

## Analisi dell'epidemia COVID19 in Italia, usando il modello epidemiologico SIR

Roberto Battiston  
Università di Trento  
16 aprile 2020  
V.1.1

Cos'è un modello **SIR** ?

Si tratta di un modello epidemiologico che analizza l'andamento dell'infezione nel tempo, fornendo alcuni dati in ingresso e osservando lo sviluppo nel tempo di variabili osservabili nel tempo come il numero di contagiati, il numero di guariti, il numero di morti nel tempo.

Nel caso del modello SIR, l'acronimo sta per Suscettibile (**S**), Infetto (**I**), Risolto o Resistente (**R**).

Si tratta di tre equazioni differenziali

$$\begin{aligned}S'(t) &= -\beta \cdot S(t) \cdot I(t) / N \\I'(t) &= \beta \cdot S(t) \cdot I(t) / N - \gamma \cdot I(t) / D_i \\R'(t) &= \gamma \cdot I(t) / D_i\end{aligned}$$

Condizioni iniziali:

$$\begin{aligned}S(0) &= p - I(0) - R(0) \\I(0) &= i_0 \\R(0) &= 0\end{aligned}$$

E' un modello largamente studiato in letteratura. Si risolve per via numerica con relativa facilità e si può confrontare con i dati, a patto che i dati siano acquisiti con un certo grado di sistematicità, affidabilità e regolarità.

I grafici che vengono ricavati con questo modello sono più efficaci degli studi fatti cercando di estrapolare i dati del passato verso il futuro, perché le equazioni alla base descrivono un tipo di relazione tra le variabili tipica del virus.

Data una popolazione di **N** individui, la possiamo infatti dividere in tre parti:

- coloro che possono essere contagiati (**S**),
- coloro che sono infetti e possono contagiare, (**I**),
- coloro che sono guariti o deceduti, in ogni caso resistenti al virus (**R**).

In ogni momento la somma di **S+I+R** fa **N**.

Il numero **S** scende nel tempo, mano a mano che si moltiplica l'infezione. Il numero **I** cresce nel tempo a causa del moltiplicarsi dell'infezione meno il tasso di guariti e di morti (**R**), il

numero  $R$  cresce nel tempo. Alla fine dell' epidemia, buona parte della popolazione sarà passata per la fase di infezione e sarà poi o guarita o deceduta. Il tasso con cui il contagio si propaga e il tasso con cui gli infetti guariscono o muoiono, dipendono da due parametri fissati,  $\beta$  e  $\gamma=1/D_i$ , che non variano con il tempo e che possono essere dedotti dai dati. La frazione della popolazione che sarà stata risparmiata dall'epidemia dipende dal parametro chiamato  $R_0=\beta/\gamma=\beta*D_i$ , il numero di infetti che ogni contagiato crea a sua volta.

All' inizio dell' epidemia  $R(0)=0$ ,  $I(0)$  è un numero molto piccolo (unità-decine di persone) è  $S(0)$  un numero molto grande, per esempio la popolazione di una nazione come l' Italia o di una regione come l'Emilia o la Lombardia.

Dopo il lockdown, improvvisamente  $S$  cala di un fattore molto grande diventando  $S'$ . Si tratta di un calo anche più di 100 volte, in quanto la popolazione diventa compartimentata e gran parte di essa non è più raggiungibile dall' infezione. Nel caso delle regioni italiane e dell'Italia nel suo insieme, questa riduzione, ricavabile dai dati è di circa 300 volte.

Da quel punto in poi, il contagio si sviluppa e si spegne all'interno della popolazione ridotta, se non ci sono modifiche nelle condizioni di lockdown. La maggior parte della popolazione  $S'$  si infetta e poi guarisce o muore: in queste condizioni il contagio termina da solo in un tempo che è collegato a quanto è stretta la condizione di lockdown. Un lockdown non efficace allunga il tempo del contagio in quanto non spegne del tutto il contagio: ad esempio la non separazione dei contagiati dalle famiglie contribuisce all' allungamento della durata dell' epidemia, perché dopo un certo numero di giorni passati in casa assieme con un ammalato, il rischio di contagio diventa molto alto per coloro che sono sani. Una quarantena stretta e di tipo individuale delle persone nell' ambito della popolazione  $S'$  riduce i tempi di arresto del contagio, salvaguardando anche parte della popolazione  $S'$  dal contagio stesso. Questo è quanto è stato osservato in Cina, in Austria e in Germania.

Il modello SIR è stato applicato ai dati forniti dalla protezione civile. In alcune regioni l'analisi risulta piuttosto accurata, nel senso che i dati osservati e le previsioni del modello sono molto ben correlati.

La qualità e la quantità dei dati provenienti da Veneto ed Emilia permette una analisi particolarmente precisa.

Le due immagini mostrano gli andamenti dei casi non risolti ( $I$ ) e dei casi risolti ( $R$ ) per le due regioni, aggiornati al 15 aprile, assieme alla popolazione ridotta ( $S'$ ).

Vediamo che in entrambi i casi è stato raggiunto il massimo della curva arancio (casi di infezione in corso), la curva blu (popolazione suscettibile) è ridotta a circa il 5% di quella iniziale ( $S'$ ), e quella verde ( $R$ , casi risolti), sta crescendo rapidamente.

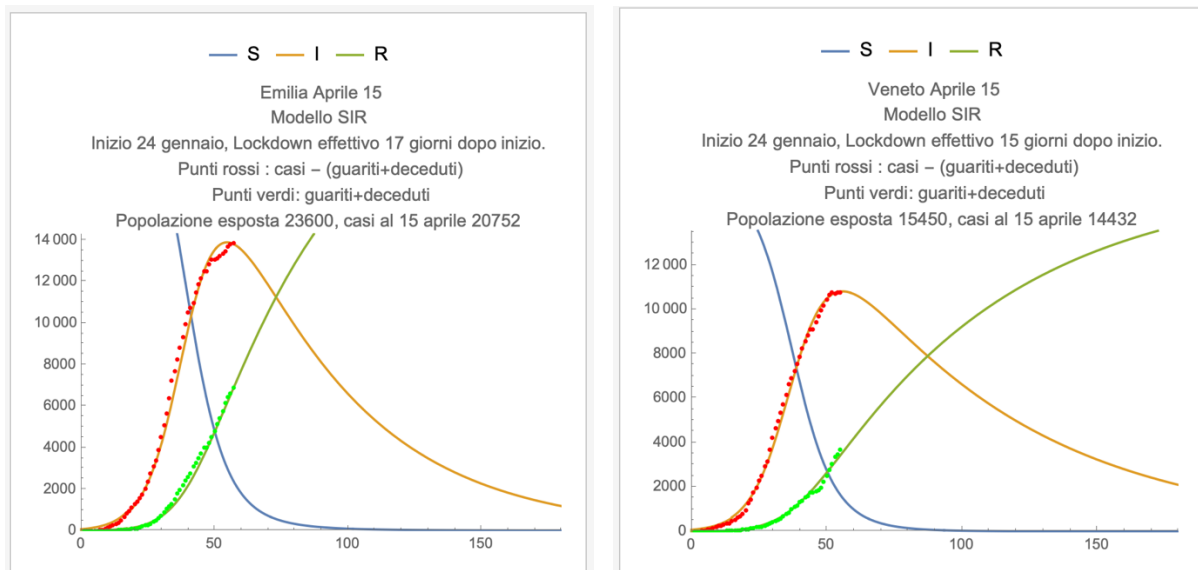


Figura 1a,b. Modello SIR basato sui dati dell'Emilia e del Veneto.

Il numero i casi totali hanno raggiunto circa il 95% della popolazione  $S'$ , da cui si può dedurre che nei giorni a seguire vi saranno nuovi casi di infezione, vale a dire passaggi dalla curva blu alla curva arancio, ( $S'$  non è ancora arrivato a 0) ma che il ritmo di casi risolti ogni giorno, (passaggi dalla curva arancio alla curva verde), sta superando il ritmo di nuovi casi di contagio (passaggi dalla curva blu alla curva arancio). Entro venti giorni circa i nuovi casi di contagio si azzereranno (perché sarà completamente esaurito  $S'$ ) e da lì in poi si tratta solo della risoluzione di casi esistenti, con il relativo passaggio dalla curva arancio alla curva verde fino ad esaurimento della curva arancio.

Risulta evidente, credo, come questo modello permetta di capire molto meglio la dinamica di quello che accade, dell'effetto del lockdown e della durata dell'epidemia in queste condizioni.

Analizzando il significato di queste curve si capisce come una serie di variabili che vengono usate regolarmente per valutare l'andamento e la data di riapertura nelle discussioni sulla stampa sono inappropriate.

Esempio #1: **il numero dei morti**, non è una quantità significativa, se non per il carico di dolore che si porta addosso. Il numero di morti che vi saranno in Emilia, in Veneto, è già fissato da molte settimane (a meno di cambiamenti/miglioramento delle cure mediche), come frazione fissa della curva verde che a regime raggiungerà (la curva verde) un valore molto vicino a  $S'$ , parametro determinato dai dati (la popolazione esposta equivalente,  $S'$ , risulta pari circa 23600 persone in Emilia e 15450 persone in Veneto e la mortalità osservata fino ad oggi nelle due regioni si manterrà sostanzialmente invariata fino alla fine di questa fase del contagio). Continueremo ad avere morti per parecchi mesi, purtroppo ci sono persone che sono morte dopo più di un mese di ospedalizzazione ma questo ha poco a che vedere con l'evoluzione del contagio.

Esempio #2: la **percentuale di crescita giornaliera dei casi sul totale dei casi**, non è una quantità particolarmente utile. Il modello prevede che fino a che non sarà esaurita la popolazione  $S'$ , avremo ogni giorno dei casi di infezione. Quale è il criterio epidemiologico

per cui ci si dovrebbe fermare ad un per cento, mezzo per cento o un per mille di nuovi casi giornalieri? Un per mille di 60.000 casi (regione Lombardia) sono 60 nuovi casi di infezione al giorno.

Esempio #3: **il punto a mezza altezza a sinistra della curva arancione** (che, ricordiamo, rappresenta  $I$ , il totale dei contagiati, sottratti i guariti ed i deceduti), rappresenta semplicemente il giorno in cui inizia a diminuire l'aumento giornaliero (sembra un gioco di parole, leggere bene). E' stato oggetto di molte discussioni un paio di settimane fa ma non è un punto particolarmente significativo.

Esempio #4: **il punto al massimo della curva arancione** (che, ricordiamo, rappresenta  $I$ , il totale dei contagiati, sottratti i guariti ed i deceduti), rappresenta semplicemente il giorno in cui inizia a diminuire il numero complessivo dei contagiati attivi.

Esempio #4: **Il punto in cui la curva arancione va a zero**, è un po' come aspettare l'ultimo morto, ci possono volere molti mesi e non ha nessun senso.

Osserviamo come cala lentamente la curva arancione: ha un tempo di decadimento che per alcune regioni è di 40 giorni, per altre di fino a 70-80 giorni. Dipende moltissimo dalla definizione di *guarito* e dal tipo di cure che vengono attivate per i pazienti negli ospedali, che possono anche prendere molte settimane. Che senso ha aspettare che la curva arancione si azzeri?

Il vero motore del virus sta piuttosto nell'andamento della curva  $S'$ , vale a dire il serbatoio di nuovi contagi che derivano dalla fase di lockdown. Innanzitutto la curva  $S'$  non va necessariamente a zero, ma raggiunge un valore piccolo, ma non nullo, dopo di che il contagio si spegne, perché non ci sono altri suscettibili da contagiare. Quindi eventualmente sarebbe da usare questa curva per analizzare quando il contagio sotto lockdown si spegne.

**In realtà il problema completamente un altro!** Noi guardiamo ai numeri del bollettino della protezione civile. Ma tutte quelle persone che fanno parte di  $S'$  sono sostanzialmente sotto controllo o lo possono essere. I contagiati, i guariti ed i morti sono certamente sotto controllo, isolamento o sono non più contagiosi. La frazione di  $S'$  che non è ancora contagiata sta diventando ogni giorno più piccola, lo dicono i grafici: la possiamo ricavare dall'ottimizzazione dei dati osservati: ad oggi in Emilia e Veneto questa frazione è già al 5%.

***E' su quella frazione che dovremmo concentrarci:***

- ***familiari o conviventi di persone ammalate che vivono in casa e non sono state separate per la quarantena,***
- ***operatori sanitari esposti al rischio di contagio,***
- ***operatori e lavoratori risposti al rischio di contagio per il tipo di attività che svolgono.....***

Se noi potessimo attivare per questa frazione residua di  $S'$  lo stesso controllo del resto ( $I$  e  $R$ ), potremmo riaprire anche domani: **infatti oggi, grazie a più di un mese di lockdown l'epidemia è attiva solo in una frazione della popolazione che è circa 300 volte più piccola dell'insieme della popolazione, sia a livello nazionale che regionale.** Il 95% di queste persone

sono già sotto stretto controllo. Dobbiamo cercare di estendere lo stesso controllo al 5% rimanente.

Nella misura in cui questo sarà possibile, allora potremmo parlare davvero di ripartenza.

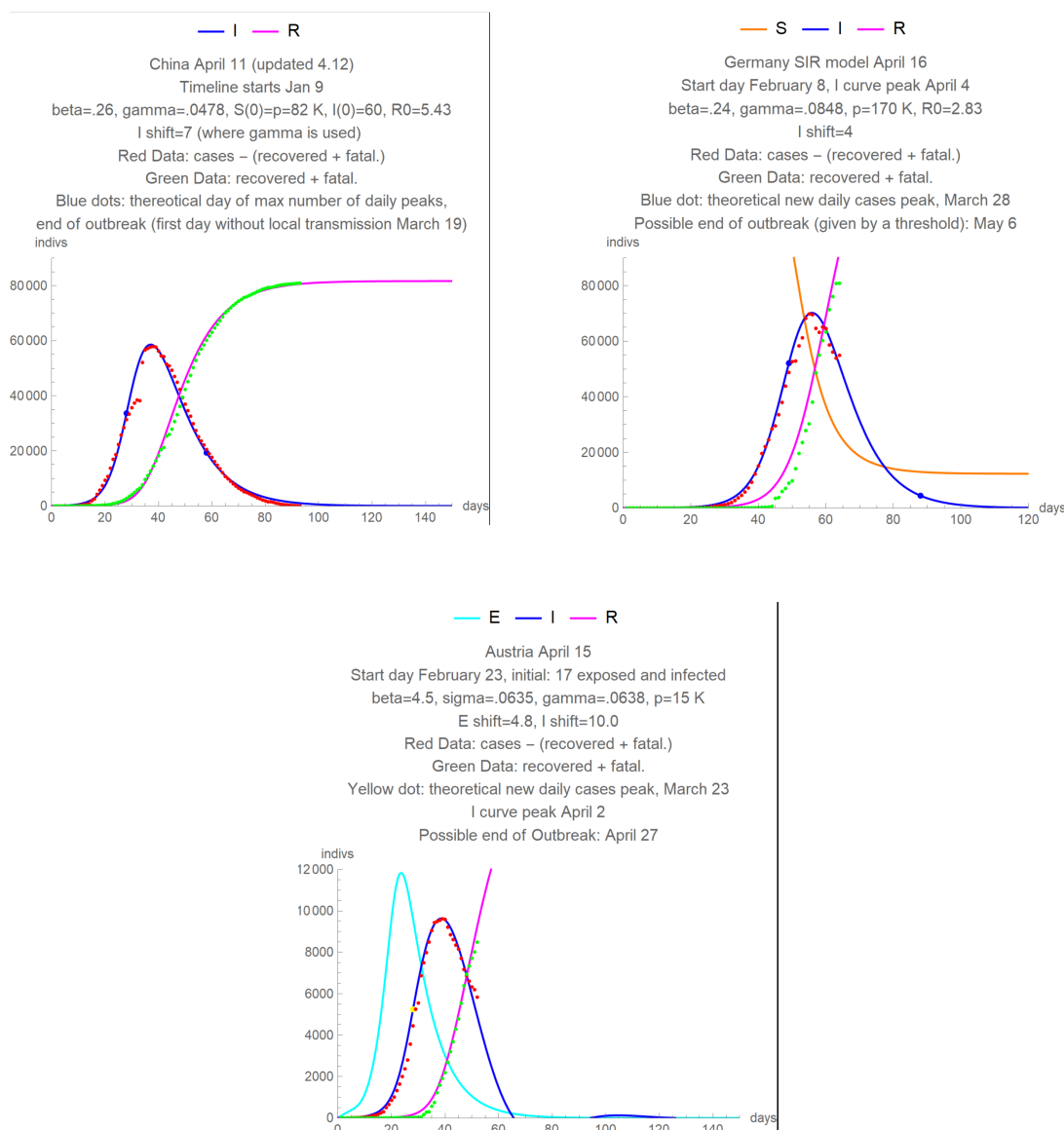


Figura 2a,b,c. Modello SIR basato sui dati della Cina, della Germania e dell' Austria (nel caso dell' Austria il modello usato è un modello a quattro equazioni, SEIR). Si tratta di 3 paesi che hanno spento l'epidemia tracciando e isolando sistematicamente tutta la popolazione a rischio contagio. La curva  $I$  in questo caso scende molto rapidamente, ma occorre ricordare che questo dipende anche dalla pratica medica che ha portato alla dimissione dei guariti. Notare i punti blu delle curve cinesi e tedesca: il punto sulla fase ascendente rappresenta il giorno con il massimo numero di nuovi contagi, il punto sulla fase discendente il primo giorno senza nuovi contagi (quando la popolazione  $S'$  si è esaurita). Si capisce bene dal caso della Germania come questo corrisponda al punto in cui la curva di  $S'$  ha raggiunto un valore minimo (ma non nullo - circa 18000 persone, pari al 10% di  $S'$ ): l'azione tedesca è così efficace da riuscire a limitare il contagio anche all' interno della popolazione ridotta  $S'$ . Questi andamenti appaiono diversi dal caso italiano, ma occorre sottolineare che la differenza è sostanzialmente dovuta all'andamento decrescente, che è fortemente correlato alle pratiche di tracciamento e di trattamento. In realtà, dal punto di vista del raggiungimento del plateau di  $S'$ , momento terminale dell' epidemia, non sono molto diverse