

# Considerazioni sull'andamento temporale della diffusione di COVID-19 Italia basate su un esercizio di regressione logistica

Bruno CAUDANA -- Adaptive Systems  
b.caudana@adaptive.it  
<http://www.adaptive.it/covid19>

ver. 1 – 15-mar-2020  
rev. 5 – 19-may-2020  
<work in progress>



## Abstract

Si è costruito un semplice modello parametrico basato sulla **funzione logistica** per tentare di descrivere l'andamento temporale dell'epidemia COVID-19 in Italia, usando i dati ufficiali del *Ministero della Salute* italiano. Il modello evidenzia anomalie nell'indice di contagio, rispetto a quanto ci si potrebbe aspettare da un contagio che si diffonde così velocemente in una popolazione presunta vergine al virus. Interrogativi sulla velocità di contagio e sulle stime di  $R_0$  vengono evidenziati ed esposti. Considerazioni vengono espone sulla base delle evidenze del fitting parametrico. Dubbi vengono esposti sulla efficacia delle azioni di contenimento della diffusione del virus a causa della velocità di contagio, almeno per Lombardia e zone limitrofe. L'epidemia potrebbe esaurirsi prima che le azioni di contenimento alla diffusione del virus possano diventare efficaci. Cautela va comunque usata nel fidarsi di modelli matematici di fenomeni molto complessi e abbastanza ignoti.

# Indice

<b>1</b>	<b>Modello logistico (<i>sigmoide</i>)</b>	<b>4</b>
1.1	Cenni teorici . . . . .	6
1.2	Implementazione del modello logistico . . . . .	9
1.3	Interpretazione fisico-epidemica del modello logistico	10
<b>2</b>	<b>Dati COVID-19 utilizzati</b>	<b>15</b>
2.1	Definizione <i>Lombardia+province limitrofe</i> . . . . .	15
<b>3</b>	<b>Stime parametriche</b>	<b>16</b>
3.1	Caso Lombardia estesa con $i_c$ ottimale . . . . .	17
<b>4</b>	<b>Analisi di stabilità del modello</b>	<b>18</b>
<b>5</b>	<b>Analisi dell'effetto di <i>lockdown</i></b>	<b>19</b>
<b>6</b>	<b>Considerazioni</b>	<b>22</b>
6.1	Se modello inadeguato . . . . .	22
6.2	Se modello plausibile . . . . .	22
6.2.1	Campione deviato . . . . .	22
6.2.2	Stress sul sistema di terapia . . . . .	23
6.2.3	Anomalia tra $i_c$ e $R_0$ . . . . .	24
6.2.4	Stima di $R_0$ dalla curva epidemica . . . . .	26
6.2.5	Considerazioni su $i_c$ e $R_0$ di COVID-19 . . . . .	36

6.2.6	Protezione Immunitaria Pregressa? . . . . .	38
6.2.7	Il <i>lockdown</i> funziona? . . . . .	41
6.3	Nota evoluzionistica . . . . .	44
6.4	Nota su mortalità e indice di letalità COVID-19 . . .	46
<b>7</b>	<b>Bibliografia Cautele e Ringraziamenti</b>	<b>50</b>

# 1 Modello logistico (*sigmoide*)

Il modello logistico (o *curva logistica*, o *sigmoide*) può essere usa-

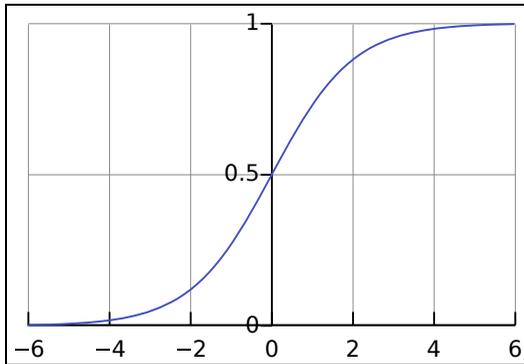


Figura 1: Curva *sigmoide*

to per descrivere fenomeni diffusivi/epidemici in un sistema chiuso dove si innescano fenomeni di saturazione (es: sopavveniente resistenza immunitaria nota o ignota, fenomeni o comportamenti che ostacolano la diffusione, ec.). Difficile è individuare il grado di chiusura del sistema da descrivere mediante modello logistico, cioè quale sia il fattore di normalizzazione da adottare nel modello.

[https://it.wikipedia.org/wiki/Equazione\\_logistica](https://it.wikipedia.org/wiki/Equazione_logistica)

La prima apparizione della funzione logistica risale a Pierre François Verhulst (1838-1847) in relazione a studi sulle crescite di popolazione.

[https://it.wikipedia.org/wiki/Pierre\\_Fran%C3%A7ois\\_Verhulst](https://it.wikipedia.org/wiki/Pierre_Fran%C3%A7ois_Verhulst)

[https://en.wikipedia.org/wiki/Pierre\\_Fran%C3%A7ois\\_Verhulst](https://en.wikipedia.org/wiki/Pierre_Fran%C3%A7ois_Verhulst)

Vito Volterra (circa 1926) nei suoi studi sulle dinamiche di popolazione definì equazioni differenziali di questa famiglia, la più semplice delle quali genera la funzione logistica.

[https://it.wikipedia.org/wiki/Vito\\_Volterra](https://it.wikipedia.org/wiki/Vito_Volterra)

[https://it.wikipedia.org/wiki/Equazioni\\_di\\_Lotka-Volterra](https://it.wikipedia.org/wiki/Equazioni_di_Lotka-Volterra)

Il comportamento a sigmoide fu usato da Enrico Fermi, tra altri, nell'ambito di fenomeni di diffusione (statistica di Fermi-Dirac).

[https://en.wikipedia.org/wiki/Fermi%E2%80%93Dirac\\_statistics](https://en.wikipedia.org/wiki/Fermi%E2%80%93Dirac_statistics)

[https://it.qwe.wiki/wiki/Logistic\\_function](https://it.qwe.wiki/wiki/Logistic_function)

[https://it.qwe.wiki/wiki/Fermi%E2%80%93Dirac\\_statistics](https://it.qwe.wiki/wiki/Fermi%E2%80%93Dirac_statistics)

Negli algoritmi *ANN* (*Artificial Neural Networks*) si usa la funzione logistica (diffusione in ambiente chiuso che satura) per rappresentare il segnale di output sull'assone simulato di ogni singolo neurone simulato dalla rete *ANN*, in funzione del segnale di input ricavato integrando i vari segnali provenienti dalle sinapsi simulate che convogliano al neurone simulato segnale proporzionale al loro peso sinaptico (schematicamente corrispondente alla capacità di una sinapsi biologica di modulare chimicamente il segnale che passa da un neurone all'altro in una rete neurale biologica).

[https://en.wikipedia.org/wiki/Artificial\\_neural\\_network](https://en.wikipedia.org/wiki/Artificial_neural_network)

Una panoramica storico-teorica sulle sigmoidi si trova in *J.S. Cra-mer*, "*The Origins of Logistic Regression*".

<https://papers.tinbergen.nl/02119.pdf>

Per una trattazione esauriente delle logistiche e relative regressioni si può consultare:

[https://it.wikipedia.org/wiki/Equazione\\_logistica](https://it.wikipedia.org/wiki/Equazione_logistica)

[https://en.wikipedia.org/wiki/Logistic\\_regression](https://en.wikipedia.org/wiki/Logistic_regression)

[https://en.wikipedia.org/wiki/Logistic\\_function](https://en.wikipedia.org/wiki/Logistic_function)

[https://en.wikipedia.org/wiki/Population\\_model](https://en.wikipedia.org/wiki/Population_model)

Esistono modelli epidemici molto più accurati e sofisticati, ma per essi servirebbero dati molto più accurati e molto più tempo per essere messi a punto e calibrati numericamente di quanto la velocità di COVID-19 consenta per ottenere qualche informazione in tempo utile, prima che l'epidemia si estingua autonomamente.

[https://en.wikipedia.org/wiki/Basic\\_reproduction\\_number](https://en.wikipedia.org/wiki/Basic_reproduction_number)

[https://en.wikipedia.org/wiki/Compartmental\\_models\\_in\\_epidemiology](https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology)

## 1.1 Cenni teorici

Nella fig.1 a pag.4, la variabile dipendente normalizzata  $y = f(x)$  (es: contagio cumulato espresso tra  $0 \div 1$ , cioè:  $n/\text{contagi totali}$ ) rispetto alla variabile indipendente (es: *tempo*, nel caso di stime di andamento temporale), assume la seguente forma:

$$y = f(x) = \frac{e^x}{e^x + 1} = \frac{1}{1 + e^{-x}} \quad (1)$$

Questa forma emerge come soluzione della seguente equazione differenziale di P.F. Verhulst:

$$\dot{W}(t) = \beta W(t)(\Omega - W(t)) \quad (2)$$

la quale descrive la crescita esponenziale di un fenomeno secondo  $\beta$  che nel suo crescere incontra però una resistenza crescente dovuta al limite di saturazione  $\Omega$  dell'ambiente in cui cresce, *man mano che cresce, come effetto del suo crescere*. Cioè un ambiente chiuso e non illimitato. Questa equazione differenziale è ben illustrata nel paper di J.S. Cramer, "*The Origins of Logistic Regression*" citato in apertura.

Riferendosi alla fig.1 e alla funzione  $f(x)$  (1), risolvente dell'equazione differenziale (2), si vede che la *sigmoide*  $f(x)$  è centrata sul punto di flesso ( $x = 0, f(x) = 0.5$ ). Questo punto di flesso rappresenta il punto di picco della derivata prima della *sigmoide* (una *simil-gaussiana*). Questo rappresenta il punto di picco dei nuovi contagiati, nel caso di questo esercizio. Il parametro  $c_t$  usato qui corrisponde al parametro  $\Omega$  nell'equazione differenziale generatrice.

Il *modello logistico* può essere linearizzato trasformando le misure della variabile dipendente normalizzata  $y$  nel seguente modo, tramite la funzione inversa di  $f(x)$ :

$$x = \ln \left( \frac{f(x)}{1 - f(x)} \right) = \ln \left( \frac{y}{1 - y} \right) \quad (3)$$

e poi scalando  $x$  in modo che si conformi ai dati del problema.

Nel caso di proporzionalità semplice (parametro  $\beta$  nell'equazione differenziale generatrice 2), si può ottenere questo assoggettando  $x$  ad una regressione lineare (*linear fitting*) rispetto ai punti  $(t, x)$ , cioè  $(t, \ln(y/(1 - y)))$ :

$$x = \ln \left( \frac{y}{1 - y} \right) = \beta(t) = \beta_0 + \beta_1 t \quad (4)$$

dove:

$t$  è il tempo dove si sono misurati gli  $y$  dell'insieme dei punti  $(t, \ln(y/(1 - y)))$ . Nel caso qui trattato sono *giorni dal 24-feb-2020*.

$y$  è la variabile dipendente dal tempo  $t$ , in questo caso espressa come conteggio normalizzato sulla popolazione incognita  $c_t$  dei contagiati con manifestazione di *patologia macroscopica*. Nel caso qui trattato sono i *casi totali* cumulati di contagio (*contagiati + guariti + morti*) normalizzati sui contagi totali  $c_t$  nella popolazione di riferimento  $p_{rif}$  a partire da  $t_0$ , fatto coincidere con il *24-feb-2020*, coerentemente con l'insieme di dati ufficiali del contagio COVID-19 disponibili per l'Italia al largo pubblico.

$\beta_0, \beta_1$  sono rispettivamente l'intercetta e il coefficiente angolare della retta di regressione.

[https://it.wikipedia.org/wiki/Regressione\\_lineare](https://it.wikipedia.org/wiki/Regressione_lineare)

Tutto ciò consente di esprimere nel modo seguente la variabile dipendente della curva logistica scalata nei termini delle misure originali normalizzate:

$$f(t) = \frac{1}{1 + e^{-\beta(t)}} \quad (5)$$

con  $\beta(t) = \beta_0 + \beta_1 t$

A partire da fine marzo inizio aprile, l'andamento dell'epidemia ha mostrato che il fattore di proporzionalità semplice insito nel fattore  $\beta$  della (2) – e perciò una regressione lineare ipotizzata come  $\beta(t)$ – non coglie più lo sviluppo dell'epidemia. In altre parole, l'ipotesi di semplice proporzionalità assunta dall'equazione differenziale generatrice (2) non rende ragione completamente dell'andamento dell'epidemia.

Osservata la curva dopo la applicazione della trasformata  $\ln(y/(1-y))$ , si è ipotizzato per  $\beta(t)$  un polinomio:

$$\beta(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_4 t^4 + \beta_5 t^5 \quad (6)$$

dove:

$\beta_0 \dots \beta_5$  sono i parametri del polinomio di quinto grado a rappresentare l'andamento a S asimmetrico della trasformata.

[https://it.wikipedia.org/wiki/Regressione\\_lineare](https://it.wikipedia.org/wiki/Regressione_lineare)

[https://it.wikipedia.org/wiki/Interpolazione\\_polinomiale](https://it.wikipedia.org/wiki/Interpolazione_polinomiale)

[https://it.wikipedia.org/wiki/Regressione\\_nonlineare](https://it.wikipedia.org/wiki/Regressione_nonlineare)

La de-normalizzazione per ottenere la stima dei contagiati nelle unità di misura originali si fa banalmente moltiplicando  $c_t \cdot f(t_k)$ .

Il *modello logistico* perciò dipende dal parametro libero dei contagi totali  $c_t$ , nel caso qui trattato, il quale rappresenta il confine del sistema chiuso (più realisticamente: *semi-chiuso*, nella dinamica spazio-temporale che genera progressivamente i dati oggetto di modellazione) descritto dalla *sigmoide* e dalla associata equazione differenziale. A sua volta,  $c_t$  può essere fatto dipendere dall'indice di contagio  $i_c$  agente su una popolazione di riferimento (popolazione su cui si ipotizza sia stato estratto il campione di casi), e dalla dimensione della popolazione  $p_{rif}$  su cui si ipotizza insistere l'indice di contagio  $i_c$ . Cioè  $c_t = i_c \cdot p_{rif}$ .

Le estrapolazioni così computabili sono ovviamente soggette a tutte le incertezze e cautele tipiche di un fitting lineare o più complesso, alla adeguatezza del modello rispetto al fenomeno in esame, alla qualità dei dati di input e alla formulazione del campione da cui provengono.

In sintesi le equazioni principali del modello sono le seguenti:

Casi totali cumulati (oggetto di fitting non lineare):

$$f(t) = c_t^* \frac{1}{1 + e^{-(\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5)}} \quad (7)$$

Casi giornalieri (derivata prima di  $f(t)$ ), derivata analiticamente:

$$f'(t) = c_t^* \frac{(\beta_1 + 2\beta_2 t + 3\beta_3 t^2 + 4\beta_4 t^3 + 5\beta_5 t^4) \cdot e^{-(\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5)}}{(1 + e^{-(\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5)})^2} \quad (8)$$

## 1.2 Implementazione del modello logistico

La implementazione del *modello logistico* calcola la trasformazione del numero dei contagi cumulati totali rilevati  $n$  (i "casi totali" della definizione dati ministeriali), normalizzati rispetto alla popolazione contagiata incognita  $c_t$ , con la formula  $\ln(y/(1-y))$ , dove  $y = n_r/c_t$ . Poi procede a trovare la curva di regressione  $\beta(t)$  sull'insieme di punti  $(t, x)$ , cioè  $(t, \ln(y/(1-y)))$ . I parametri  $\beta_0, \beta_1$  così ricavati vengono poi utilizzati per calcolare le stime di contagio sulla base dei seguenti 2 parametri:

1. popolazione  $p_{rif}$  a cui il conteggio dei casi intende riferirsi, che può essere oggetto di ipotesi o stimato dal modello.
2. indice di contagio  $i_c$ , cioè (*contagiati/popolazione*), che può essere oggetto di ipotesi o stimato dal modello.

La stima ottimale di  $i_c$ , oppure di  $p_{rif}$ , è fatta dal calcolo mediante la minimizzazione di  $err_m$ , definito come il quadrato delle differenze tra i  $k$  conteggi dei *casi giornalieri* rilevati  $n$  e gli associati conteggi  $\hat{z} = f'(t)$  (tutti in scala logaritmica) stimati dalla *regressione logistica*:

$$err_m = \sqrt{\frac{1}{N} \sum_{k=1}^N (\ln(n_k) - \ln(\hat{z}_k))^2} \quad , \quad \sigma_{min} = \min(err_m) \quad (9)$$

col vincolo  $\hat{z}_k > 0$  per ogni conteggio ricalcolato dal modello.

A fine epidemia/diffusione, il temine  $\Omega$  dell'equazione differenziale (il  $c_t$  di questo modello) è perfettamente stabile, se il sistema è chiuso. Perciò la sua stima ottimale durante la diffusione tende ad approssimarsi al valore finale man mano che la diffusione procede.

La stima di  $i_c^*$  mediante minimizzazione di  $err_m$  è fatta con algoritmo evolutivistico DEPS (*Differential Evolution and Particle Swarm Optimization*).

[https://en.wikipedia.org/wiki/Differential\\_evolution](https://en.wikipedia.org/wiki/Differential_evolution)

[https://en.wikipedia.org/wiki/Particle\\_swarm\\_optimization](https://en.wikipedia.org/wiki/Particle_swarm_optimization)

[https://en.wikipedia.org/wiki/Evolutionary\\_computation](https://en.wikipedia.org/wiki/Evolutionary_computation)

Il processo descritto qui produce tutti i grafici e le stime quantitative riportate in questo documento.

### 1.3 Interpretazione fisico-epidemic del modello logistico

In una popolazione di ospiti che un patogeno parassita (virus, batteri, altro organismo o reazione chimica autocatalitica replicante, ec.) contagia, quanti si ammalano di *patologia macroscopica* durante una epidemia dipende da:

1. quanti sono suscettibili al patogeno (*suscettibilità iniziale*): non si sa, e non si può determinare direttamente;

2. quanti restano suscettibili al patogeno man mano che l'epidemia si propaga: non si sa, dipende dalla *suscettibilità iniziale* e da quanti progressivamente ne diventano immuni;
3. quanti si infettano perché incontrano il patogeno e quanti non;
4. quanti di questi ultimi hanno bisogno di terapie tali da essere intercettati dal sistema terapeutico;

**L'andamento temporale dei casi cumulati di *patologia macroscopica* contiene in modo aggregato questa informazione incognita.** Non è semplice disaggregare questa informazione. Forse impossibile. Ma non si può dire che chi **non** ha sviluppato anticorpi protettivi è automaticamente suscettibile ad infettarsi. Questa è una semplificazione prudenziale, ma troppo conservativa. Implica assumere arbitrariamente che tutta una popolazione sia interamente suscettibile al virus. Di fatto non si sa quale è la suscettibilità iniziale di una popolazione ad un pur nuovo virus. Potrebbero agire fattori di protezione genetica o indotta da pregressa storia immunitaria soggettiva; per esempio: questo virus potrebbe avere bersagli biochimici comuni ad altri Corona Virus già riconosciuti dallo spettro immunitario della popolazione ospite. Inoltre, storicamente, non si è mai visto una epidemia infettare una popolazione completa.

Quanti muoiono o guariscono dipende da:

- gravità della patologia indotta dalla infezione (un Corona Virus del raffreddore non è grave come Ebola);
- soggettivi stato di salute generale e reazione immunitaria di chi si infetta;
- tecniche di cura in evoluzione;

Questo modello non affronta la problematica di stimare malati o morti. Affronta solo l'analisi dell'andamento temporale dei casi rilevati di *patologia macroscopica*.

Semplificando all'osso, una epidemia in evoluzione libera ha una fase di accensione, raggiunge un picco, infine si spegne in modo più o meno simmetrico a come è iniziata. Questa è la cornice del modello matematico adottato.

Il modello qui formulato è una estensione della *equazione differenziale di Verhulst*:  $\dot{W}(t) = \beta W(t)(\Omega - W(t))$  utilizzata per descrivere un processo di diffusione in un ambiente chiuso, dove il progredire della diffusione riduce progressivamente lo spazio di diffusione.

La *equazione differenziale di Verhulst* (2) viene adottata come cornice a rappresentare un fenomeno di diffusione in ambiente a risorse limitate, quale una epidemia è.

L'estensione si ispira all'approccio adottato nella statistica di Fermi-Dirac, dove il fattore di proporzionalità  $\beta$  è sostituito da una funzione più complessa  $\beta(t)$ .

L'estensione si è resa necessaria perché il criterio di proporzionalità semplice della *equazione differenziale di Verhulst*, dopo un periodo iniziale ha smesso di rappresentare correttamente i dati misurati. I motivi possono essere molteplici: modifiche ai criteri di diagnosi, aumento delle diagnosi, ripetizioni di test su stessi soggetti calcolati come nuovi soggetti, modifiche nel comportamento della popolazione; dinamiche non lineari di espansione territoriale, ec. Ad ora, la funzione ipotizzata è un polinomio di terzo grado per rappresentare l'oscillazione osservata sui dati trasformati secondo *inversa della Funzione Logistica*.

L'uso di un polinomio è un espediente che coglie la correlazione del caso, ma non dice nulla sulle cause che producono quell'andamento. Con ulteriori analisi con anche altri insiemi di dati di realtà territoriali diverse forse potrà emergere qualche regolarità più chiara.

Passare da 2 a 4 parametri liberi per  $\beta(t)$  non dovrebbe intro-

durre causare troppi problemi di overfitting.

Il parametro  $i_c^*$  equivale al parametro  $\Omega$  della (2) e rappresenta il confine del sistema chiuso. Col procedere dell'epidemia ho visto che esso tende a crescere. Questo ha la sua ragion d'essere nel fatto che l'epidemia nasce in piccoli cluster e si espande territorialmente a coinvolgere sempre più popolazione fino alla popolazione totale di area di riferimento di cui si assume costante la popolazione. Inoltre, presumibilmente tende ad attenuare la sua crescita man mano che i clusters di epidemia tendono a espandersi in zone a più bassa densità territoriale di popolazione, cioè in aree dove la probabilità di interazione interpersonale è meno densa. Potrebbe essere sensato tentare di correlare questo andamento del parametro  $i_c^*$  ad un andamento secondo  $\sqrt{t}$ .

Studiando in modo comparato altri insiemi di dati COVID-19 di altre aree, e studiando la capacità predittiva del modello man mano che l'epidemia evolve, altre funzioni più generali e significative dal punto di vista fisico-epidemico possono essere scovate con l'ausilio di fitting simbolico e altre tecniche che facciano un uso più profondo di algoritmi evolutivisti e altro (tecniche di *Deep Learning*, per dirla con una locuzione di moda).

[https://en.wikipedia.org/wiki/Evolutionary\\_algorithm](https://en.wikipedia.org/wiki/Evolutionary_algorithm)

L'equazione differenziale più generale di Verhulst 10, menzionata dal paper di J.S. Cramer, che contiene la componente funzionale  $\phi(W(t))$ , al posto della semplificazione introdotta da  $\Omega$ , forse descriverebbe in modo più dettagliato il fenomeno di diffusione con saturazione in un sistema chiudentesi o semi-chiuso (es: fenomeno epidemico).

La funzione  $\phi(W(t))$  svincola l'equazione dalla proporzionalità diretta a  $\beta$ , imposta dal confine rigido  $\Omega$ .

$$\dot{W}(t) = \beta W(t) - \phi(W(t)) \quad (10)$$

Ma  $\phi(W(t))$  contiene praticamente quasi qualsiasi cosa di un fenomeno epidemico/diffusivo, compreso il *learning* che si sviluppa col

procedere della diffusione epidemica – sia sui comportamenti consci, semi-consci e inconsci; sia sui meccanismi biochimici e immunitari sottostanti. Difficile, comunque, sarebbe la sua interpretazione fisica e la sua determinazione sperimentale affidabile. Questo per quanto mi riesce di capire.

Forse una accuratezza maggiore si ottiene sostituendo  $\omega(t)$  al confine rigido  $\Omega$ , ad indicare che la diffusione tende a gonfiare il confine incognito  $\Omega$  col procedere dell'epidemia (per esempio linearmente rispetto a  $t$ , in prima approssimazione), senza tentare di rendere ragione degli effetti a *loop chiuso* di  $W(t)$  insiti in  $\phi(W(t))$ . Si può tentare. Proverò.

<todo: automatizzare fitting con altri insiemi di andamenti temporali COVID-19 e scovare funzioni comuni più generali>

## 2 Dati COVID-19 utilizzati

In questo esercizio di regressione vengono utilizzati i dati ufficiali del Ministero della Salute italiano e della Protezione Civile della grandezza *casi totali* o *Totale positivi* ( $n$ ) come definita dal Ministero (*contagiati + guariti + morti*).

<https://github.com/pcm-dpc/COVID-19/tree/master/dati-province>

<http://www.salute.gov.it/portale/nuovocoronavirus/dettaglioContenutiNuovoCoronavirus.jsp>

<http://opendatadpc.maps.arcgis.com/apps/opsdashboard/index.html#/b0c68bce2cce478eaac82f>

Una modellazione più accurata risulterebbe dall'applicazione di *regressioni logistiche*, o equivalenti, nello spazio, oltre che nel tempo, in modo da vedere, oltre la dinamica temporale, anche la dinamica spaziale. Questo fanno modelli epidemiologici più sofisticati. Ma questo, per non generare overfitting e avere qualche affidabilità, richiederebbe dati in quantità e accuratezza tali che non sono a mia disposizione.

### 2.1 Definizione *Lombardia+province limitrofe*

*Lombardia, Modena, Parma, Piacenza, Reggio, Rimini, Pesaro, Alessandria, Asti, Novara, Verbanò, Vercelli, Padova, Treviso, Venezia.*

Totale popolazione assunta come dal modello per "*Lombardia+province limitrofe*":  $p_{rif} \approx 17e6$  (17\_mln di abitanti). Totale popolazione assunta come dal modello per "*Italia*":  $p_{rif} \approx 60e6$  (60\_mln di abitanti).

### 3 Stime parametriche

Col modello descritto è possibile fare ipotesi variando due parametri del modello di calcolo:

1. popolazione  $p_{rif}$  a cui si ipotizza sia riferito l'insieme di dati utilizzati
2. indice di contagio  $i_c$ , cioè *contagiati/popolazione* ( $c_t/p_{rif}$ ).

Due tipi grafici caratterizzano il fitting:

1. i dati e la curva della stima dell'evoluzione epidemica cumulata e la sua derivata prima (simil-gaussiana)
2. il grafico dati (trasformati secondo l'inversa della curva logistica) e la proiezione del modello, da cui si evince facilmente la mediana del contagio, quando la retta di regressione incrocia l'asse  $t$ , e l'estinzione come punto simmetrico all'origine dei dati, nonché altre informazioni.

### 3.1 Caso Lombardia estesa con $i_c$ ottimale

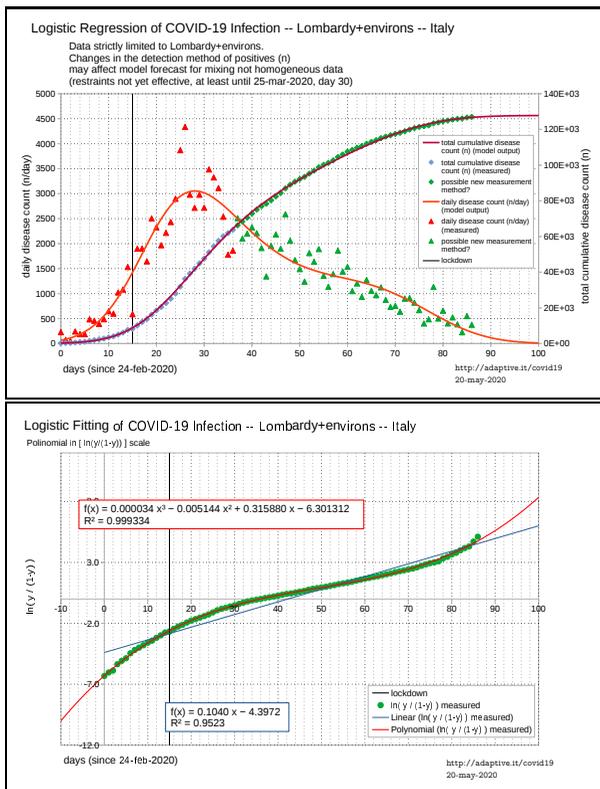


Figura 2: Fitting della linearizzazione della logistica: Lombardia+province limitrofe: 17\_mln ab ( $p_{rif} = 17 \cdot 10^6$ ,  $i_c^* = 0.00752$ ,  $i_c^*$  con tendenza a crescere lievemente nel tempo, di fitting in fitting, coerentemente con il fatto che i dati misurano nel tempo un'epidemia che si espande territorialmente, cioè un sistema semi-chiuso e non perfettamente chiuso, almeno nelle fasi iniziali di una pandemia)

## 4 Analisi di stabilità del modello

Uno dei pregi più importanti di un modello è la sua capacità predittiva. In altre parole la sua capacità di estrapolare correttamente al di fuori dell'insieme di dati utilizzati per stimare i parametri liberi che gli consentono di adattarsi ai dati rilevati. Una corretta previsione è indice di una ipoteticamente corretta cattura di regolarità intrinseche del fenomeno misurato.

Ne caso di questo modello si può valutare la capacità predittiva osservando gli errori che il modello compie nella estrapolazione . Questa indicazione consente anche di cercare quei parametri e quelle funzioni che minimizzano l'errore di estrapolazione. Si usano cioè i dati futuri per valutare/correggere i parametri e le funzioni del modello fino a un certo tempo che stimano il comportamento futuro del modello.

<todo>

Tendenza della stima ottimale di  $i_{\{c\}}$

## 5 Analisi dell'effetto di *lockdown*

Il grafico di fig.3 a pag.19 mostra che il *lockdown* della popolazione non ha avuto effetto misurabile almeno fino al 31-mar-2020, data in cui probabilmente si sono cambiati metodi, criteri e strumenti per contare i positivi. Questo almeno per la zona lombarda.

Il grafico in fig.3 è molto sensibile alle variazioni di andamento. Appena qualcosa si muove rispetto alla evoluzione libera dell'epidemia, si vede subito. Se il *lockdown* avesse avuto effetto, si sarebbe

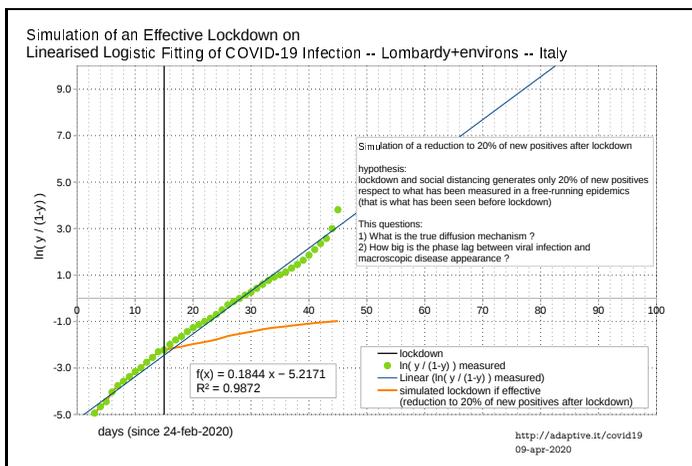


Figura 3: Questo grafico mostra l'efficacia ipotetica di un *lockdown* che riduca al 20% l'insorgenza di nuovi contagi, rispetto a quanto misurato, su una epidemia che si espande naturalmente senza modifiche nel comportamento degli infettati/infettanti, cioè come si evidenzia **prima** del *lockdown*. Il grafico mostra l'efficacia ipotetica di un *lockdown* che abbia azione immediata. In realtà un *lockdown* avrà una efficacia differita nel tempo per un tempo di latenza proprio dell'infezione virale, stimata in 15 giorni.

vista la sequenza dei punti verdi misurati deviare verso destra a determinare un coefficiente di regressione della retta più basso, come esemplificato dalla simulazione. Così non appare.

La inefficacia del *lockdown* rivelato dalla curva indica: o un

notevole ritardo di fase tra l'infezione virale e la manifestazione della patologia macroscopica; oppure meccanismi di infezione sconosciuti o non ipotizzati che passano le maglie del *lockdown*, come l'acqua passa in un colabrodo; oppure una popolazione che non si infetta per pregresse protezioni immunitarie ignote o non ipotizzate; oppure una popolazione che si è infettata e non se ne è accorta; oppure una miscela di queste cose.

Può darsi che le azioni di contenimento abbiano funzionato dove l'epidemia era in fase più iniziale. Si potrà vedere dalla diversa inclinazione delle regressioni dei dati trasformati secondo inversa della *logistica*:  $(t, \ln(y/(1 - y)))$  fatta per aree diverse.

In altre parole, **pare un po' arbitraria e senza evidenza misurabile l'affermazione secondo cui le misure di *lockdown* hanno contenuto l'infezione**. Avrebbero contenuto l'infezione rispetto a quale riferimento? L'epidemia invece sembra seguire una evoluzione naturale. Si potrà capire meglio con confronti più ampi.

<todo>

<: ulteriori regressioni mobili e confronti

-- ampliamento con grafici, regressioni mobili ec.>

## **Efficacia del *lockdown*?**

Altri elementi fanno dubitare dell'efficacia del lockdown. In nazioni diverse, il *lockdown* è stato declinato in vari modi: da poco o nessun *lockdown* a massimo *lockdown* possibile. Al momento non si notano correlazioni evidenti di minor contagio a fronte di maggior *lockdown*. Anzi.

Per esempio: il Brasile evidenzia un indice di contagio (nessun *lockdown*, 1 su 10\_mila) di 50 volte inferiore rispetto alla zona lombarda (massimo *lockdown*, 50 su 10\_mila).

Segno che qualcosa di fondamentale non si è capito della dinamica epidemica di questo virus. La cosa merita seri approfondimenti.

Dubbi sorgono circa l'impostazione generale dell'azione anti-epidemia: *l'assioma di voler/poter impedire il contagio ad ogni costo.*

## 6 Considerazioni

Sulla base delle stime parametriche con il *modello logistico* proposto, si possono cominciare ad articolare alcune considerazioni sulla base dei dati alla data del 13-mar-2020.

### 6.1 Se modello inadeguato

Il *modello logistico* adottato qui può non essere adeguato a descrivere aspetti dell'andamento temporale del fenomeno epidemico COVID-19, nonostante un alto coefficiente di correlazione della *sigmoide* con i dati disponibili ( $R^2 = 0.98 \div 0.99$ ). Il *modello logistico* descrive fenomeni diffusivi in ambiente chiuso con saturazione conseguente. Incerto è il grado di chiusura e incerti sono i meccanismi di saturazione del sistema COVID-19, soprattutto nel caso di dati parziali di popolazioni parziali. Se il *modello logistico* è inadeguato, questo scritto è, al più, un piccolo esercizio di matematica e di logica.

### 6.2 Se modello plausibile

Se invece il *modello logistico* ha qualche validità descrittiva di COVID-19, allora si possono articolare alcune considerazioni per *circoscrivere* e *confinare concettualmente* le congetture fatte.

#### 6.2.1 Campione deviato

Il conteggio dei contagi sta avvenendo sulla base di tamponi fatti principalmente per orientare terapie e proteggere dal contagio le strutture sanitarie. I conteggi perciò rappresentano un campione estremamente deviato, rispetto alla popolazione  $p_{rif}$  di cui si intenderebbe stimare l'indice di presenza del virus. Di fatto i dati

disponibili sono il conteggio dei casi di *patologia macroscopica* così come sono stati intercettati dal sistema terapeutico. Per questo motivo nel modello si sono considerati  $i_c$  e  $p_{rif}$  come parametri liberi con cui fare ipotesi *what if*.

Ovviamente tamponi per individuare la presenza del virus fatti su un campione randomizzato rappresentativo della popolazione servirebbe a capire meglio la diffusione del virus e la dimensione del contagio da raffrontare a dimensione e andamento della *patologia macroscopica*.

Ma, al momento, se sono state fatte queste analisi, non si hanno notizie dei risultati. Forse una simile analisi non sembra neppure essere praticabile, in stato di emergenza sanitaria.

In assenza di informazioni dirette su dimensione, distribuzione e andamento della presenza del virus in popolazione, le uniche cose che si possono tentare di fare:

1. con i dati disponibili,
2. con modelli di regressione
3. e con ragionamenti logici e matematici,

sono di circoscrivere e confinare ipotesi e congetture sulla base della forma delle curve che fittano i dati di *patologia macroscopica*.

## 6.2.2 Stress sul sistema di terapia

I dati misurano l'estensione della *patologia macroscopica*. Non misurano l'estensione del *contagio virale*.

Per come sono stati formulati e rilevati i casi di contagio, si può ipotizzare che essi correlino bene con lo stress sulle unità di terapia di vario tipo. In fondo. I tamponi in Italia sono stati fatti più

per orientare per orientare azioni terapeutiche, che per altro. I casi sintomatici e/o accertati con tampone positivo, insieme alla forma della curva con fitting ottimale di  $i_c$ , in un certo senso selezionano quella quota di popolazione  $c_t$  intercettata dal sistema terapeutico che va a insistere sulle le unità di terapia e a stressarle.

### 6.2.3 Anomalia tra $i_c$ e $R_0$

Un'epidemia che si diffonde velocemente senza incontrare ostacolo alla diffusione perché il suo agente infettante è, in ipotesi, sconosciuto al sistema immunitario della popolazione, dovrebbe manifestare rapidamente un numero elevato di contagi nella popolazione.

Invece, la stima ottimale dell'indice di contagio  $i_c^*$  fatta dal modello dai dati di *manifestazione macroscopica* del contagio intercettati dal sistema terapeutico, insieme alla alta velocità di salita della curva, mostra un  $i_c$  dei contagi evoluti in *patologia macroscopica* molto più basso di quanto ci si potrebbe aspettare. Solo l'aver a che fare con un campione molto deviato di accertamento del contagio non sembra bastare a spiegare questa anomalia.

Bisogna aggiungere l'ipotesi che il virus abbia infettato molte più persone in modo asintomatico, o quasi asintomatico, tanto da non essere intercettato dal sistema terapeutico, e che anche asintomatici possano contagiare; e/o che parte della popolazione sia coperta da meccanismi di protezione ignoti o non ipotizzati.

Il che farebbe propendere per l'ipotesi che il denominatore degli indici di letalità e mortalità intrinseca del virus siano molto più bassi di quanto si evidenzia dai dati di *patologia macroscopica*.

**Questo ha conseguenze sulla valutazione delle strategie di azione:** se davvero il virus contagia velocemente e con efficacia alta, allora le azioni di contenimento alla diffusione rischiano di essere poco efficaci. Inoltre, in questo caso, ci dobbiamo aspettare di veder rispuntare, come funghi, nuovi focolai più avanti nel tem-

po, in funzione del tempo di decadimento delle difese immunitarie sviluppate nei contagiati.

## 6.2.4 Stima di $R_0$ dalla curva epidemica

### Interrogativi su $R_0$ di COVID-19

Le stime di  $R_0$  a me note ad oggi per COVID-19 sono:  $R_0 \approx 2.0 \div 2.5$ , (fonte: prof. Burioni e altri che ho ricavato dai media, ma senza che fosse indicata la durata di infettività assunta per quello  $R_0$ , da quanto ricordo). Purtroppo raramente si dice a quale, e quanto lungo, periodo di infettività si riferisce la stima di  $R_0$  che viene proposta.

### Definizione epidemiologica di $R_0$ e suo andamento $R_t$

$R_0$  è  $R_t$  all'istante  $t = 0$  di inizio epidemia. Ogni istante può essere considerato inizio di una epidemia.  $R_0$  assume come iniziali le condizioni di suscettibilità della popolazione a partire da quell'istante.

**definizione:**  $R_0$  (e più in generale  $R_t$ ) esprime il numero atteso di nuove infezioni generate da un singolo individuo infetto nel corso del suo intero periodo di infettività, in una popolazione interamente suscettibile omogenea, fatte costanti le altre condizioni al contorno.

In altre parole  $R_0$  dipende, in prima approssimazione, dalla infettività istantanea e dalla durata del periodo di infettività, dallo spettro di suscettibilità della popolazione, e dalla invariabilità del contesto spazio-temporale in cui avviene il processo di infezione.

In una formulazione alternativa, si può dire che  $R_0$  rappresenta il numero di infezioni secondarie causate da un singolo caso di malattia durante il suo periodo di infettività specifica considerato periodo unitario. Il che è equivalente al concetto di tasso di rendimento periodico in regime di capitalizzazione composta, dove  $R_0$  il montante di infetti prodotti dal capitale di 1 infetto.

Va sottolineato che  $R_t$  esprime la capacità infettante di un agen-

te patogeno, ad un certo istante  $t$ , calato nel modello microbiologico e comportamentale della popolazione infettabile.  $R_0$  è  $R_t$  all'istante  $t = 0$  di inizio epidemia. Per esempio: una popolazione di infettabili che ha poche relazioni interpersonali esprimerà un diverso  $R_0$  rispetto ad una popolazione che ha intensi scambi interpersonali, ad esempio frequenti interscambi commerciali, se l'infezione si propaga per qualche tipo contatto/vicinanza con un infetto.

### Idea di stimare $R_0$ ed $R_t$ dalla curva dei casi accertati

$R_0$  esprime un concetto intricato e una unità di misura complicata e difficile da determinare con accuratezza.

D'altra parte l'andamento temporale dei casi giornalieri contiene in sé  $R_0$  e l'andamento di  $R_t$ . Di fatto, **l'andamento della curva epidemica è l'effetto diretto della azione dell'indice di contagiosità  $R_t$  durante i precedenti  $g_i$  giorni di infettabilità.**

L'idea che si propone qui è quella estrarre quella informazione e di stimare  $R_0$  ed  $R_t$  direttamente dall'andamento dei *casi giornalieri* derivati dalla curva dei *casi totali cumulati*, curva che automaticamente compensa alcuni errori nel conteggio giornaliero dei casi.

Per stimare  $R_t$  dalla curva occorre introdurre preliminarmente il concetto di  $R_{ist}$ , misura che si può ricavare facilmente, accuratamente e automaticamente dalle curve dei casi, man mano che vengono aggiornati i conteggi.

Inoltre la curva dei casi cumulati tiene conto di tutti i casi, cosa che è impossibile da ottenere con una stima diretta di  $R_t$ , che può essere fatta solo su base campionaria, con molta difficoltà e molta incertezza di misura e con risultati che arrivano sfasati di almeno  $g_i$  giorni.

### Definizione di $R_{ist}$

La definizione di  $R_0$  (e di  $R_t$ ) indica una progressione esponenziale. Un infetto, dopo il suo periodo di capacità infettiva avrà infettato un nuovo infetto più (o meno) una quota di nuovi infetti. Diciamo, ad esempio, un infetto più un altro infetto e mezzo, pari a due infetti e mezzo ( $1 + 1.5 = 2.5$ ). Dopo 2 periodi di infettività gli infettati saranno quelli del periodo precedente ( $2.5$ ) ciascuno dei quali avrà infettato nuovi ( $1 + 1.5 = 2.5$ ): cioè gli ( $1 + 1.5$ ) del periodo 1, moltiplicato ( $1 + 1.5$ ) del periodo 2; e così via. . . .

In generale: dopo il periodo 1: [ $1 \cdot (1+r) = R_t$ ]; dopo il periodo 2: [ $1 \cdot (1+r) \cdot (1+r) = 1 \cdot (1+r)^2$ ]; e così via: [ $(1+r)^1, (1+r)^2 \dots (1+r)^n$ ]. Di fatto si tratta di un processo analogo alla capitalizzazione composta del tasso di interesse  $r$ , dove  $R_t$  è il montante dopo il periodo 1.

Per ottenere quale tasso  $r$  si debba usare per una capitalizzazione composta continua di  $n$  frazioni di periodo che dia, dopo 1 periodo, il montante  $R_t$ , si può scrivere così:  $R_t = (1 + \frac{r}{n})^n$ , equivalente a  $R_t = [(1 + \frac{r}{n})^{\frac{n}{r}}]^r$ . Passando al limite per  $n \rightarrow \infty$  si ottiene:

$$R_t = \lim_{n \rightarrow \infty} [(1 + \frac{r}{n})^{\frac{n}{r}}]^r = e^r \quad (11)$$

dato che

$$\lim_{n \rightarrow \infty} (1 + \frac{r}{n})^{\frac{n}{r}} = e$$

da cui

$$r = \ln(R_t)$$

Cioè  $r$  è l'esponente da dare a  $e$  per ottenere  $R_t$  dopo un periodo di durata infettiva = 1. Cioè:

$$R_{ist} \cdot 1 = r = \ln(R_t)$$

Se voglio determinare  $R_{ist}$  per una unità di tempo più convenzionale, ad esempio i giorni (o le ore) con cui misuro la durata del periodo di infettività di un soggetto infettante  $g_i$ , posso scrivere:

$$R_{ist} \cdot g_i = r = \ln(R_t)$$

da cui:

$$R_{ist} = \frac{\ln(R_t)}{g_i} \quad (12)$$

In questo modo ho il parametro  $R_{ist}$  che caratterizza la crescita esponenziale (come da definizione di  $R_0$  e  $R_t$ ) nel punto che, all'istante  $t$ , l'incremento (o decremento) dei casi giornalieri va a determinare.

Cioè la *curva dei casi giornalieri*, che è la derivata prima  $f'(t)$  della *curva dei casi totali cumulati* (8), diventa la base per calcolare le variazioni di  $R_{ist}$  nel decorrere dell'epidemia. Dato l'andamento esponenziale di  $R_t$  determinato dalla sua definizione epidemiologica, faccio notare che  $R_{ist}$  è l'inclinazione della retta tangente alla curva dei casi giornalieri rappresentata in scala logaritmica. Introduco perciò:

$$h(t) = \ln(f'(t))$$

da cui l'inclinazione  $R_{ist}$ :

$$R_{ist} = h'(t) \quad (13)$$

Ricavato numericamente  $R_{ist}$  dalla curva dei casi giornalieri, si possono calcolare facilmente i parametri  $R_0$ ,  $R_t$  e tempo di raddoppio o dimezzamento dei casi giornalieri lungo tutta la durata dell'epidemia.

$$R_t = e^{(R_{ist} \cdot g_i)} \quad (14)$$

Tempo di raddoppio  $g_{d\vee h}$  (o dimezzamento se  $g_{d\vee h} < 0$ )

$$g_{d\vee h} = \frac{\ln(2.0)}{R_{ist}} \quad (15)$$

Risolvendo  $h(t) = 0$  per  $t_i$ , si ricava il tempo stimato di quando si è avuto il primo caso, antecedente al momento da cui si cominciano ad avere i dati. Da cui:

$$R_0 = e^{(h'(t_i) \cdot g_i)}. \quad (16)$$

## Considerazioni usando le equazioni ricavate

Assumendo:

1.  $R_0 = 2.0$ , come asserito (raddoppio di casi durante la durata di infettività, la stima della durata da accompagnare a  $R_0$  non mi è pervenuta),
2. costante il contesto spazio-temporale di infezione (il che non è, perché basta che il soggetto infettante infetti ad esempio 3 persone in un contesto affollato all'inizio della sua infezione e non al termine di essa per fare differenza, soprattutto in presenza di un alto  $R_0$  e una duratura capacità infettante),
3. costante la suscettibilità di una popolazione (il che è una assunzione molto forte, che non tiene conto dei fenomeni di ostacolo all'infezione insorgenti ad infezione attiva in un ambiente a risorse finite per il contagio, cioè si assume una popolazione infinita vergine al virus indefinitamente nel tempo),
4. costante la infettività per tutto il periodo di infettività dell'individuo infettante, ipotizzato di circa 15 giorni (dato che non mi è noto su quale base temporale di infettività sia stato stimato lo  $R_0 = 2.0$  dichiarato) — in realtà una stima più corretta della infettività media sarebbe esprimibile come  $\bar{v} = \frac{1}{t} \int_0^t v(\tau) d\tau$ , con  $v$  che indica la virulenza istantanea, ma probabilmente è molto difficile stabilire sperimentalmente la curva  $v(t)$ ,

si ha, con queste ipotesi, un indice di infettività istantaneo medio  $R_{ist} \approx 0.04261$  n/g.

Cioè,  $R_{ist} = \ln(R_0)/g_i$  espresso in contagi al giorno in crescita esponenziale durante il periodo di infettività di un soggetto, con  $g_i$  espresso in giorni, valore da usare in espressioni esponenziali del tipo  $R_0 = e^{R_{ist}x}$  dove  $x$  è il tempo espresso in giorni.  $R_{ist}$  è il coefficiente

angolare della retta di una curva esponenziale rappresentata in scala logaritmica con la stessa base dell'esponenziale in questione.

Se ricalcoliamo per 15 giorni di infettività  $R_0 = e^{0.04261 \times 15}$ , otteniamo ovviamente  $R_0 = 2.0$ , cioè 1 infettato infetta 2 soggetti. Dopo 30 giorni di infezione (2 periodi di  $R_0$ ), in esplosione esponenziale senza fattori limitanti, come implicito nella definizione di  $R_0$ , abbiamo il raddoppio degli infettati rispetto ai precedenti 15 giorni, cioè 4 infettati. E così via...

Si può usare l'andamento esponenziale insito nella definizione di  $R_0$  (e associato  $R_{ist}$ ) per cercare di stimare il suo valore numerico a partire dal conteggio giornaliero dei casi di *patologia macroscopica*.

All'inizio del fenomeno epidemico, quando si presume non siano ancora intervenuti fattori limitanti generati dall'agente infettante, la sua esplosione è esponenziale: **il virus non incontra ostacoli generati dalla sua azione**, salvo l'ostacolo della suscettibilità iniziale della popolazione, non ancora modificata dall'azione virale stessa sulla popolazione. Quindi il limite iniziale del fenomeno epidemico è una buona approssimazione della sua crescita esponenziale secondo la definizione di  $R_0$ .

**Se e quando** l'accertamento della *patologia macroscopica* associata al virus è fatto **in modo e con metodo costante nel tempo**, si può ipotizzare che l'andamento temporale di essa rifletta abbastanza bene la diffusione virale sottostante nella popolazione. Non è proprio così, perché intervengono fattori soggettivi nel sottoporsi a diagnosi, e perché il processo di accertamento non è omogeneo sull'insieme della popolazione. Ma non abbiamo di meglio, al momento.

Forse un accertamento diretto della presenza virale su un campione rappresentativo della popolazione ci direbbe di più. Ma anche questo metodo ha molte incertezze (cosa è un campione *biologicamente* rappresentativo di una popolazione rispetto allo spettro dei fattori distorsivi non lineari che intervengono? ec.). Difficile dire.

In pratica, l'operazione di stimare  $R_0$  dalla curva corrisponde a far coincidere l'andamento dei casi rilevati con la curva esponenziale **al suo limite iniziale**. Cioè si fa coincidere la curva dei casi di *patologia macroscopica* censiti in Lombardia e dintorni (che, per definizione, sono stati contagiati da qualcuno), cercando il limite iniziale su cui stimare  $R_0$ .

Se facciamo questo, rispetto alle ipotesi iniziali elencate sopra, allora dovremmo avere  $R_{ist} \approx 0.04261$  n/g, se vale  $R_0 = 2.0$  su 15 giorni. D'ora in poi uso la notazione  $R_0 = 2.0 \parallel 15$  g per specificare il valore numerico di  $R_0$  e periodo espresso in *giorni di infettività ad esso associato*.

Dall'andamento della curva, invece si ricava  $R_{ist} \approx 0.4225$  n/g, pari a  $R_0 = 565.2 \parallel 15$  g. Oppure,  $R_0 = 2.0 \parallel 1.6$  g, cioè un raddoppio di infetti ogni circa un giorno e mezzo.

Per dare una sensazione qualitativa di quale sia la capacità di questo virus di infettare una popolazione vergine al virus molto grande — cioè di infettare un numero illimitato di infettabili **senza** che si formino meccanismi di barriera e immunizzazione —, si può stimare che esso può contagiare nell'ordine di **22\_milioni di persone in 40 giorni**, oppure **1.5\_miliardi di persone in 50 giorni**, oppure **565 persone in 15 giorni**.

Perciò il contagio virale potrebbe avvenire prima che ce ne si possa accorgere dal manifestarsi dell'insorgere di *patologie macroscopiche*, considerando un periodo di incubazione infettante di 15 giorni e in assenza di barriere immunitarie specifiche preesistenti, o altro meccanismo protettivo automatico.

**Ovviamente non succede questo** perché si formano barriere immunitarie e/o comportamentali che sottraggono suscettibili alla diffusione virale, come descrive l'equazione differenziale di Verhulst (2) e associata *funzione logistica* (1). Ma il fatto che l'infettato ospite del virus si muova rapidamente, e possa raggiungere rapidamente territori ancora vergini al virus fa aumentare la potenza

diffusiva iniziale dell'agente infettante. Questo potrebbe spiegare il rapido spuntare parallelo di focolai ovunque nel mondo.

**Per riassumere:** o non vale 2.0 su 15 giorni, come da ipotesi. Oppure vale 2.0 per una infettività che dura per 1.6 g. Ma 1.6 g di infettività mi sembra poco, visto che si chiede una quarantena di 15 giorni o più, e soprattutto **visto che non pare accordarsi con le misure di carica virale** nel tempo fatte su individui con *patologia macroscopica*. Questo almeno da quello che si può capire dalle informazioni virologiche passate sui media.

Questo è quanto riesco a ricavare dalla curva misurata dei casi di *patologia macroscopica* secondo la definizione di  $R_0$ .

Oppure ho sbagliato i conti (ma li ho ricontrollati in modo incrociato mille volte), oppure non ho capito il concetto di  $R_0$ , e non vale la definizione data sopra.

Oppure ancora, si può considerare in prima approssimazione che  $R_0$  evolva nel tempo totale dell'infezione e assuma valore  $R_t = 1.0$  nel punto di flesso della *curva logistica* (cioè quando l'infezione attraversa lo stato stazionario), dopodiché l'epidemia si spegne in modo più o meno simmetrico a come è nata, per acquisita immunità della popolazione, o altro. Cioè  $R_t$  varia nel tempo, come è ovvio, pensandoci su.

**Altra questione è il tempo di decadimento dell'immunità acquisita dalla popolazione**, il quale può essere più o meno breve e avere enormi implicazioni sull'andamento futuro dell'epidemia da COVID-19. In questo semplice modello non si considera questo aspetto.

La *funzione logistica* descrive la resistenza progressiva incontrata dall'agente infettante che modera la sua crescita esponenziale nell'ecosistema chiuso "*infezione sulla popolazione*", come descritto dall'equazione differenziale generatrice della *funzione logistica* e dalla sua soluzione (cfr. 1.1 cenni teorici sopra), nella assunzione

di proporzionalità semplice implicita nel parametro  $\Omega$ .

La resistenza è principalmente dovuta alla immunità acquisita per esposizione al virus, al modello di decadimento di essa e allo spettro di suscettibilità della popolazione oggetto di infezione. A questi si possono aggiungere le modifiche efficaci di comportamento della popolazione e le azioni per indurle, che agiscono in modo simile al processo di immunizzazione.

Se il decadimento delle difese immunitarie acquisite dalla popolazione infettata è rapido, o il virus muta rapidamente nei bersagli biochimici che un vaccino sollecita nel sistema immunitario a riconoscere, allora un eventuale vaccino potrebbe essere di scarsa efficacia. E questo, in presenza di un virus altamente e velocemente infettante, potrebbe essere il caso, per quanto capisco. Allora più efficaci potrebbero essere terapie per le patologie macroscopiche, o terapie che aggrediscano direttamente il virus in un infettato, oltre l'azione autonoma del sistema immunitario.

In presenza di azioni invarianti di contrasto alla diffusione virale, la curva dei casi cumulati di *patologia conclamata* sale e si estingue con una velocità specifica. Azioni aggiuntive, per così dire "artificiali", alterano ovviamente in modo non lineare l'azione costante di contrasto descritta dall'equazione differenziale che genera la *funzione logistica*.

Le eventuali azioni "artificiali" di contrasto (come il congelamento della popolazione mediante *lockdown*) però andrebbero mantenute perennemente per avere l'effetto equivalente di una protezione immunitaria che non sia soggetta a decadimento. Lo stesso decadimento di una protezione immunitaria è il risultato evolutivo di un bilanciamento costi/benefici scovato dal processo evolutivo nell'esplorazione dell'*iper-vastissimo possibile* chimico e biologico.

Qualora quelle azioni "artificiali" di contrasto fossero allentate, allora l'epidemia rispunterebbe come i funghi e riprenderebbe fino a infettare tutti gli infettabili. E un congelamento della popolazione

così rigoroso, come servirebbe per questo virus contagiosissimo e veloce, non può essere mantenuto a lungo, pena morire di fame.

L'atteggiamento para-militare di imporre quarantene e coprifuoco nasce dal presupposto, di dubbia consistenza, che un virus o un agente infettante (e particolarmente un virus veloce a infettare come questo) possa essere debellato con una campagna militare che individua un nemico e lo annienta.

Invece no: nessuno annienta entità come i virus o i batteri. Sono lì da qualche miliardo di anni e la "sanno lunga"... E nessuno può pensare di impedire mutazioni o comportamenti o voleri di altre entità per sempre. Nessuno ferma il meccanismo evolutivo. Si è sempre nel mezzo di *ongoing* processi evolutivi e co-evolutivi darwiniani. Solo una dubbia, indebita e presuntuosa astrazione etica, al limite del delirio di onnipotenza, può pensare di poterlo fare; illudendosi e sbagliando la comprensione di un fenomeno, per come capisco.

Può essere sensato imporre temporanei coprifuoco/isolamento per virus o agenti infettanti lenti e altamente letali come *Ebola*, fino a capire e trovare un rimedio vaccinale e/o terapeutico. Se un virus è lento a infettare, allora è più probabile che la caccia al *paziente zero* e l'*isolamento della intera catena dei contagi* possa essere efficace, fino alla sopravvivenza/estinzione della popolazione infettata e isolata. Ma anche in questi casi è una soluzione temporanea che serve solo a prendere tempo. La cosa ha senso nel caso di virus lenti rispetto al processo e costo di trovare rimedi. Ma nel caso di virus veloci, dubito.

Se invece le azioni di contrasto "*artificiali*" diventano abitudini culturali consolidate in una popolazione (es: distanziamento sociale, confinamento, riduzione di spostamenti futili e frequenti, *smart working*, adozione di diffusa di pratiche igieniche semplici ed efficaci, ec.) allora essi diventano armamentario aggiuntivo analogo ai meccanismi immunitari e ai comportamenti che funzionano a livello inconscio e microscopico. Ma la adozione di tali comportamenti è

un bilanciamento di pro e contro che solo il meccanismo evolutivo può trovare efficacemente. Difficile che una *presunta razionale* norma morale o giuridica imposta possa funzionare. Le stesse norme morali e giuridiche praticate sono effetto di processi evolutivi.

### **In sostanza:**

Se le considerazioni fatte qui su  $R_0$  e sulla velocità di contagio di questo virus sono corrette, allora non dovremmo vedere, a fine contagio, con la dovuta calma e attenzione, sostanziali differenze tra i paesi che hanno tentato di adottare misure drastiche per contenere il virus (come *Italia* e *Cina*) e i paesi che invece lo stanno facendo meno o per niente (come *UK*, *USA*, *Brasile*, ec.), ma si sono concentrati esclusivamente, o quasi esclusivamente, sul trattamento dei casi più o meno gravi di patologia macroscopica, se questo hanno davvero fatto. Diverse invece potrebbero essere le conseguenze e le ricadute dei due opposti atteggiamenti.

### **6.2.5 Considerazioni su $i_c$ e $R_0$ di COVID-19**

<todo>

<rivedere espandere per  $R_t$ >

I ragionamenti su  $i_c$  e  $R_0$ , se sono corretti, e, più in generale, i ragionamenti derivanti da questo esercizio di *regressione logistica*, fanno supporre che ci troviamo di fronte a un virus *iper-iper-infettivo* che ha già probabilmente infettato gran parte della popolazione lombarda, quella da cui emerge la curva dei casi di *patologia macroscopica*. Queste analisi si riferiscono alla manifestazione in *Lombardia+limitrofe*. Probabilmente si possono estendere ad altre zone.

Questa è sia **buona** che **cattiva** notizia.

**Cattiva** perché pare trattarsi di un virus pressoché incontenibile con azioni di contrasto alla diffusione. L'unico contrasto

sensato per rallentare un po' l'onda del contagio sembra essere una autodisciplina nel "distanziarsi", senza paranoia, e nell'adottare le pratiche igieniche indicate. Sperare di debellare il contagio mi pare una illusione: questo virus è qui e impareremo a convivere, e a curare i suoi effetti avversi.

**Buona** perché i, pur tanti, morti improvvisi sono solo una piccolissima parte dei contagiati, se valgono i ragionamenti fatti qui sulla contagiosità del virus. Fanno aumentare grandemente il denominatore dell'indice di letalità  $i_l = (\text{morti}/\text{contagiati})$ . In altre parole, chi viene contagiato dovrebbe avere scarsissime probabilità di avere un malanno grave.

Va sottolineato che l'onda dei casi della *patologia conclamata* è sfasata rispetto al contagio virale, non si sa bene di quanto. E il carico sulle strutture di terapia è ulteriormente sfasato e aggravato dal fatto che la terapia non è istantanea, ma occupa le strutture di terapia per un tempo molto più lungo del periodo di raddoppio dei contagiati dal virus. Insomma il problema fondamentale pare essere il carico sulle strutture di terapia.

Se valgono i ragionamenti fatti qui, pare conveniente concentrarsi sulla terapia emergenziale dei casi di *patologia macroscopica*, piuttosto che sulle azioni di prevenzione del contagio; le quali azioni sono, per quanto mi pare di capire, illusorie, se il virus ha lo  $R_{ist}$  stimato dalla curva.

Infatti, anche se si riuscisse a pensare di veder spegnere il contagio in un focolaio, poco dopo esso riemergerebbe con la stessa virulenza altrove, come un fungo, in funzione della durata della protezione immunitaria acquisita, fino allo spegnersi dell'infezione per intrinseca dinamica dell'infezione virale (cioè: per risposta naturale immunitaria della popolazione). Insomma: parrebbe la dinamica dei raffreddori, con però rischi di complicazioni gravi, anche se relativamente non molto frequenti rispetto alla dimensione della popolazione che parrebbe contagiarsi.

Inoltre, azioni ferree di congelamento di una popolazione non possono essere fatte, né tanto rigorosamente quanto servirebbe per un virus che appare contagiosissimo dai ragionamenti fatti su  $R_0$  e  $i_c$ , né *sostenute rigorosamente per lungo tempo per niente affatto futuri motivi*, pena, ad esempio, morire di fame molto di più che per COVID-19.

Sensati invece paiono il distanziamento interpersonale, lo *smart working*, ove possibile, e le pratiche igieniche indicate dai medici, preferibilmente decisi e praticati autonomamente per autodisciplina da una popolazione consapevole dei rischi. Autodisciplina indotta da un razionale e ragionevole timore del contagio e dei suoi effetti. Speriamo nei costruttori di farmaci, vaccini e dispositivi terapeutici... *Voler sconfiggere* un virus non vuol dire *sapere sconfiggerlo* o *potere sconfiggerlo*.

Se invece il virus risultasse meno contagioso di quanto sembra emergere dal *fitting logistico*, allora le azioni di contenimento acquisterebbero maggior senso.

Insomma, il contenimento del contagio secondo una consolidata prassi anti-epidemic, potrebbe non funzionare contro un virus ad altissima velocità di propagazione.

In ogni caso, c'è da dire che i morti per conseguenza indirette delle azioni di contenimento organizzato del contagio non potranno mai essere contati. Si avrà perciò l'illusione che le azioni di contenimento del contagio avranno avuto effetto. Questo può rassicurare molti, ma illusoriamente, temo, per come capisco. Un po' come *l'acqua di Lourdes*, ma enormemente più costosa.

### **6.2.6 Protezione Immunitaria Pregressa?**

Come sappiamo che una quota di popolazione non abbia già qualche protezione immunitaria pregressa o rapidamente in formazione?

L'ipotesi precedente (6.2.2), insieme alla bassa quota di contagiati rilevati ' indice di contagio stimato ottimale dal modello, rispetto a quanto ci si potrebbe aspettare da una infezione da virus nuovo con come analizzato nella considerazione 4 ' alla contagiosità del virus e alla velocità di diffusione del contagio conseguenti, fa supporre qualche meccanismo (ignoto o non ancora ipotizzato) che ostacola la diffusione e/o la sua manifestazione patologica macroscopica in COVID-19 in almeno buona parte della popolazione generale di un territorio. Per esempio: uno spettro ampiamente diversificato della reazione infiammatoria al virus, forme di immunità pregressa, ec. La maggiore o minore suscettibilità di base della popolazione influenza grandemente la quota di popolazione che subisce infezione virale e, da quest'ultima, quella che poi sviluppa patologia macroscopica.

So poco o nulla dei meccanismi biochimici del sistema immunitario, perciò chiedo scusa di eventuali stupidaggini che ipotizzo.

Una possibile ipotesi è che, pur trattandosi di un virus nuovo, il sistema immunitario di una parte consistente della popolazione abbia già sviluppato anticorpi verso altri agenti patogeni che presentano bersagli biochimici analoghi o uguali a potenziali bersagli sul Corona Virus. In fondo, il sistema immunitario riconosce piccoli bersagli biochimici, rispetto alla dimensione e complessità biochimica di un pur piccolo virus. Questo potrebbe spiegare una rapida o latente reazione immunitaria in chi affronta il contagio virale.

In altre parole, potrebbe non essere completamente vera la affermazione secondo cui "non ci sia immunità in una popolazione, dato che il virus è nuovo". Non so valutare quanto sia azzardata questa mia ipotesi, ma la propongo alla discussione. Però questa ipotesi renderebbe ragione della incongruenza tra alta e veloce contagiosità, da una parte, e stranamente piccolo indice di contagio evidenziato dall'analisi dell'andamento temporale rilevato del contagio nella sua manifestazione in *patologia macroscopica*. Analisi immunologiche post-epidemia forse potranno chiarire meglio la

cosa.

Un'altra possibile ipotesi potrebbe essere uno spettro ampiamente diversificato della reazione infiammatoria al virus nella popolazione. Questo potrebbe far emergere solo una piccola frangia che evolve in patologia macroscopica. Per di più questa frangia si potrebbe ammalare in conseguenza di un meccanismo di reazione all'agente infettante che va in *overshoot*, con caratteristica di regolazione oscillante o amplificante per *overshoot* della reazione (nella teoria e pratica dei controlli a feedback questo è un fenomeno ben noto. Esempio: il fischio *effetto Larsen* del sistema *microfono-amplificatore-altoparlante*). Questa ipotesi di *overshoot* aggiunge caoticità e non-linearità alla relazione:

*infezione da virus*  $\Rightarrow$  *patologia macroscopica*

Queste ipotesi potrebbero spiegare perché relativamente pochi manifestino patologia macroscopica. In presenza ipotetica di una agente altamente e velocemente infettante, al momento, il modello epidemico dai dati rilevati mostra manifestarsi *patologia macroscopica* in una quota relativamente piccola di popolazione ( $i_c = 0.5\%$  della popolazione).

Inoltre, se vale l'analogia con un sistema di controllo a loop chiuso che va in *overshoot*, allora lo smorzamento della reazione (in questo caso sarebbe quella infiammatoria, penso) serve a far riguadagnare stabilità al sistema controllato, anche se questo fa rallentare la risposta del sistema (per esempio: in un controllo PID, se si ha *overshoot*, si riottiene stabilità riducendo o addirittura annullando le componenti P (*Proportional*) e D (*Derivative*), lasciando inalterata e/o rimodulando la componente I (*Integral*), a scapito della prontezza di risposta del controllo).

[https://en.wikipedia.org/wiki/PID\\_controller](https://en.wikipedia.org/wiki/PID_controller)

### 6.2.7 Il *lockdown* funziona?

Circola un dogma: è *evidente* che il **lockdown ha funzionato**.

Io non vedo una evidenza misurabile di questo. Anzi semmai il contrario. Che il lockdown abbia funzionato è un puro atto di fede basato su una assunzione di meccanismo del contagio tutt'altro che sperimentalmente evidente, per quanto riesco a capire. Un atto di fede che attribuisce a questo virus un meccanismo di diffusione accertato per virus molto diversi. Espongo il mio scetticismo in merito con alcune questioni che offro alla riflessione.

#### **Funziona sulla base della curva epidemica?**

I dati esposti, e le stime fatte da essi, al momento non riflettono ancora l'effetto di azioni di contenimento del contagio. Inoltre, tra contagio virale e manifestazione della patologia macroscopica c'è un ritardo di fase, che non so stimare, durante il quale è incerto quanto il virus possa infettare da soggetti asintomatici. Con un  $R_0$  come stimato sopra, è possibile che il virus sia già diventato endemico, almeno in Lombardia e dintorni, vanificando qualsiasi azione di contenimento.

Le azioni di contenimento del contagio sarebbero apprezzabili qualora si potesse notare una riduzione del coefficiente angolare della retta di regressione del fitting logistico linearizzato (inclinazione della retta di regressione che si riduce) calcolata esclusivamente sui nuovi dati che emergeranno col ritardo di fase dovuto al periodo di incubazione tra contagio virale e patologia macroscopica intercettata dal sistema terapeutico. In altre parole, si dovrebbe poter notare una sterzata delle rette e delle curve di fig.3.

La mia prima impressione, sulla base dei dati e della forma della curva e dei ragionamenti fatti, è che le azioni di contenimento del contagio siano servite a poco o nulla. Anzi. La curva epidemica sembra indicare addirittura il contrario: un aumento dei contagi

dopo la adozione delle misure di contenimento. Non penso che questo sia il caso, ma che questo dipenda da un rilevante cambiamento nel metodo di rilevamento dei contagi dopo il 31-mar-2020. Fino a quella data nulla si nota di diverso nella curva epidemica tra prima e dopo l'adozione del *lockdown*.

Infatti, tra:

- tempo per metterle in attuazione;
- $R_0$  elevatissimo;
- rispetto non sempre rigoroso delle stesse, per ragioni tutt'altro che futili;
- sfasamento tra infezione virale e patologia macroscopica dovuto all'ipotetica incubazione e altro;

esse rischiano di poter avere eventualmente effetto quando l'epidemia della *patologia macroscopica* è esaurita e il virus diventato endemico. Se poi l'immunizzazione avesse un decadimento rapido, allora la *patologia macroscopica* rispunterebbe periodicamente con analoga virulenza e velocità, magari smorzata in ampiezza, e con frequenza in funzione della rapidità di decadimento della protezione immunitario/comportamentale acquisita.

**Funziona sulla base della confronto tra diverse azioni fatte da diversi paesi?**

<todo>

confronto ra paesi Brasile VS Lombardia, New York, ec.

**Caccia all'untore nel post epidemia?**

La presunzione che le misure di contenimento abbiano funzionato condizionano la sensatezza di ciò che si prospetta nel dopo-epidemia: *la caccia all'untore per proteggere i non-ancora-contagiati.*

<todo>

espandere

### 6.3 Nota evoluzionistica

Meccanismi evolutivi fondamentali evidenziano che se un patogeno parassita è altamente letale per l'ospite, allora tende a diffondersi con difficoltà, in quanto tende a uccidere il suo ospite e di conseguenza a causare la propria difficoltà a diffondersi.

Viceversa, un patogeno parassita tendenzialmente innocuo per l'ospite può facilmente evolvere per avere alta e veloce diffusività, avvantaggiandosi sui suoi competitors meno veloci e meno inclini a diffondersi, a pari innocuità per l'ospite e a pari mobilità della popolazione degli ospiti. Non nuocere all'ospite genera un ambiente favorevole al patogeno più veloce e incline a diffondersi proprio perché non nuoce all'ospite.

Un virus non nasce dal nulla. Arriva da altri serbatoi in cui è evoluto. E per evolvere non li ha distrutti. Virus e batteri sono lì da alcuni miliardi di anni.

La distanza evolutiva tra ominidi e pipistrelli (o altro serbatoio di mammiferi) è molto piccola, se vista dal punto di vista del virus. Perciò un antenato prossimo di questo virus, così efficiente e veloce a contagiare, non può essere stato troppo nocivo per quei mammiferi, pena la sua estinzione conseguente. Molto improbabile che possa esserlo per noi.

Se si ragiona in termini evoluzionisti, la strategia del virus è molto sensata. O, meglio, il meccanismo di propagazione di questo virus mostra essere molto efficace, dato che attribuire intenzionalità al meccanismo evolutivo è profondamente sbagliato: il processo evolutivo si limita ad esplorare per tentativi il possibile chimico-fisico, tentativi setacciati dall'idoneità contingente all'ambiente.

Manifestarsi in patologia macroscopica con forte ritardo di fase — insieme a infettare velocemente; insieme a fare pochi danni sull'ospite, soprattutto se giovane e femmina, cioè in modo tale da non danneggiare la potenzialità riproduttiva dell'ospite, dato che,

almeno tra i mammiferi, non servono molti maschi per popolare un ambiente, mentre femmine sì; insieme a una probabile scarsa impronta a rapido decadimento sul sistema immunitario dell'ospite — pare una strategia di successo, dal punto di vista del virus. Ed è probabilmente quello che succede nel serbatoio animale da cui si pensa provenga il Corona Virus di COVID-19. In fondo, dal punto di vista del virus, i pipistrelli non sono molto distanti dagli ominidi, come ambiente da colonizzare.

Tutto ciò avviene per fondamentali caratteristiche del meccanismo evolutivo scoperto da *Charles Darwin* e ulteriormente sviluppato dal pensiero evoluzionista: *Ernst Mayr*, *Richard Dawkins*, *Daniel Dennett* (per gli aspetti teorico/algoritmici), e molti altri notevoli pensatori.

Se l'epidemia evolve come sta tratteggiando il modello logistico (rapidità di contagio latente e infettante con scarsa impronta sull'ospite), avremo una ulteriore conferma di questa inferenza evoluzionista.

In ogni caso va tenuto a mente che nulla si può mai escludere in universo chimico-fisico indifferente alla sofferenza dei suoi prodotti, quali noi siamo. La cautela è d'obbligo. Così *Charles Darwin* ad *Asa Gray*:

”Mi sembra che ci sia troppa sofferenza nel mondo. Non riesco a persuadermi che un Dio benefico e onnipotente avrebbe potuto creare gli *Ichneumonidæ* con l'intenzione esplicita di farli alimentare dei corpi viventi dei bruchi dall'interno, o che un gatto debba giocare con i topi.”

<https://www.darwinproject.ac.uk/letter/DCP-LETT-2814.xml>

Comunque, da quanto si intravede, questo virus non sembra l'avvento dell'*Apocalisse*. Quel genere letterario è molto amato. Infatti l'onda montante degli *Apocalittici* è in fermento.

## 6.4 Nota su mortalità e indice di letalità COVID-19

Alcune premesse per inquadrare la mortalità di COVID-19 sono d'obbligo.

Questo non cancella il fatto che la necessità inattesa e incognita di extra terapie intensive e sub-intensive, di ausili respiratori, farmaci e altro, in regime di alta e veloce contagiosità specifica del virus e pandemica, causi stress al limite della tenuta del sistema terapeutico di un paese. Inoltre il numero di quanti vanno a sviluppare la *patologia macroscopica*, stimabile sulla base del metodo attuale di identificazione COVID-19, sono imponenti e concentrate nel tempo. Le conseguenze sul sistema terapeutico sono immediatamente evidenti e pesanti.

Ecco le premesse:

1. La mortalità "fisiologica" media annua per l'Italia durante gli ultimi 10 anni si attesta intorno all'1.0% della popolazione totale ( $\mu = 0.01026, \sigma = 0.000328$ ), pari a circa 620\_mila persone/anno per l'Italia intera con 1 deviazione standard che vale circa 20\_mila persone/anno. Analoghe considerazioni si possono estendere agli altri paesi.

[https://it.wikipedia.org/wiki/Demografia\\_d%27Italia](https://it.wikipedia.org/wiki/Demografia_d%27Italia)

2. Questa "mortalità fisiologica" è grossolanamente ripartibile: 1/3 per cause cardio-vascolari, 1/3 per patologie tumorali, 1/3 per il resto. Questo fa circa 52\_mila morti/mese in Italia che, con incertezza pari a  $\pm 3\sigma$ , fa una incertezza di  $\pm 5000$  morti/mese. Questa incertezza/fluttuazione statistica tra un anno e l'altro di quasi 10\_mila morti "fisiologici" al mese va raffrontata e sovrapposta ai **morti CON Corona Virus** durante lo stesso periodo detti invece **morti PER Corona Virus** dai dati del Ministero della Salute.
3. I morti **conteggiati CON Corona Virus** sono, prevalentemente, anziani pluri-70-enni e affetti da una o più patologie

gravi, di quelle che caratterizzano la "mortalità fisiologica": oltre il 96% dei *morti con CON Corona Virus* ha almeno una patologia grave. Inoltre l'accelerazione relativa di mortalità di questo periodo (febbraio, marzo, aprile) potrebbe facilmente essere seguita da un rallentamento relativo di mortalità subito dopo per depauperamento del bacino di morituri (può sembrare cinico ragionare così, ma così è, che ci piaccia o no). Prime considerazioni sulla mortalità associata al Corona Virus si trovano sul sito ISS:

<https://www.epicentro.iss.it/coronavirus/sars-cov-2-decessi-italia>

[https://www.epicentro.iss.it/coronavirus/bollettino/Report-COVID-2019\\_20\\_aprile.p](https://www.epicentro.iss.it/coronavirus/bollettino/Report-COVID-2019_20_aprile.p)

4. Meccanismi evolutivi fondamentali evidenziano che se un patogeno parassita è altamente letale per l'ospite, allora tende a diffondersi con difficoltà, come descritto nella *Nota Evoluzionistica* sopra 6.3.

Queste considerazioni sembrano poter far dire che la *mortalità PER Corona Virus*, al momento, si presenta annegata nel rumore di fondo della "mortalità fisiologica". Al momento, sembra di poter dire che solo attente e sofisticate analisi statistiche sulle anomalie della "mortalità fisiologica" a fine epidemia potranno eventualmente evidenziare l'effetto di COVID-19 sulla mortalità, considerando la sovrapposizione della co-morbilità. Altre indicazioni potranno venire da analisi patologiche sui deceduti e immunologiche sulla popolazione post-epidemia.

In altre parole, date le premesse fatte, al momento è molto difficile stimare con la necessaria accuratezza l'estensione delle sovrapposizioni degli effetti di patologie plurime in anziani *morti CON Corona Virus* per attribuire causa principale di morte a questa o quella causa. Per questo, anche il numeratore dell'indice di letalità  $i_l$  (*morti/contagiati*) e dell'indice di mortalità  $i_m$  (*morti/popolazione*) è molto incerto.

Per esempio:

1. Se si decide di conteggiare come morti **PER Corona Virus** tutti quelli che alla morte hanno il Corona Virus, allora si ottiene un certo  $morti_{COVID19}$ . Ma questo non chiarisce ancora la questione, perché lascia incerto il numero di tutti quelli che muoiono senza che venga fatto loro il tampone, per esempio.
2. Se si decide invece di escludere dai **morti PER Corona Virus** chi invece è morto per almeno una patologia grave o ha più anni dell'aspettativa di vita alla nascita, allora si ottiene un  $morti_{COVID19}$  molto più piccolo, al limite dello statisticamente trascurabile, per come stiamo vedendo.
3. Lì in mezzo ci stanno tutte le miscele di criteri che si desiderano adottare.

Circa il denominatore di  $i_l$ , le congetture fatte qui mostrano come la stima degli esposti a contagio sia incerta per almeno  $1 \div 2$  ordini di grandezza ( $10 \div 100$  volte, forse 200, più 200 che 10, a naso). Al 2-apr-2020, una prima indagine sierologica sembra indicare una prevalenza di infettati del tra i donatori di sangue di Castiglione d'Adda. Uno studio cinese in proposito pare stimi l'80%. Simili stime pare arrivino da uno studio tedesco. Vedremo con stime più accurate.

La maggiore o minore suscettibilità di base della popolazione influenza grandemente la quota di popolazione che subisce infezione virale e, da quest'ultima, quella che poi sviluppa *patologica macroscopica*.

Analoghe considerazioni valgono per l'indice di mortalità  $i_m$  specifica per COVID-19, ( $morti_{COVID19}/popolazione$ ), in quanto il suo numeratore è incerto.

Un indice di contagio  $i_c^* \approx 0.0045 \div 0.0055$ , cioè lo 0.4%  $\div$  0.5% stimato dal modello sulla curva dei casi di *patologia macroscopica*, mal si accorda con la velocità di contagio mostrata dalle considerazioni su  $R_0$  fatte qui (6.2.4, 6.2.4), e con i numeri di altre epidemie nuove e dilaganti, o ipotizzati per questa: ( $i_c \approx 0.3 \div 0.7$ ), cioè 30%  $\div$  70%, secondo autorevoli stime di contagio virale atteso.

Questo non vuol dire che non si debba cercare di curare tutti, per quanto si può. E non vuol dire che lo stress sul sistema terapeutico sia stato enorme, ed enorme gratitudine vada espressa a chi ci lavora. E non vuol dire che la perdita di un proprio caro non sia dolorosa e urti il senso di compassione di ciascuno.

Ma capita a tutti di morire, una volta o l'altra. E indignarsi contro un virus o contro i limiti della capacità di cura o di azione non vuol dire aver capito come funziona un meccanismo. Forse ricordare ogni tanto che non siamo onnipotenti aiuta.

Mischiare etica e scienza non porta mai buono. Da un *essere* non deriva un *dover essere*, ci ammoniva *David Hume*. Da una *descrizione*, da una *osservazione*, non deriva automaticamente una *prescrizione*. E una *prescrizione* non è una *descrizione*, né una *osservazione*. In mezzo ci sta un *volere*, una *scelta*, una *decisione*, un *arbitrio*. Sembra banale, ma sovente lo si dimentica, in tempi concitati.

## 7 Bibliografia Cautele e Ringraziamenti

Questo documento è fornito "as is". Esso è liberamente scaricabile e utilizzabile.

<http://www.adaptive.it/covid19>

Non ho avuto modo di consultare letteratura specifica sull'argomento. Inoltre non ho facile accesso a tale letteratura. **I riferimenti da me utilizzati sono menzionati lungo il documento.** Mi scuso perciò in anticipo se quanto detto qui si sovrappone involontariamente ad altri studi effettuati, o in corso, che non so menzionare.

Devo al fisico **Aldo Gotta**, allievo di *Carlo Castagnoli* e conosciuto alla *Softtech* di Torino allora condotta da **Roberto Pagani** (tutti gravitanti attorno alla figura di **Lorenzo Matteoli**), il fatto di avermi parlato, nel 1990, della curve logistiche e relative regressioni. All'epoca egli lavorava allo *IIASA* di Vienna dove conobbe *Cesare Marchetti* che si ingegnava a fittare ogni cosa con curve logistiche, e ci azzecava.

Devo a: **Lorenzo Matteoli, Massimo Foti, Roberto Pagani, Gabriella Peretti, Gianvincenzo Fracastoro, Marco Masoero, Federico Butera, Donatella Marino, Giorgio Ceragioli, Franco Ossola**, e a tutti gli altri di quel gruppo, dove aleggiava la figura ieratica di *Giuseppe Ciribini*, cibernetico, come *Marina Gariboldi, Luciana Conforti, Antonella Marucco, Paola Caccia, Andrea Ketoff, Andrea Aparo, Gabriella Funaro, Elena Montacchini, Giancarlo Pavoni, Grazia Zanchetta, e altri, facenti capo, in un modo o nell'altro al *Politecnico di Torino*, al *Politecnico di Milano* e all'*ENEA*, quello che ho cominciato a capire di computazione, fisica e matematica dei *Sistemi Complessi* e dei *Sistemi Energetici*.*

Devo ad **Anna Ostanello** del *Politecnico di Torino* se ho imparato qualcosa della matematica dei *Sistemi Complessi*, della *Ricerca*

*Operativa della Teoria dei Grafi e della matematica dei reticoli.*

Devo a **Sergio Vallero** e **Gualtiero Franzino** dello *CSELT* di Torino e poi *EDS*, quello che so di elettronica per sistemi di misura e controllo.

Devo a **Flavio Conti**, **George Helké** e **Stefano Cocchi** del *JRC di Ispra*, insieme a **Roberto Pagani**, allora *Softech*, di avermi sostenuto durante i miei primi esperimenti con algoritmi *ANN* (*Artificial Neural Networks*) nel lontano 1995-1997, insieme a quanti del *JRC* con cui ho potuto discuterne.

Devo quanto ho imparato sui sistemi di misura complessi e sui controlli a loop chiuso a **Gualtiero Balossini**, **Vittorio Barara**, **Antonio Celozzi**, **Giuseppe Musso**, **Vittorio Lignarolo**, **Pierluigi Teodori**, **Carlo Dionisio** e a tutti coloro con cui ho avuto modo di lavorare in *Microtecnica-Collins Aerospace*.

Devo a **Loredana Loconte** e **Roberto Merletti** (*MRC di Cambridge, Boston University, LISiN, Politecnico di Torino*) di avermi introdotto alle problematiche del *sistema neuromuscolare*, della *elettromiografia non invasiva* e ad altre problematiche di bioingegneria.

Ovviamente la responsabilità delle eventuali stupidaggini scritte qui è esclusivamente mia.

Graditi riscontri e critiche. Vogliate scusare la scarsa formalità di scrittura. Si tratta di considerazioni, calcoli e note elaborate e scritte in fretta come eventuale spunto di discussione.