

Forse stiamo leggendo male i dati... ma il risultato non cambia, anzi!

20 marzo 2020

Sono giorni che, come tutti, leggo e controllo ogni aggiornamento sui numeri di questa maledetta emergenza, ...e visto che da nessuna parte ho trovato le risposte che cercavo, provo a condividere con voi alcune considerazioni.

Partiamo dalle basi. Vediamo quali sono i primi ingredienti che considereremo per cercare di ricavare delle conclusioni:

- **Il numero di morti:** è un dato importante ai nostri fini ed è quello che considereremo come il più affidabile. Non può comunque essere considerato come un dato certo. È evidente che i criteri di classificazione dei decessi variano ampiamente tra nazione e nazione (basta confrontare i decessi della Germania con i nostri..).
- **Il numero di contagiati:** sarebbe fondamentale per valutare due parametri di interesse quali il tasso di propagazione e quello di mortalità. Purtroppo, il dato sui contagiati è proprio quello meno attendibile: oggi è ancora troppo condizionato dal contesto e dalle politiche di screening di ogni singolo focolaio. Nessuna nazione può dire di avere un dato minimamente attendibile su questo punto.
- **Il tasso di mortalità:** banalmente, è il rapporto tra i primi due ingredienti. Data l'incertezza sul numero di contagiati, è evidente che nemmeno questo parametro può essere, al momento, considerato attendibile. Si tratta comunque di un valore fondamentale per correlare il numero di decessi con il numero di contagi e capire in che direzione stiamo andando. Considero questo parametro come l'unico sostanzialmente invariante e quindi quello che meglio caratterizza ogni singolo virus. Ovviamente anche il tasso di mortalità può variare in funzione di molteplici parametri e, in particolare, lo stato di saturazione del sistema sanitario potrebbe portare ad una crescita incontrollata anche di questo. Per il nostro fine, trarremo spunto da quanto accaduto a Wuhan e assumeremo che il tasso di mortalità, in condizioni critiche, possa arrivare a raddoppiare rispetto alle condizioni "standard". Per tasso di mortalità in condizioni "standard", intenderò pertanto il rapporto tra numero totale di decessi e numero totale di contagiati a epidemia conclusa e nell'ipotesi di disporre di un sistema sanitario "normale". Un'ultima nota: poiché ogni virus può colpire con maggiore o minore aggressività specifiche categorie di età, in linea di principio anche il tasso di mortalità potrebbe variare da nazione a nazione in funzione della distribuzione demografica. In questo primo ragionamento, non terremo comunque in considerazione questa variabilità.. valuteremo successivamente se sia necessario aggiungerla.

Tutto ciò premesso, il mio primo obiettivo sarebbe quello di fare chiarezza proprio sul tasso di mortalità. L'esigenza nasce dalla grande variabilità che si riscontra analizzando i dati quotidianamente pubblicati su internet; come si vede nella tabella¹, già tra le prime 10 nazioni per contagio l'indice varia tra il nostro 8,3% e lo 0,3% della Germania. Sarebbe come se ad ogni decesso tedesco corrispondessero quasi 30 morti in Italia.

| Nazione | Contagiati | Ricoverati | Decessi | Ricoverati/ Contagiati | Decessi/ Contagiati |
|-------------|------------|------------|---------|---------------------------|------------------------|
| China | 80.967 | 69.614 | 3.248 | 88% | 4,0% |
| Italy | 41.035 | 4.025 | 3.405 | 11% | 8,3% |
| Iran | 18.407 | 571 | 1.284 | 32% | 7,0% |
| Spain | 18.077 | 1.081 | 831 | 6% | 4,6% |
| Germany | 15.320 | 86 | 44 | 1% | 0,3% |
| USA | 14.366 | 154 | 217 | 1% | 1,5% |
| France | 10.995 | 602 | 372 | 12% | 3,4% |
| S. Korea | 8.652 | 106 | 94 | 26% | 1,1% |
| Switzerland | 4.222 | 15 | 43 | 0% | 1,0% |
| UK | 3.269 | 65 | 144 | 2% | 4,4% |

¹ Fonte: <https://www.worldometers.info/coronavirus/>. Aggiornamento riferito a: March 20, 2020, 07:24 GMT

Le prime cronache provenienti dalla Cina ci hanno insegnato che la mortalità, purtroppo, può variare in funzione della capacità del sistema sanitario di far fronte ai malati critici. Stando a quanto riportato dall'OMS², nel picco dell'epidemia cinese, la mancanza di posti in terapia intensiva ha fatto sì che il tasso di mortalità dell'area di Wuhan risultasse circa 8 volte più alto di quello delle aree affette da focolai minori (5,8% contro 0,7%).

Altro aspetto da tenere in considerazione è il tempo: il Ministero della salute riferisce che *il periodo di tempo che intercorre fra il contagio e lo sviluppo dei sintomi clinici* vari tra 2 e 11 giorni³, mentre, stando ad un articolo scientifico pubblicato a fine gennaio⁴, il tempo che mediamente intercorre tra la comparsa dei sintomi e il decesso è di circa 14 giorni. Per i miei fini, considerato quanto detto, approssimerò a 20 il numero di giorni che intercorrono tra il contagio e il decesso. Tutto questo per chiarire che, in teoria, a epidemia in corso il numero di morti dovrebbe essere correlato al numero di contagi di 20 giorni prima e che, quindi, a meno che l'epidemia non possa ritenersi conclusa, non è comunque corretto estrarre il tasso di mortalità dalla tabella vista prima. Delle 10 nazioni riportate in tabella, soltanto Cina e Sud Corea presentano un andamento decrescente di nuovi casi, pertanto sono gli unici casi che considereremo per raggiungere il nostro primo obiettivo che rimane quello di individuare il tasso di mortalità "standard"... per farlo, continua a mancarci comunque il denominatore, visto che abbiamo detto che il numero di contagiati non può essere considerato un dato attendibile (nel caso della Cina tale aspetto è risaltato dall'abnorme rapporto tra ricoverati e contagiati).

In nostro soccorso è giunto un interessante articolo pubblicato su Science il 16 marzo⁵ in cui vengono stimati i casi di contagio non documentati in Cina. Prenderò pertanto questi risultati per buoni salvo, poi, commentare i risultati che ne conseguiranno. Vediamo intanto dove ci portano...

La ricerca stima che, al momento del lockdown cinese (23 gennaio), l'86% dei casi di contagio non era stato documentato. Purtroppo, il campione disponibile a quella data è troppo poco rappresentativo per tentare delle estrapolazioni, quindi provo a spingermi oltre, diciamo di 14 giorni. Confido sul fatto che, fino al 11 febbraio, la curva dei contagi rilevati mostra una crescita regolare e solo il 12 febbraio i risultati di uno screening massivo fanno impennare il dato. Applico dunque i risultati della ricerca ai dati del 7 febbraio, consapevole che il punto più fragile di tutte le considerazioni è proprio questa ipotesi e che, anche per questo, i risultati ottenuti andranno confrontati con tutti gli altri dati disponibili. Dunque: al 7 febbraio risultavano 34.546 casi di contagio, assumendo che l'86% dei casi non sia stato documentato, vorrebbe dire che avevamo in realtà 246.757 contagiati. Richiamando quanto detto in merito al fattore tempo, vado a vedere il numero di decessi di 20 giorni dopo: 2.788, divido ed ottengo un tasso di mortalità del 1,1 %. Se avessi certezza di quale fosse la direzione in cui sono state fatte tutte le precedenti assunzioni/approssimazioni (eccesso o difetto?), potrei arrotondare il valore.. purtroppo così non è, quindi per coerenza lo prendo per buono così com'è venuto.

Lo step successivo consiste nel tentare di estrapolare dati a partire da quanto ricavato, verificandone i risultati.

Prima estrapolazione: quanti sono i contagiati effettivi oggi in Italia ? per quanto detto, per il momento possiamo provare a rispondere in merito al dato di 20 giorni fa, la cosa interessante però, è che adesso non dobbiamo più preoccuparci degli effetti del lockdown perché siamo riusciti a creare un legame diretto tra

² Fonte: <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>
Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19),

³ Fonte: <http://www.salute.gov.it/portale/malattieInfettive/dettaglioFaqMalattieInfettive.jsp?lingua=italiano&id=228>

⁴ Fonte: <https://onlinelibrary.wiley.com/doi/full/10.1002/jmv.25689?af=R> *Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China*

⁵ Fonte: <https://science.sciencemag.org/content/early/2020/03/13/science.abb3221> *Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2)*

morti (dato “attendibile”) e contagiati effettivi. Devo però aggiungere un'altra considerazione: negli ultimi giorni purtroppo la concentrazione di morti in zone limitate ci porta ad ipotizzare che il tasso di morte stia già iniziando a risentire dell'“effetto Wuhan”, e cioè si stia alzando in virtù del fatto che non siamo in condizioni “standard” ma che gli operatori sanitari si stiano già trovando nella condizione di non poter curare adeguatamente tutti i pazienti critici. Ipotizzato quindi, arbitrariamente, che il tasso sia passato dal 1,1% al 2%. Ieri il numero di decessi totali registrati in Italia è stato pari a 3.405, quindi i casi di contagio al 28 febbraio, sarebbero stati **170.250** (secondo i dati ufficiali, alla stessa data avevamo 889 contagiati). Per ora tengo da parte questo dato, vedremo poi come utilizzarlo e come arrivare a stimare i casi di contagio odierni (quelli sì, saranno condizionati dagli effetti del lockdown).

Seconda estrapolazione: cosa sta accadendo in Germania ? Analizzando la solita tabella, i dati dichiarati dalla Germania appaiono incongruenti con le altre nazioni. Proviamo ad applicare il nostro tasso, in questo caso, così com'è: i 44 morti di ieri dovrebbero quindi corrispondere a 4.000 contagiati effettivi alla data 28 febbraio, contro i 74 dichiarati dal governo tedesco alla stessa data. Il fatto che anche in Germania stiano attuando politiche di distanziamento sociale restrittive, lascia supporre che nemmeno loro siano troppo convinti del basso tasso di mortalità che emergerebbe dai dati ufficiali e che, anzi, stiano adottando una politica di screening a tappeto ben più consistente di quanto fatto dalle nostre parti. Se così fosse, l'apparente bassa mortalità potrebbe essere semplicemente condizionata dal ritardo con cui la dinamica dei decessi segue quella dei contagi e nei prossimi giorni potremmo assistere all'allineamento di questo dato con quello delle altre nazioni. Resta comunque il dubbio in merito ai criteri di classificazione dei decessi, visto che non tutte le nazioni associano le morti al virus, in presenza di quadri clinici già compromessi.

Da sole, queste estrapolazioni non ci stanno aiutando molto: è necessario provare a stimare i contagiati odierni in modo da poterli confrontare con il dato dei decessi per capire davvero se i nostri ragionamenti sono sulla strada giusta e, se sì, a che livello di diffusione siamo realmente arrivati. Come si evolve, dunque, questo maledetto virus ? Proviamo a costruire un modello basandoci sugli unici elementi che, arbitrariamente, consideriamo attendibili. Nel nostro caso sono 3:

- L'istante t_0 nel quale la propagazione ha avuto inizio. Nel nostro caso (Italia) prenderemo a riferimento la data del primo contagio rilevato in Italia, registrato il 6 febbraio, a cui, per coerenza con quanto fatto in precedenza, sottrarrò i 6 giorni relativi al periodo contagio-primi sintomi. Per semplicità, assumiamo il 1° febbraio come il t_0 italiano.
- L'istante t_{rif} nel quale la propagazione ha raggiunto un determinato valore. È qui che potremo sfruttare i risultati ricavati in precedenza, disponendo di un'ampia gamma di valori.
- Il tipo di propagazione che caratterizza la natura del processo. Poiché i contagiati non aumentano di un numero fisso per unità di tempo (giorni), bensì di un numero che è proporzionale alla grandezza raggiunta fino a quel momento, stiamo pertanto parlando di *crescita esponenziale*.

Adesso, non potendo contare su rimembranze universitarie ormai troppo lontane nel tempo, ho provato inizialmente a fare le cose semplici limitandomi ad ipotizzare che la funzione che rappresenta la propagazione del virus sia rappresentata da x^t dove x sarà il nostro “fattore di crescita”, inteso come l'aumento giornaliero di contagiati rispetto al giorno precedente, da determinare, e t il tempo, espresso in giorni. Se non fosse già chiaro, va precisato che tale formula può essere presa in considerazione esclusivamente nell'ipotesi di diffusione libera del contagio e quindi cesserà di valere al momento del lockdown. Vi risparmio i passaggi anticipando il risultato: il modello esponenziale non è adeguatamente rappresentativo dell'andamento dei dati calcolati a partire dal nostro tasso di mortalità. Il motivo è facilmente comprensibile: in una comunità chiusa, come è schematicamente rappresentabile una nazione, esiste una soglia, relativamente vicina, oltre la quale la crescita non è più “libera” in quanto la diffusione del virus inizia a risentire della presenza dei soggetti già contagiati, rallentando così la propria crescita. Per completezza, riporto comunque la soluzione individuata che meglio approssima i dati:

$$91(1,36)^t .$$

Procedo quindi alla ricerca di un modello leggermente più fine, quello logistico, che tiene in considerazione la presenza di un limite alla crescita del contagio. Nemmeno in questo caso, lo anticipo, sono riuscito ad approssimare adeguatamente la dinamica di crescita del virus nel primo mese “libero”; prendo atto dei miei limiti matematici e abbandono pertanto il sogno di ottenere il modello perfetto. Retrocedo quindi l’obiettivo cercando di ottenere almeno una forbice sui cui far ricadere i dati. Passo subito al risultato: con il modello logistico, i parametri che meglio sembrano approssimare i dati sono i seguenti:

$$\frac{ab}{a + (b - a)e^{-kt}}$$

Dove:

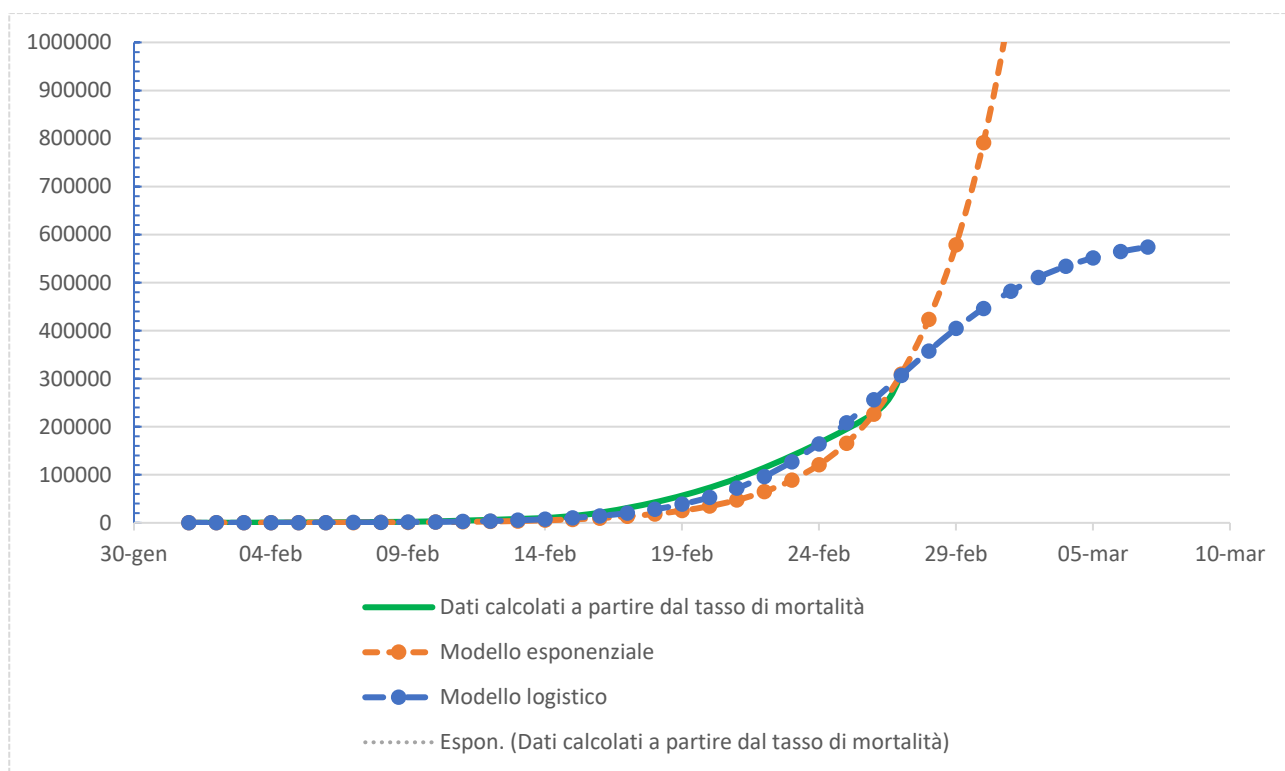
$$a = 91$$

$$b = 600.000$$

$$k = 0,34$$

$$t = \text{giorni a partire dal 1/2/2020}$$

L’ultimo step consiste nel sovrapporre le curve ricavate con i due modelli ai dati calcolati sperimentalmente, per vedere di quanto ci stiamo sbagliando.



Per il momento concludo, accontentandomi del rapporto semplicità/bontà dei dati ricavati. Ok, adesso abbiamo i modelli e, benché abbiamo capito che non rappresentino fedelmente la realtà, possiamo comunque provare ad utilizzarli per trarre almeno una conclusione: i parametri utilizzati nel modello logistico sono stati inseriti empiricamente con l’obiettivo di ottenere l’estremo inferiore della forbice che contiene il dato relativo ad oggi. Affidandoci a questa ipotesi, osserviamo alla data di ieri, 19 marzo, la logistica ha già raggiunto il 99,9% del suo asintoto che, ricordo, è pari a **600.000**. Osservando l’andamento delle curve, la mia personalissima sensazione è che tale numero sia comunque abbastanza inferiore a quello reale. Per avere un’idea di ciò di cui stiamo parlando, secondo il modello esponenziale tutta la nazione sarebbe già stata contagiata a partire dal 12 marzo.

Dando quindi per buono il risultato ricavato, tra 20 giorni potremmo registrare un numero totale di decessi pari ad almeno 6.600. Purtroppo, però, stiamo ipotizzando di trovarci sulla parte inferiore della forbice e non stiamo modificando il tasso di mortalità che, come abbiamo visto, dovrebbe essere portato almeno al 2%, in considerazione del fatto che non ci troveremo più nelle condizioni standard. Lasciando la forbice bassa e applicando il 2%, i decessi salirebbero a 12.000 e, ripeto, gli effetti del lockdown saranno successivispero davvero di sbagliarmi.

Ipotizzando coerenti e sensate tutte le considerazioni sin qui effettuate (mi rendo conto che è proprio questa l'ipotesi più azzardata e spero pure lo sia) mi verrebbe da trarre due immediate conclusioni:

- Notizia “buona”: il tasso di mortalità di questo maledetto virus, in condizioni “standard”, potrebbe essere più basso di quello che sembra leggendo i bollettini che quotidianamente vengono emessi.
- Notizia “cattiva” e con impatto, purtroppo, maggiore di quella “buona”: la velocità di propagazione del virus è letteralmente devastante.

Nei prossimi giorni continuerò a monitorare i dati, come tutti, cercando di capire se, come e quando terminerà mai questo incubo, nel frattempo.. #stiamoacasa!