patologie, sarebbero rimasti in vita (e per quanto tempo) se non fossero stati infettati dal Corona virus.

Come avviene per tutte le epidemie con un certo grado di letalità, il numero totale dei morti evolve inizialmente secondo una "legge" esponenziale (i morti di oggi sono in rapporto costante, e maggiore di uno, con i morti di ieri), che poi muta in legge "logistica" (curva a "esse"; v. di nuovo neodemos.info), il che significa che alla fine l'incremento, e cioè il numero di morti giornalieri, scende a 0. (NdA Non necessariamente una buona notizia: potrebbe anche voler dire che siamo ormai morti tutti.)

E in effetti, dal 27 febbraio al 10 marzo, e cioè all'inizio dell'epidemia in Italia, il numero di morti ha seguito quasi perfettamente una legge esponenziale (tab. 1):

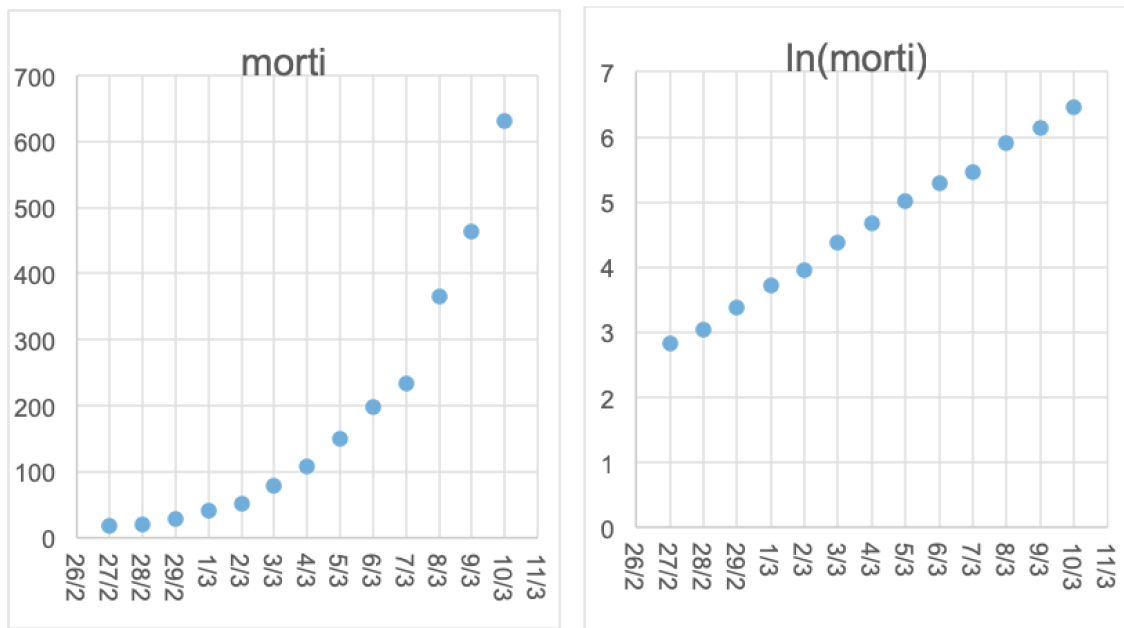
**Tab. 1. Morti in Italia con Corona Virus,
dal 27 febbraio al 10 marzo**

data	morti	ln(morti)
27/02	17	2,83
28/02	21	3,04
29/02	29	3,37
01/03	34	3,53
02/03	52	3,95
03/03	79	4,37
04/03	107	4,67
05/03	148	5,00
06/03	198	5,28
07/03	233	5,45
08/03	366	5,90
09/03	463	6,14
10/03	631	6,45

Se si mettono questi dati su un grafico
(in scala naturale a sinistra, e logaritmica a destra),
si ottiene un allineamento quasi perfetto dei punti
lungo una retta di pendenza 0,3071 ($R^2=0,9981$).

La pendenza della retta stimata sulla trasformata logaritmica ci dice che nel periodo considerato il numero di morti giornalieri è stato circa il 31% del valore (complessivo) aggiornato al giorno precedente. Ma questa evoluzione regolare (lineare nel logaritmo, e quindi esplosiva nei numeri naturali) deve prima o poi trasformarsi in una curva logistica, e non (più) esponenziale. E, per fortuna, questo è quello che sta in effetti succedendo.

Fig. 1. Morti in Italia con Corona Virus, dal 27 febbraio al 10 marzo, in scala naturale e logaritmica

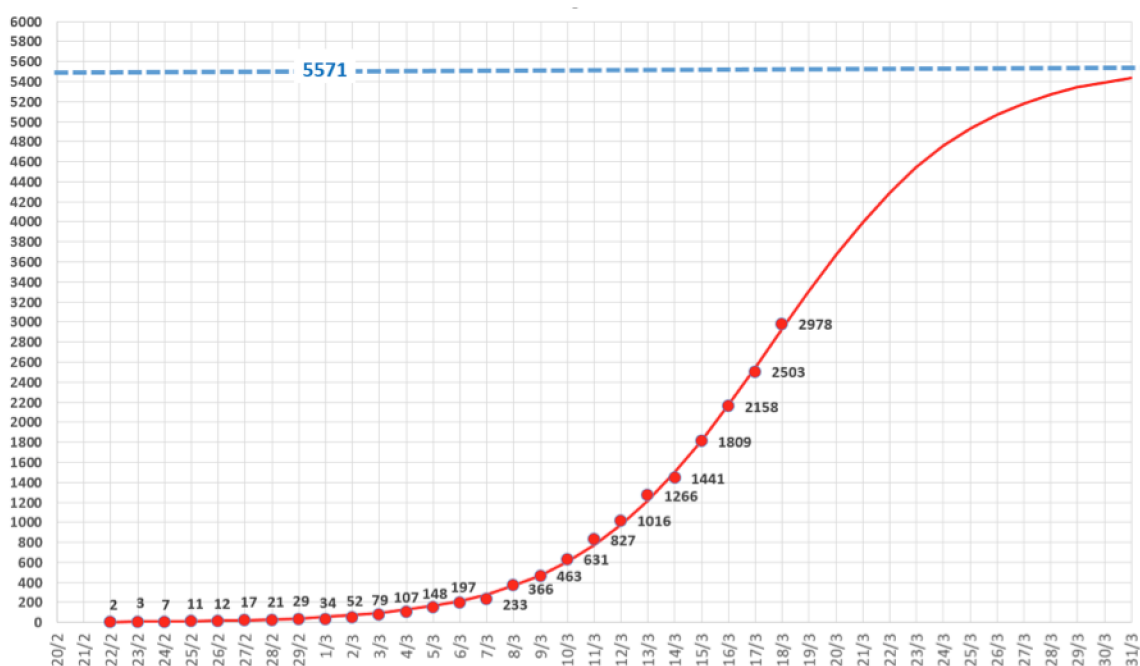


Fonte: Ministero della Salute*

La stima della curva logistica e del ... D-Day

Se aggiungiamo i dati degli ultimi giorni otteniamo la figura 2. Su questa base, gli statistici come me possono fare il loro lavoro e stimare la curva logistica della cumulata dei decessi. Se, come in questo, si osserva un buon adattamento per il passato, si può anche azzardare una previsione sul futuro, e in particolare sull'esaurimento del processo (=fine della pandemia, almeno in Italia).

Fig. 2. Stima e previsione del numero dei morti in Italia con Corona Virus, dal 20 febbraio al 31 marzo.



Fonte: Elaborazioni dell'A. Dati del Ministero della Salute *

Beh, in senso stretto, a dire il vero, una logistica non “finisce” mai: tende a un asintoto orizzontale (il numero totale dei morti), senza però mai raggiungerlo. Ma diciamo che ci accontentiamo di stimare un giorno che possiamo convenzionalmente definire come di “fine della pandemia”: quello dopo il quale si avranno, complessivamente, meno di 10 morti *con* Corona virus. Ebbene, questo giorno dovrebbe essere ormai vicino: il 10 aprile, o giù di lì.

Cautela!

Beninteso, questa è solo una stima: posso sbagliare, e tra pochi giorni si saprà di quanto. Possiamo intanto dire che se la stima fosse perfetta, il dato di oggi pomeriggio (19 marzo) dovrebbe portare a un numero (cumulato) di morti pari a 3.300, e cioè circa 330 più di ieri (18 marzo).

Non dobbiamo poi dimenticare che, man mano che arrivano nuovi dati, il modello si adatta e le stime si aggiustano. Inoltre, questo esercizio di previsione sottintende che siamo tutti omogenei, e cioè tutti ugualmente soggetti a mortalità da questo virus, mentre sappiamo così non è: ci sono persone più o meno resistenti.

Ci sono fattori che possono mostrare discontinuità, collassando (la sanità?), o invece dando spinte positive (la politica di confinamento, i cui effetti sui morti, ancora, in pratica, non si sono visti, e la scoperta di nuove cure).

E c'è il mistero del Sud d'Italia, finora relativamente risparmiato dell'epidemia: terra vergine (e quindi soggetta a rapido peggioramento) o terra protetta (dalle misure di confinamento)? La risposta, ormai, tra pochi giorni.

Fonte figure

*salute.gov.it - il portale del Ministero della Salute