

Previsioni a breve termine dei posti occupati giornalmente in terapia intensiva da pazienti con 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 un tentativo di analisi e previsione a breve termine dei posti occupati giornalmente in terapia intensiva da pazienti con Covid-19. Per poter permettere una buona pianificazione delle risorse sanitarie, abbiamo validato il nostro approccio su base regionale e per un orizzonte temporale di tre/cinque giorni.

Il predittore ensemble. Combiniamo due metodi previsivi. Il primo è ottenuto tramite un GLMM di Poisson, con intercetta e trend quadratico casuali e distribuiti come una Gaussiana multivariata con matrice di covarianza, in cui gli elementi fuori dalla diagonale per intercetta e componente quadratica, e per componenti quadratica e lineare, sono nulli. Abbiamo inoltre utilizzato i residenti al 31/12/2019 come offset. I dati regionali sono quindi utilizzati come un unico panel, sebbene potenzialmente ogni regione abbia un suo trend e una sua intercetta nella riparametrizzazione logaritmica del parametro della Poisson. Il secondo metodo utilizza modelli INAR(1) non-stazionari con innovazioni di Poisson, valutati separatamente per ciascuna regione. Il modello stazionario, quello con trend lineare, e quello con trend quadratico sono confrontati tramite BIC per ciascuna regione; e la previsione regionale è quella corrispondente al modello con BIC minimo. Per la Lombardia e l'Emilia Romagna si considera inoltre un termine cubico. Per limitare l'influenza di effetti di lungo periodo, utilizziamo solamente le osservazioni giornaliere degli ultimi 18 giorni.

Le previsioni sono poi combinate in maniera convessa. Il peso è scelto ri-stimando il modello lasciando fuori l'ultima osservazione, e minimizzando la somma dei valori assoluti dell'errore di previsione sull'ultimo tempo. Gli intervalli di previsione al 99% sono ottenuti con bootstrap non-parametrico, aggregati tramite l'operatore unione. L'approccio ha ampi spazi per il miglioramento ed è puramente approssimativo, non tiene conto infatti del reale processo generatore dei dati (che è probabilmente un INAR(1) non-stazionario disomogeneo, con innovazioni Poisson-Binomiali, a parametri condivisi) perché le informazioni pubbliche non permetterebbero di stimare questa specificazione.

Abbiamo validato la tecnica negli ultimi sette giorni, con successo, su tutte le 20 regioni. In ciascun giorno producevamo la previsione la mattina, e la sera verificavamo che il dato osservato cadesse nell'intervallo di previsione. Ciò nonostante, data la scarsità di informazioni a nostra disposizione e la mispecificazione del modello, è plausibile che qualche intervallo di confidenza possa non coprire il valore vero; anche a seguito ad esempio dell'imprevedibile trasferimento di pazienti da una regione all'altra, di cluster di comportamenti a rischio avvenuti nei giorni passati, o di onsets di efficacia delle misure restrittive.

I Dati. I dati utilizzati sono i dati ufficiali condivisi dalla Protezione Civile, disaggregati per regione.

Risultati. La tabella è pubblicata al seguente link:

<http://afarcome.altervista.org/ICU.predictions.html>

e verrà aggiornata, con le nuove previsioni, a quell'indirizzo. La tabella riporta, a titolo di esempio, i ricoveri in terapia intensiva previsti per il giorno 18.3, per regione, con relativi intervalli di confidenza; insieme, a titolo esemplificativo, alla disponibilità di posti.

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)