

Esercizio di Analisi dei Dati sui Contagi COVID-19

Obiettivo: Questo esercizio ha come scopo l'analisi dell'andamento dei contagi COVID-19 in una regione prestabilita, utilizzando i dati forniti dalla Protezione Civile Italiana (utilizzare il file "dpc-covid19-ita-regioni.csv" contenuto in "dati-regioni").

Parte 1: Analisi dell'Andamento dei Nuovi Contagi

1. Scaricare il dataset dei contagi COVID-19 dal repository della Protezione Civile.
2. Graficare l'andamento dei nuovi contagi giornalieri (flag 'nuovi_positivi') per la regione prestabilita.

Parte 2: Implementazione dei Modelli Epidemiologici

Nell'analisi epidemiologica, particolarmente in situazioni di focolai infettivi come la pandemia COVID-19, è fondamentale disporre di modelli matematici per prevedere e comprendere l'andamento della diffusione del virus. Due modelli, benché semplici, comunemente utilizzati in questo contesto sono la funzione logistica e la funzione di Gompertz.

- **Funzione Logistica:**
Questa funzione è spesso impiegata per modellare la crescita di una popolazione in un ambiente dove le risorse sono limitate. In epidemiologia, la funzione logistica può rappresentare l'andamento di un'epidemia in cui l'aumento del numero di casi (o di un'altra variabile di interesse) inizialmente cresce in maniera esponenziale, ma poi rallenta man mano che la popolazione suscettibile si riduce o intervengono misure di contenimento. La forma della curva logistica, caratterizzata da una fase iniziale di rapida crescita seguita da una stabilizzazione, è coerente con molti andamenti epidemici osservati nella realtà.
- **Funzione di Gompertz:**
Questa funzione, usata originariamente in biologia per modellare la crescita di tumori o di popolazioni animali, è stata adattata per descrivere l'andamento delle epidemie. La funzione di Gompertz è particolarmente utile per modellare la fase di decrescita dei contagi, in cui il ritmo di diffusione del virus rallenta ma non si arresta improvvisamente. È quindi efficace nell'analizzare le fasi più avanzate di un'epidemia, dove la crescita non è più esponenziale ma segue un andamento più graduale.

Matematicamente vengono descritte da:

$$\text{Eq. Logistica: } N(< t) = \frac{N_{\text{tot}}}{1 + e^{-k(t-t_0)}}$$

$$\text{Eq. di Gompertz: } N(< t) = N_{\text{tot}} e^{-\ln(N_{\text{tot}}) e^{-k(t-t_0)}}$$

Dove i parametri N_{tot} , k , e t_0 descrivono rispettivamente il totale dei contagi a fine ondata epidemica, la velocità di propagazione del contagio, e da quanto tempo l'epidemia è in corso (il giorno di contagio del paziente 0).

Scrivere delle funzioni in Python per implementare:

- La funzione logistica
- La funzione di Gompertz
- I nuovi contagi giornalieri predetti dalle due funzioni (la differenza dei contagi predetti tra il giorno $i+1$ ed il giorno i).
- Utilizzare queste funzioni per modellare l'andamento dei nuovi contagi, analizzando come i contagi cambiano al variare dei parametri di queste funzioni.

Parte 3: Analisi della Prima Ondata

1. Concentrarsi sull'andamento dei contagi nella regione prestabilita durante i primi 35 giorni dalla comparsa dei primi casi (30 marzo 2020).
2. Supporre che i nuovi contagi giornalieri siano una realizzazione poissoniana con un valore di aspettazione dato dalla funzione logistica (e di Gompertz).
3. Scrivere una funzione per la likelihood poissoniana e utilizzarla per vincolare i valori dei parametri del modello.
4. Confrontare le previsioni del modello con i dati effettivi, propagando l'incertezza teorica associata al modello. Analizzare come la predizione fatta dopo i primi 35 giorni possa corrispondere all'andamento osservato nei primi 100 giorni (sia giornaliero sia cumulativo).
5. Commentare i risultati e i limiti dell'analisi

Parte 4: Modellizzazione dell'Intervallo 3 Giugno 2020 - 3 Giugno 2021

1. Considerare l'intervallo temporale tra il 3 giugno 2020 e il 3 giugno 2021.

2. Alla luce di quanto verificato precedentemente, come modellereste l'andamento epidemiologico in questo periodo?

Parte 5: Codice e Documentazione

1. Gli studenti devono scrivere un codice per eseguire le analisi richieste e commentarlo adeguatamente. Il codice può anche essere scritto e commentato tutto all'interno di un Jupyter Notebook.
2. Caricare il codice su un repository GitHub personale, fornendo una documentazione chiara sulle procedure seguite e i risultati ottenuti.