

Comportamento logistico dell'infezione Covid-19 in Italia al 7 marzo 2020

Vincenzo Fiorentini

Dipartimento di Fisica, Università di Cagliari, Cittadella Universitaria, Monserrato, I-09042 Cagliari, Italy

(Dated: March 8, 2020)

Tramite interpolazione non lineare con funzione logistica, si forniscono stime dell'andamento temporale del numero di infetti, pazienti ICU, e (tentativamente) di decessi nell'epidemia Covid-19. **Al 7 marzo 2020**, l'evoluzione è chiaramente logistica e non esponenziale. La stima del cumulativo finale è 16000 infetti e 1600 ICU, con saturazione intorno al 20-25 marzo.

È in corso in Italia una epidemia causata dal virus Covid-19. Per evitare sottovalutazioni e allo stesso tempo isterismi ingiustificati (che abbondano nei media) è utile fare chiarezza sul comportamento temporale del numero di infetti, di pazienti, o di decessi,

In questa nota, fornisco una stima dell'andamento temporale (quindi della durata) del processo, e del numero massimo delle diverse categorie affette alla presunta fine del processo; in particolare, discuto il totale degli infetti e il numero di pazienti in terapia intensiva o ICU (statisticamente più consistenti) e, tentativamente, dei decessi.

Nelle fasi iniziali di una epidemia il numero di casi (e quindi ospitalizzati, ICU, decessi, etc.) aumenta esponenzialmente con il tempo t :

$$n = \exp(a * t + m).$$

I parametri k e m , e quindi n , sono diversi per ogni categoria (una stima è stata fornita da Bucci e Marinari [1] alcuni giorni fa, e ne fornisco una corrente più oltre). La popolazione affetta essendo finita, il comportamento esponenziale non può applicarsi in generale, e va ipotizzato un comportamento, a tempi sufficientemente lunghi, di tipo "a gradino": la funzione logistica o sigmoide ne è un esempio. Per semplificare, utilizzo

$$n = \frac{L}{1 + \exp(-k * (t - t_0))}. \quad (1)$$

I tre parametri sono la popolazione finale $L \equiv n$ o valore di saturazione, il tasso di salita k e il punto di inflessione t_0 (anche qui, diversi per ogni categoria). Ho replicato i calcoli per quattro classi di funzioni diverse (logistica, Brillouin, Langevin, error-function) ottenendo risultati pressochè identici per le prime due classi. (La terza e la quarta tendono a produrre stime rispettivamente minori e maggiori, di un fattore circa 1.5-2.)

Ho utilizzato una ottimizzazione non-lineare ai minimi quadrati, che determina (in soldoni) i valori dei parametri della funzione che minimizzano la distanza della stessa dai dati epidemiologici. (Per l'esponenziale ho usato un fit lineare ai dati logaritmizzati.) Per un definito set di dati epidemiologici, i risultati sono robusti, cioè indipendenti dal punto di partenza dell'ottimizzazione (i valori iniziali di L , k , t_0), dalle ipotesi sugli errori, e dai diversi metodi utilizzati (Levenberg-Marquardt, dogbox, TRF).

Con i parametri così determinati si possono calcolare i valori di n a tutti i tempi futuri, e la qualità delle previsioni dipende dalla consistenza e accuratezza dei dati utilizzati. Qui utilizzo i dati forniti dalla Protezione Civile

(e utilmente collazionati da Davide Magno [2]) che assumo, appunto, essere accurati e consistenti oltre ogni dubbio. In particolare, **questa versione riporta previsioni basate sui dati fino al 7 marzo 2020**. Aggiungerò il report a cadenza regolare in base ai nuovi dati.

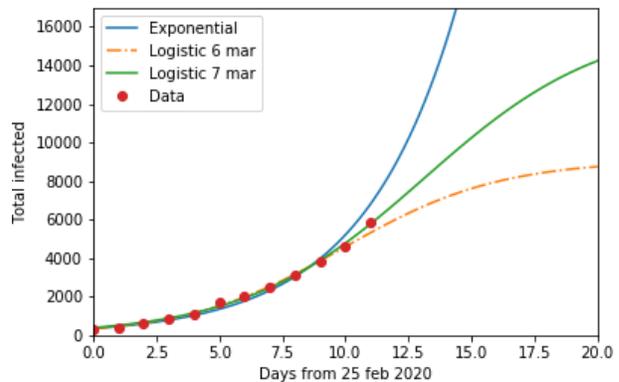


FIG. 1. Andamenti esponenziale e logistico vs dati epidemiologici (totale infettati).

I risultati per n di totale infettati, pazienti ICU, e decessi sono mostrati nelle Figure 1-3. In tutti i casi il comportamento è logistico e non esponenziale (per ragioni di spazio evito cambiamenti di scala o di tipo di scala; la cosa rimane evidente anche in altre rappresentazioni). La predizione in Fig.1 (infettati totali) è molto sensibile: con i dati fino al 6 marzo la predizione finale è 9000 circa, mentre con quelli fino al 7 marzo circa 16000. La differenza è dovuta tutta all'ultimo punto; nei giorni scorsi la predizione era stabile sotto i 10000. Ovviamente, variazioni nei criteri di reporting e assessment medico possono influenzare l'andamento dei dati totali. La stima potrebbe ancora cambiare di molto sia in eccesso che (più probabilmente) in difetto. Gorini [3] dà una stima simile alla mia di ieri, ma con i dati fino a oggi; il punto del 7 marzo è forse un accumulo di dati "vecchi".

In Fig.2 (pazienti ICU, un dato relativamente solido e obiettivo), la logistica descrive bene i dati giornalieri già da alcuni giorni, e la predizione del valore di saturazione è circa 1500 sia per i dati del 6 che del 7 marzo.

In Fig.3 sono riportati i dati per i decessi. Anche qui, la logistica predice i valori giornalieri con buona precisione già da un po', e il valore di saturazione si sta progressiva-

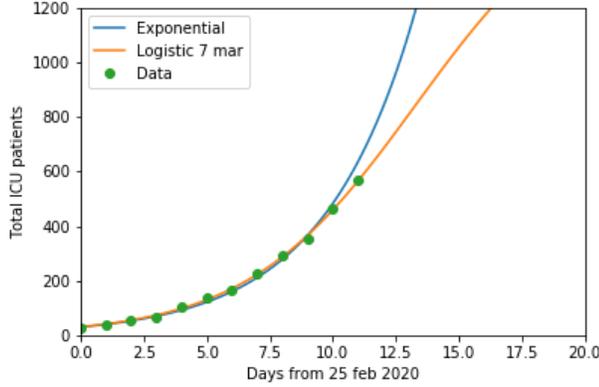


FIG. 2. Come Fig.1, per pazienti ICU.

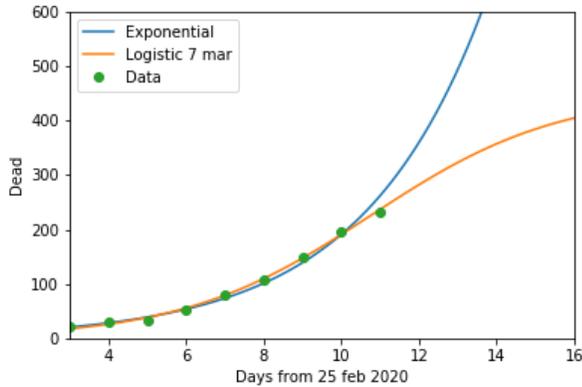


FIG. 3. come in Fig.1, per decessi.

mente stabilizzando grazie ai nuovi dati (nei giorni scorsi l'estrapolazione era completamente inaffidabile, con più morti che contagiati, ad esempio). Il numero di decessi potrebbe naturalmente essere sovrastimato (ad esempio, per decesso dovuto ad altre patologie etc.), ma, di nuovo, assumo la consistenza interna dei dati pubblici.

Le predizioni sui valori di saturazione e i punti di inflessione, con i dati odierni, sono

- $\Rightarrow n_{\text{total}}=16300$, con inflessione $t_0=11$, cioè l'8 marzo;
- $\Rightarrow n_{\text{ICU}}=1688$, con inflessione $t_0=13$, cioè l'8 marzo;
- $\Rightarrow n_{\text{dead}}=451$, con inflessione $t_0=13$, cioè il 6 marzo.

I dati sono riportati anche in Tabella I. Rozzamente, questi numeri indicano che per le attuali condizioni e mis-

ure preventive, siamo circa alla metà dell'evento, sia in termini temporali che numerici. Che le stime siano almeno plausibili è suggerito anche dai rapporti dei valori di saturazione consistenti con quelli medi storici dei dati:

- \Rightarrow decessi/ICU: medio 0.33, saturazione 0.27;
- \Rightarrow ICU/totale: medio 0.09, saturazione 0.10;
- \Rightarrow decessi/totale: medio 0.03, saturazione 0.027.

Naturalmente, per quanto confortanti quanto all'efficacia delle misure, alle prospettive temporali e al carico sanitario, **tutti questi numeri sono in aggiornamento e fortemente dipendenti dai dati analizzati**; vanno perciò presi *cum grano salis*.

TABLE I. Parametri di logistica ed esponenziale per totale, ICU, decessi (dati 7 marzo 2020)

	Totale	ICU	Decessi
Logistica			
$n \equiv L$ (saturazione)	16300	1688	451
k (salita, 1/giorni)	0.28	0.30	0.41
t_0 (inflessione, giorni)	13.1	13.3	10.8
$1/k$ (salita, giorni)	3.57	3.33	2.43
Esponenziale			
a	0.268	0.275	0.316
m	5.877	3.429	2.088

TABLE II. Valori statistici (dati 7 marzo 2020) per logistica (L) ed esponenziale (E)

	S-p	paired-S-p	χ^2	χ^2-p	R^2
Totale					
L	0.988	0.722	54.05	0	0.997
E	0.936	0.546	175.50	0	0.969
ICU					
L	0.995	0.844	2.38	0.967	0.998
E	0.960	0.595	8.25	0.69	0.984

Essendo ancora dibattuto se il comportamento sia logistico o esponenziale, ho usato diverse misure statistiche di correlazione tra dati veri e stimati; il campione in Tabella II suggerisce che la logistica è un modello migliore dell'esponenziale. (En passant, il χ^2-p pressochè nullo—peraltro più piccolo per l'esponenziale che per la logistica—suggerisce ancora grandi incertezze nei dati del numero totale, mentre per le ICU, come anche per i decessi, la logistica risulta una buona ipotesi.)

Ringrazio Edoardo Gorini (U Salento) [3] per discussioni e confronto dati.

[1] E. Bucci ed E. Marinari, <https://bit.ly/32XIC4Y>[2] D. Magno, <https://bit.ly/330s21h>.[3] E. Gorini, <https://bit.ly/2It5mAv>.