

Previsioni a medio termine dell'andamento dell'epidemia di Covid-19

StatGroup-19¹

In questi giorni il nostro paese è sottoposto a forti pressioni riguardanti la salute pubblica per l'azione epidemica del virus SARS-CoV-2. Presentiamo qui un approccio per la previsione a medio termine dell'andamento dell'epidemia.

Il GLM di Richards. Per ciascuna serie di indicatori pubblici cumulati (nuovi casi, guariti, ricoverati in terapia intensiva, deceduti, ecc., che indicheremo genericamente con Y), essendo conteggi, possiamo alternativamente ipotizzare una distribuzione di Poisson o, più credibilmente, una distribuzione Binomiale Negativa. In entrambe i casi, ipotizziamo per la media della variabile aleatoria al giorno t il seguente andamento:

$$E[Y] = b + (T - b) / [1 + 10^{(h(p-t))}]^s$$

dove b è un asintoto inferiore, T un asintoto superiore, h la pendenza (o tasso di crescita), p il punto di picco, ed s un parametro di asimmetria che regola la differente velocità di salita e discesa al picco per la derivata prima. Una versione estesa in cui è previsto un picco sostenuto nel tempo, tramite un operatore di *soft-thresholding*, è stata inoltre valutata e verrà discussa in un successivo documento tecnico. Inoltre, i cinque parametri sopra vengono poi riparametrizzati per l'inclusione di variabili esplicative e l'eliminazione di vincoli. Ad esempio per eliminare il vincolo $T > b$ e valutare l'effetto di una covariata x su T si parametrizza $T - b = r$ e $\log(r) = a + b x$.

La funzione di verosimiglianza viene massimizzata utilizzando un metodo numerico quasi-Newton.

Per il calcolo degli errori standard, e conseguentemente degli intervalli di confidenza, utilizziamo l'approssimazione numerica della matrice di informazione nella maniera usuale. La matrice di informazione viene inoltre utilizzata per il controllo dell'identificabilità locale.

Il codice *R* utilizzato per l'implementazione del modello verrà messo a disposizione nei prossimi giorni. Ne verrà data pubblicità sulla pagina StatGroup-19 e verrà aggiornato questo documento con il link al codice.

I Dati. I dati utilizzati sono i dati ufficiali condivisi dalla Protezione Civile, disaggregati per regione. Dal momento che i dati sui positivi giornalieri sono ormai eccessivamente distorti (fortemente selezionati sui casi con gravi sintomatologie, riferiti a pazienti il cui tampone sia stato effettuato da 1 a 10 giorni prima), ci focalizziamo sul conteggio dei ricoverati in terapia intensiva.

Risultati. Per il picco epidemico in Lombardia, utilizzando come indicatore il numero di ricoveri in terapia intensiva, l'ingresso nel periodo di massimo per la derivata prima viene previsto per il giorno 26/3 (22-30/3). Pertanto, ipotizzando un ritardo di 7-10 giorni tra il picco di ricoveri e il picco di casi positivi, possiamo attenderci che il picco per i nuovi casi giornalieri in Lombardia sia iniziato il 18/3 (12/3 - 23/3). Per tutte le altre aree e per la nazione, il punto di flesso viene stimato successivamente al 25/3, e pertanto non è attendibile in quanto non è possibile al momento tenere conto dell'effetto delle misure restrittive, che si potrà apprezzare non prima della fine del mese di Marzo.

1 **StatGroup-19:** Fabio Divino (Università del Molise), Alessio Farcomeni (Università di Roma Tor Vergata), Giovanna Jona Lasinio (Università di Roma La Sapienza), Gianfranco Lovison (Università di Palermo), Antonello Maruotti (LUMSA, Roma). Si ringrazia il Dr. Gabriele Sene (Banca D'Italia).