Modello idraulico di una epidemia

di Roberto Chiappi

46
Sarà impuro tutto il tempo che avrà la piaga; impuro; se ne starà solo*; abiterà fuori del campo.* (*Bibbia, Levitico*)

Poiché il pensiero umano è analogico, i modelli idraulici aiutano enormemente la comprensione di fenomeni che possono essere anche molto complessi.

Nello studio dell'elettricità è ben nota l'assimilazione della corrente elettrica (misurata in Ampere) con il flusso di un liquido in una tubazione. La differenza di potenziale (misurata in Volt) è spesso assimilata ad una cascata con salto più o meno ampio. La resistenza (misurata in Ohm) può essere assimilata agli impedimenti che un fluido incontra al suo scorrimento.

La *System Dynamics,* ideata da Forrester, è un potente strumento di simulazione dinamica basata sui concetti di: variabili di livello, variabili di flusso e circuiti di retroazione positivi o negativi.

Elenco dei simboli e concetti adottati:
N   Flusso giornaliero dei Nuovi positivi
Y   Livello degli attualmente positivi (contagiati)
G   Flusso giornaliero dei Guariti.
M   Flusso giornaliero dei Morti.
DY = Yn+1 – Yn = N – G – M. Variazione giornaliera del livello Y.
Fase 1 Crescita della variabile Y   (DY > 0)
Picco Massimo della variabile Y   (DY = 0)
Fase 2 Decrescita della variabile Y (DY < 0)
+   Feed – Back positivo. Legame tra due variabili (nel caso Y ed N) che si rinforzano reciprocamente sia nella Fase 1 di crescita, che nella Fase 2 di decrescita.
S   Casi Totali da inizio epidemia. Somma (o integrale) dei Nuovi positivi.
F   Flesso della la curva sigmoide (S). Coincide con il Picco dei Nuovi positivi (Nmax)
R Rimossi. Somma progressiva di G + M

Modellino idraulico di una epidemia

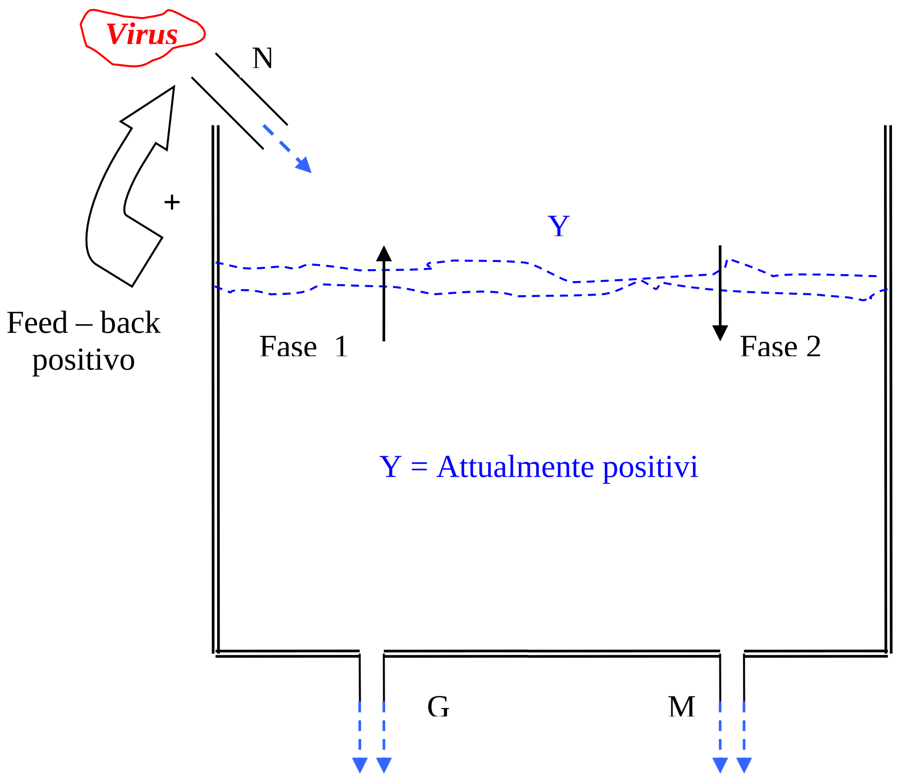
Con riferimento alla figura 1 riportata sotto si consideri un recipiente che contiene invece di litri persone attualmente positive o contagiate (Y) ad un certa data. Il recipiente è alimentato ogni giorno dal flusso dei nuovi positivi (N). Il livello Y rappresenta anche il livello di rischio per ognuno di contrarre il contagio: più persone contagiate ci sono, maggiore è il rischio di contrarre il contagio. Fondamentale è il circuito di retroazione positiva tra livello Y e flusso N: I nuovi positivi (N) aumentano il livello degli attualmente contagiati (Y) e questo livello, a sua volta, accresce il flusso di N . Questo circuito di Feed-back positivo spiega la crescita esponenziale degli infetti o contagiati

nella fase 1 della epidemia, ma spiega anche il rinforzo della decrescita nella fase 2: il calo dei positivi (Y) riduce le occasioni di nuovi contagi (N) che a loro volta riducono il livello Y.
Il serbatoio del modello presenta anche due uscite: il flusso giornaliero dei guariti (G) ed il flusso giornaliero dei morti (M). Se improvvisamente tutti i positivi (Y) guarissero (G) l'ondata epidemica cesserebbe. Ma anche (soluzione non certo auspicabile) se tutti i positivi morissero (M), analogamente l'ondata epidemica cesserebbe.
In definitiva il livello Y degli attualmente positivi è regolato dalla equazione ricorrente (o alle differenze finite):

$$Y_{n+1} = Y_n + N - G - M \qquad \text{ovvero} \qquad DY = Y_{n+1} - Y_n = N - G - M$$

Si tratta di un bilancio: a fine giornata il nuovo livello (Y_{n+1}) dipende dal livello del giorno precedente (Y_n) aumentato o diminuito di quanto entrato e uscito dal serbatoio durante la giornata: $DY = N - (G + M)$

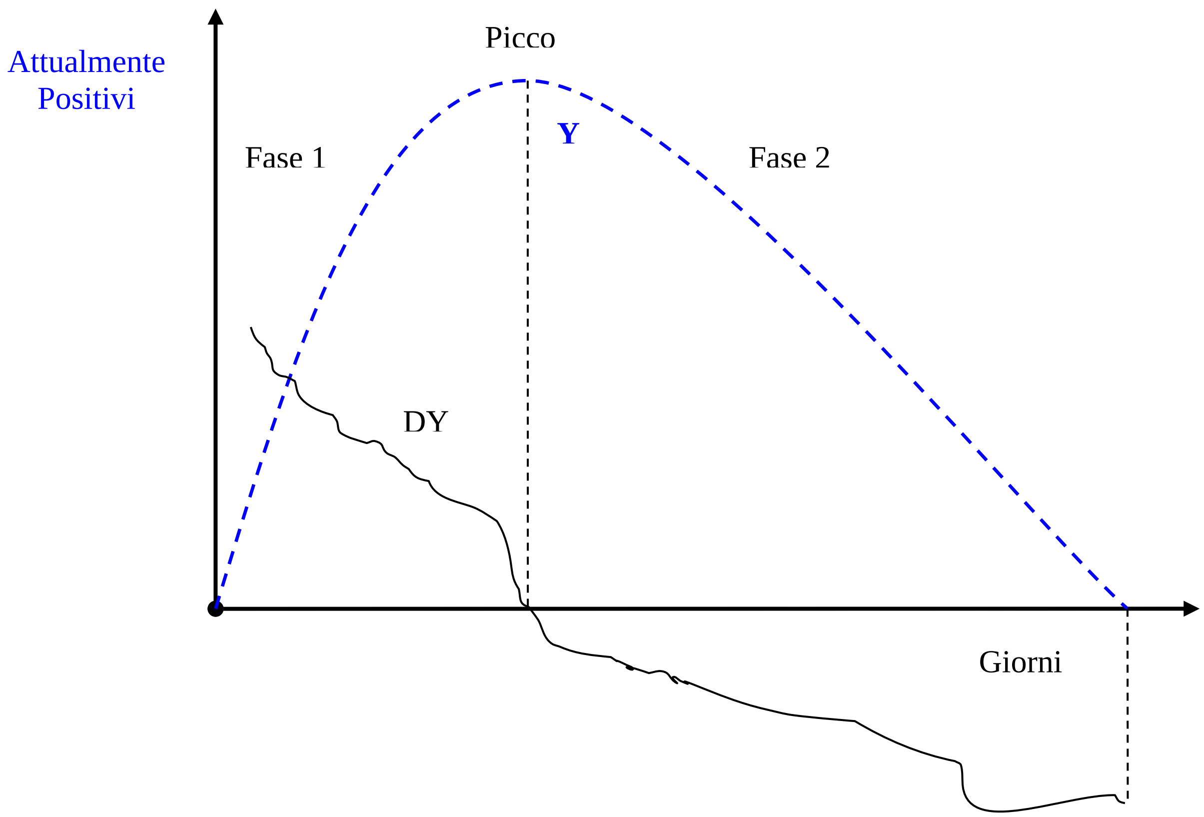
Fig.1



Avremo dunque una Fase 1 in cui il livello Y continua a crescere (anche esponenzialmente) poiché il bilancio DY è positivo. Poi avremo un Picco (o un periodo di Plateau) in cui DY = 0 (N è circa uguale ad M + G) ed infine una Fase 2 in cui il livello di Y continua a decrescere poiché DY < 0.
È chiaro, come già scritto, che il livello Y (Attualmente positivi) rappresenta anche il rischio di contagiarsi poiché esso misura il numero di persone portatori di virus, in casa, in strada, nei negozi, nelle case di riposo e negli ospedali portatori del virus.
Disporre di un farmaco efficace significa aprire maggiormente il rubinetto (G) dei guariti e chiudere maggiormente il rubinetto (M) dei morti. Se tutti i contagiati morissero immediatamente il virus (che sopravvive solo nell'organismo umano, a differenza dei batteri) si estinguerebbe assieme a loro.
A partire da questo dato alcuni virologi evoluzionisti sostengono che dal punto di vista del virus e della sua spinta a sopravvivere e prosperare far morire gli ammalati non sarebbe una buona strategia.
Disporre di un vaccino significa invece chiudere maggiormente il rubinetto dei nuovi postivi. Di solito si sostiene che, se un vaccino è efficace, basta che sia vaccinato lo 80-95% della

popolazione (la percentuale varia da malattia a malattia) per raggiungere la cosiddetta immunità di gregge.
Dunque il farmaco adotta una strategia curativa, il vaccino persegue una strategia preventiva.
In assenza di un farmaco e di un vaccino, quando il virus si trasmette da umano ad umano e non sopravvive all'esterno dell'organismo, l'unica strategia di difesa possibile è quella del distanziamento sociale (che sarebbe meglio chiamare distanziamento fisico) tra umani. Con questa strategia si contribuisce alla chiusura del rubinetto dei nuovi positivi (N).

Fig.2

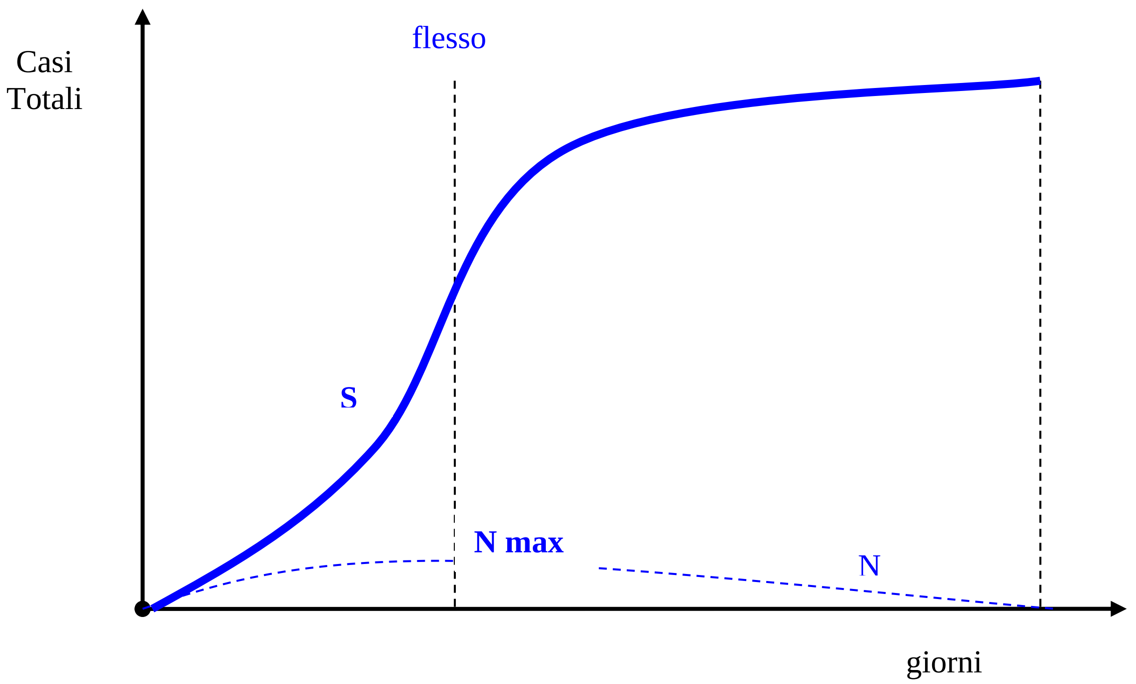


Il grafico riportato in Fig. 2, puramente qualitativo, mostra l'andamento, a simil parabola rovesciata, degli attualmente positivi (Y) e quello della spezzata DY = Yn + 1 – Yn. Dal grafico si evince che, nella Fase 1, si può tentare di prevedere il giorno del Picco di Y con una regressione lineare (analisi del trend) effettuata su DY: quando DY si annulla Y è massimo.
Nella Fase 2 si può tentare di stimare la fine della ondata epidemica sempre con una analisi del trend effettuata su DY, calcolando poi, sin quando è maggiore di zero, il valore degli attualmente positivi con la semplice equazione ricorrente: Yn+1 = Yn + DY.

I Casi totali (S) si calcolano sommando i Casi totali del giorno precedente con i Nuovi postivi (N) della giornata. L'analogia idraulica fa ora riferimento ad un recipiente identico al precedente ma senza i due rubinetti in uscita, così che il recipiente non potrà far altro che riempirsi sempre più. Avremo però anche in questo caso due fasi: nella prima si riempirà più velocemente, nella seconda più lentamente. Quindi la curva sigmoide (S) rappresentante i Casi totali (Fig. 3) avrà prima una concavità verso l'alto e poi una convessità verso l'alto. La separazione è data dal punto di flesso cui corrisponde il massimo giornaliero dei nuovi positivi (N). Si osservi che S rappresenta l'integrale dei nuovi positivi (N) che assumono la forma di una simil parabola rovesciata.
Generalmente, in assenza di un farmaco o di un siero efficace, il massimo (Nmax) dei nuovi postivi è antecedente al massimo della curva Y degli attualmente positivi.

L'andamento sigmoide dei casi totali (S) serve principalmente per il confronto con altri paesi e per avere memoria storica dell'entità dell'epidemia.
Alcuni poi considerano una seconda curva ad S dei Risolti (R), cioè del progressivo dei Guariti più il progressivo dei Morti. Questa curva, rispetto a quella di figura 3 sarà traslata in avanti nel tempo e potrà avere forma un po' diversa, ma avrà, al termine, la stessa ordinata, cioè il numero dei Casi Totali provocati dall'ondata epidemica.
Con riferimento al modello idraulico di figura 1 è come se sotto il recipiente ivi riportato ve ne fosse un altro identico, ma senza uscite ed alimentato da G ed M. Al termine della ondata epidemica questo recipiente conterrà i Casi Totali.

Fig. 3



Modelli statistici di una epidemia
Le funzioni disponibili sono molteplici e comprendono diversi tipi di curve per rappresentare gli andamenti illustrati qualitativamente nei grafici riportati sopra. Tra queste ricordiamo le funzioni: lineari, di potenza, polinomiali ed esponenziali (Es. curva logistica e curva di Gompertz).
I due principali indicatori che si usano per comprendere la bontà dell'adattamento di una curva ai dati osservati sono il coefficiente di correlazione $-1 < R < 1$ (vale 0 se non c'è correlazione, 1 se c'è perfetta correlazione lineare diretta, -1 se c'è perfetta correlazione lineare inversa) e l'indice di Fisher: F = Varianza spiegata / Varianza dei residui . Se F =1, la Varianza spiegata è uguale a quella dei residui. La funzione proposta $Y = f(x)$ non spiega nulla di più della media di Y. Più grande è il valore di F maggiore sarà il valore interpretativo della funzione proposta.
L'analisi statistica di regressione è particolarmente efficace nei problemi di interpolazione dove si tratta di minimizzare i quadrati degli scarti (residui) tra i dati osservati e la funzione proposta. Più complesso è il caso della estrapolazione (tentativo di prevedere il futuro). Il grande De Finetti metteva in guardia con un proverbio, credo lucano: "Acqua passata non macina più!". In effetti nelle estrapolazioni bisognerebbe dare pesi via via crescenti alle osservazioni più recenti.
Uno strumento più semplice della pesatura (in omaggio al rasoio di Occam), per ottenere facilmente buoni risultati, è quello di trovare una grandezza ragionevolmente approssimabile con

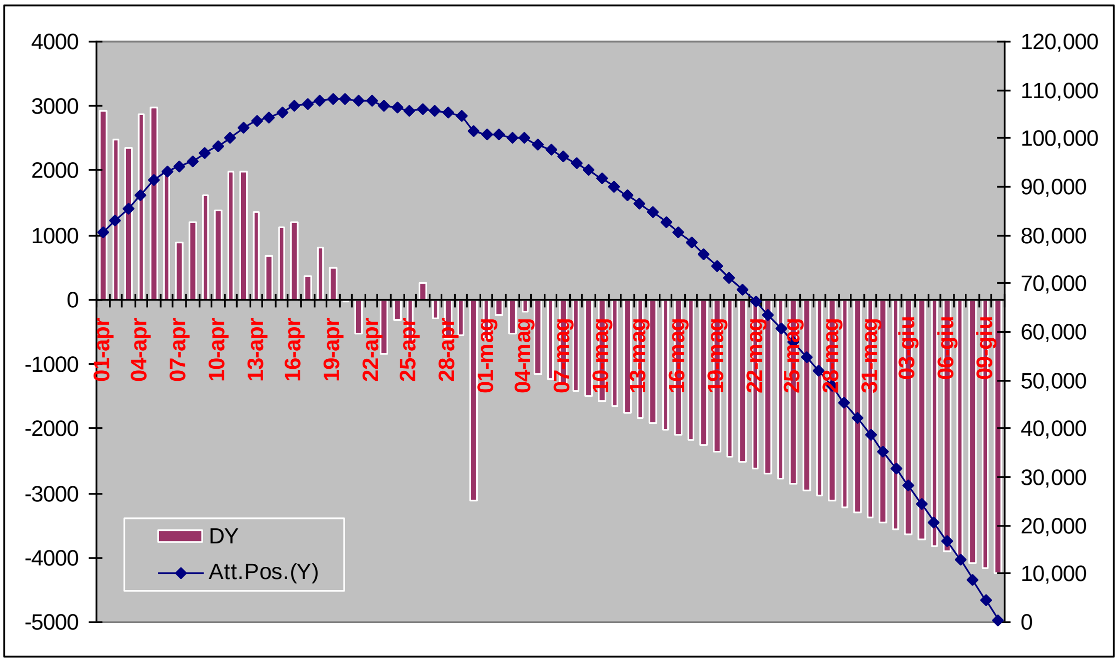
segmenti di retta (nel caso DY) e poi prendere in considerazione solo le ultime 5 o 10 o 20 o 30 osservazioni.
Ad esempio si considerino i dati del Covid 19 per l'Italia nella fase 1 in cui si ricerca il periodo di rischio massimo di contagio cioè il massimo degli "attualmente positivi Y" o "infetti" (Vedi Fig. 2). Si trova facilmente che questo massimo (108,257 di infetti nel giorno) si è realizzato il 19 Aprile, l'ultimo in cui DY risulta positivo.
Questo massimo non deve essere confuso con quello (Nmax) dei Nuovi Positivi (vedi Fig.3) che per l'Italia è stato raggiunto il 21 Marzo (Nmax = 6.557 nuovi casi nel giorno).

Si supponga ora, con riferimento alla figura 2, di essere sempre in Italia, nella seconda fase, dopo il Picco, ed in particolare alle ore 18 del giorno 4 di maggio (vedi Tab.1). Dunque tutti i numeri precedenti o uguali a quelli riscontrati il 4 di Maggio sono dati consuntivi, i numeri successivi sono previsioni (*Corsivo*)

Tab. 1

Giorno	DY	Att Posit.
01-apr	2937	80,565
02-apr	2477	83,042
03-apr	2339	85,381
04-apr	2886	88,267
05-apr	2972	91,239
06-apr	1941	93,180
07-apr	880	94,060
08-apr	1202	95,262
09-apr	1615	96,877
10-apr	1396	98,273
11-apr	1996	100,269
12-apr	1984	102,253
13-apr	1363	103,616
14-apr	675	104,291
15-apr	1127	105,418
16-apr	1189	106,607
17-apr	355	106,962
18-apr	809	107,771
19-apr	486	108,257
20-apr	-20	108,237
21-apr	-528	107,709
22-apr	-10	107,699
23-apr	-851	106,848
24-apr	-321	106,527
25-apr	-680	105,847
26-apr	256	106,103
27-apr	-290	105,813
28-apr	-608	105,205
29-apr	-548	104,657
30-apr	-3106	101,551
01-mag	-608	100,943
02-mag	-239	100,704
03-mag	-525	100,179
04-mag	-199	99,980
05-mag	*-1147*	*98,833*
06-mag	*-1233*	*97,601*
07-mag	*-1319*	*96,282*
08-mag	*-1405*	*94,878*
09-mag	*-1491*	*93,387*
10-mag	*-1577*	*91,811*
11-mag	*-1663*	*90,148*
12-mag	*-1749*	*88,399*
13-mag	*-1835*	*86,564*
14-mag	*-1921*	*84,644*
15-mag	*-2007*	*82,637*
16-mag	*-2093*	*80,544*
17-mag	*-2179*	*78,365*
18-mag	*-2265*	*76,100*
19-mag	*-2351*	*73,749*
20-mag	*-2437*	*71,312*
21-mag	*-2523*	*68,789*
22-mag	*-2609*	*66,180*
23-mag	*-2695*	*63,485*
24-mag	*-2781*	*60,704*
25-mag	*-2867*	*57,837*
26-mag	*-2953*	*54,884*
27-mag	*-3039*	*51,845*
28-mag	*-3125*	*48,720*
29-mag	*-3211*	*45,508*
30-mag	*-3297*	*42,211*
31-mag	*-3383*	*38,828*
01-giu	*-3469*	*35,358*
02-giu	*-3555*	*31,803*
03-giu	*-3641*	*28,162*
04-giu	*-3727*	*24,434*
05-giu	*-3813*	*20,621*
06-giu	*-3899*	*16,721*
07-giu	*-3985*	*12,736*
08-giu	*-4072*	*8,664*
09-giu	*-4158*	*4,507*
10-giu	*-4244*	*263*

Fig. 4



Decidiamo di selezionare, per la proiezione lineare, i dati consuntivi di DY che, partendo dal 16 Aprile, arrivano al 4 Maggio. Trasciniamo (croce in basso a destra in Excel) le celle sino, ad esempio, al 20 Giugno. Nella colonna di destra Att. Posit. sono calcolati gli Attualmente positivi od infetti con la semplice formula ricorrente Yn = Yn-1 + DY. Osserviamo che dopo il 10 Giugno Yn diventa negativo cosa che non ha alcun significato fisico (qualunque modello matematico funziona entro limiti definiti). Dunque cancelliamo le proiezioni posteriori al 10 Giugno. Il risultato ottenuto è riportato in Tab. 1 e rappresentato graficamente in Fig. 4.
In definitiva la proiezione, fatta il 4 di Maggio, prevede la fine della ondata epidemica il 10 di Giugno con ancora solo 263 infetti che rappresentano il 100*263/108,257 = 0.24 % del picco massimo.

Cenni storici conclusivi
Le considerazioni storiche conclusive sono affidate ad estratti dalla voce italiana di Wikipedia: "*Modelli matematici in epidemiologia*" citata nei riferimenti bibliografici.
Il primo scienziato che ha cercato sistematicamente di quantificare le cause della morte è stato John Graunt nel suo libro *Natural and Political Observations upon the Bills of Mortality.* del 1662. I disegni di legge studiati erano elenchi di numeri e cause di decessi pubblicati settimanalmente. Graunt fu il primo a correlare la salute dei cittadini di Londra alle loro condizioni socio-economiche attraverso un'attenta analisi dei registri delle nascite e delle morti custoditi nelle parrocchie londinesi. Successivamente la pratica della registrazione dei decessi fu adottata anche dalle autorità civili.
Il primo resoconto della modellazione matematica della diffusione della malattia fu effettuato nel 1766 da Daniel Bernoulli. Formatosi come medico, Bernoulli ideò un modello matematico per difendere la pratica dell'inoculazione contro il vaiolo. Questo modello mostrava che l'inoculazione universale contro il vaiolo avrebbe aumentato l'aspettativa di vita da 26 anni e 7 mesi a 29 anni e 9 mesi.
Il lavoro di Daniel Bernoulli ha preceduto gli studi sulla vaccinazione di Edward Jenner e la elaborazione della "teoria dei germi" da parte di Pasteur.
Nel 1840, Farr presentò una lettera all' *Annual Report of the Registrar General of Births, Deaths and Marriages in England* . In quella lettera, proponeva la applicazione della matematica ai registri delle morti durante una recente epidemia di vaiolo, sostenendo che:

"*Se non è possibile scoprire la causa latente delle epidemie, è possibile indagare sul modo in cui opera. Le leggi della sua azione possono essere determinate dall'osservazione, nonché dalle circostanze in cui si verificano le epidemie o da cui possono essere controllate.* "
William Farr fu probabilmente il primo a introdurre una teoria matematica delle epidemie, utilizzando un'equazione polinomiale di terzo grado per descrivere e predire l'andamento della peste bovina nel 1865.
Il medico tropicale inglese Ronald Ross, già premio Nobel (1902) per aver stabilito che la malaria si trasmette con le punture delle zanzare, proponeva il primo modello probabilistico utilizzato in epidemiologia correlando il diffondersi della malattia al numero di zanzare.
La propagazione della malattia e l'entità di un'epidemia in una popolazione dipendono da diversi fattori spaziali e temporali che sono stati inquadrati negli anni Venti del Novecento da Lowell Reed e Wade Hampton Frost nel cosiddetto *modello epidemico Reed-Frost*. Nel modello Reed-Frost la propagazione della malattia varia in relazione alla probabilità di contatti infettivi e di ospiti suscettibili. Tale probabilità è influenzata dalla densità della popolazione, dal tempo e dalla durata del contatto, dalla suscettibilità dell'ospite, dall'infettività dell'ospite, dalla trasmissibilità dell'agente, dall'infettività dell'agente e dalla virulenza dell'agente.
Negli ultimi decenni sono stati elaborati decine di modelli matematici sempre più sofisticati che cercano di catturare le complesse dinamiche spazio-temporali delle forme epidemiche (o endemiche) che caratterizzano le diverse malattie infettive. Tali modelli rientrano in due categorie generali: modelli statistici che hanno come unico obiettivo la descrizione della struttura dei dati e modelli meccanicistici che tentano di rappresentare i processi che si ritiene abbiano generato i dati.
La maggior parte di questi modelli rimane piuttosto astratta e sottodeterminata, e comunque appare irrealistico attendersi un modello unificato delle dinamiche epidemiche, mentre sarebbe più ragionevole cercare di valutare più analiticamente e sperimentalmente il peso delle diverse variabili nei diversi modelli epidemici che tentano di spiegare o predire le specifiche dinamiche ospite/parassita.

Riferimenti Sitografici:

Modelli Matematici in Biologia
http://www.mat.unimi.it/users/gaeta/BMT/chap_add.pdf

Modelli matematici in Epidemiologia
https://it.wikipedia.org/wiki/Modelli_matematici_in_epidemiologia

Un modello matematico di epidemie
http://www1.mat.uniroma1.it/people/pensionati/lamberti/lauree/sir.pdf

Matematica della Pandemia (min.18 – min. 30)
https://www.youtube.com/watch?v=7DTyMg3ciSE

Il modello SIR (Suscettibili, Infetti, Rimossi)
http://www1.mat.uniroma1.it/people/pensionati/lamberti/lauree/sir.pdf

Il Modello Reed-Frost
http://web.tiscali.it/mferrodir/I_matematici_sono_come_i_franc/STUDIO_DELLE_EPIDEMIE/IL_MODELLO_DI_REED_E_FROST2/il_modello_di_reed_e_frost2.htm

Idraulica
https://it.wikipedia.org/wiki/Idraulica

Analogia tra circuiti elettrici ed idraulici
http://cabestano.altervista.org/alterpages/files/circuiti_alessandra.pdf

La dinamica dei sistemi di Forrester

https://www.matematicamente.it/approfondimenti/problem-solving/la-dinamica-dei-sistemi-j-w-forrester/

Analisi di tendenza con Excel (Video)
https://www.youtube.com/watch?v=kDTpkNcg6Io

Funzioni lineari, di potenza, polinomiali, esponenziali (Logistica, Gompertz, ecc)
http://www.dm.unibo.it/~cagliari/agraria/teoria/funzioni_elementari.pdf

La Curva dei contagiati da Covid 19
https://www.neodemos.info/articoli/la-curva-dei-contagiati-da-covid-19-la-ricerca-del-punto-di-svolta/

Coronavirus nel mondo
https://www.endcoronavirus.org/countries?itemId=4afy7b4zszrcbczlfl3hay6sgah2dm

Coronavirus in Italia
https://lab24.ilsole24ore.com/coronavirus/

Il foglio elettronico come strumento per il Problem Solving
https://www.francoangeli.it/Ricerca/scheda_libro.aspx?Id=15921

Gli enti non sono da moltiplicare oltre la necessità (Occam)
https://www.matematicamente.it/approfondimenti/problem-solving/sp-15712/

Previsioni: Filosofia e Matematica
https://www.matematicamente.it/approfondimenti/matematica/previsioni-filosofia-e-matematica/

Statistica (Tec. Gen. Manag) scaricabile gratuitamente dal sito Pionieri ENI
http://www.pionierieni.it/wp/wp-content/uploads/SNR-131-General-Management.-R.-Chiappi-Scuola-Sup.-Mattei-2001.pdf

Indice di Correlazione r di Pearson
https://it.wikipedia.org/wiki/Indice_di_correlazione_di_Pearson

Analisi della varianza e test di Fisher
https://lorenzogovoni.com/anova-excel/

Differenza tra Siero e Vaccino
https://medicinaonline.co/2018/03/03/differenza-tra-vaccino-e-siero/

Differenza tra Vaccino e Farmaco (caso del Covid-19)
https://www.agi.it/fact-checking/news/2020-03-18/coronavirus-cura-tempi-7614793/