

# Esercitazione Informatica

Federico Bellisardi

13 Ottobre 2023

## 1 Istruzioni di Consegna

Inviare all'indirizzo mail federico.bellisardi2@unibo.it una relazione in formato **pdf** nominato

cognome\_nome\_corsodilaurea\_esercitazione\_informatica.pdf

dove dovete sostituire a *cognome* il vostro cognome, a *nome* il vostro nome, a *corsodilaurea* il vostro corso di laurea ("CTF" oppure "biotec") con i grafici ottenuti e le risposte alle domande.

Inoltre inviare il link al foglio di calcolo utilizzato per ottenere i grafici e per fare i calcoli.

**Consegnare entro sabato 2 Dicembre.**

## 2 Esercitazione

Collegarsi al sito del ministero della salute:

<https://github.com/pcm-dpc/COVID-19>

e scaricare i dati contenuti nel file `/dpc-covid19-ita-andamento-nazionale.csv`

### 2.1 Considerazioni Generali

1. Considerando i dati a disposizione (considerare solo la colonna "totale\_positivi"), rappresentare in un grafico l'andamento della pandemia.
2. Calcolare tramite le funzioni del foglio di calcolo quanti individui sono risultati positivi dall'inizio della pandemia complessivamente.
3. In che giorno è stato registrato il picco degli ospedalizzati? Ci sono state delle restrizioni in quel periodo da parte del governo?
4. In che periodo il rapporto tra gli individui dimessi/guariti e i nuovi positivi è stato massimo? Giustificare questa risposta con l'ausilio di un grafico.

### 2.2 Prima Ondata

Utilizzare un nuovo Foglio Google, che chiameremo "Prima Ondata".

Consideriamo adesso la prima ondata della pandemia (Febbraio 2020-Marzo 2020) e prendiamo in considerazione le prime tre settimane di dati (24 Febbraio - 01 Marzo, 02 Marzo - 08 Marzo, 09 Marzo - 15 Marzo).

1. Per il periodo preso in esame e per ciascuna settimana, riportare in un grafico la colonna "totale\_casi" in funzione del tempo.
2. Per ciascuna settimana, utilizzando la funzione "linea di tendenza", modelliamo i dati con un andamento esponenziale del tipo:

$$n(t) = a * \exp(b * t) \tag{1}$$

dove  $n(t)$  è il numero totale dei casi e  $t$  è il tempo espresso in giorni <sup>1</sup>.  
 Ricavare le equazioni delle curve ottenute e ricavare il tempo di raddoppio nelle tre settimane prese in considerazione e riportare in una tabella i risultati ottenuti.

Risultati			
Settimane	$a$	$b$	$t_{double}$
I° Settimana			
II° Settimana			
III° Settimana			

3. In un nuovo foglio (da rinominare "Analisi Logaritmica"), rappresentare in scala logaritmica le curve dei contagi nelle 3 settimane analizzate.

Il coefficiente di determinazione  $R^2$  ( $R^2 \in [0, 1]$ ) rappresenta un indice che esprime quanto il modello matematico è accurato per rappresentare la popolazione di dati e, in particolare, tanto è più vicino a 1 migliore è il modello.

Per ciascuna settimana:

- Quale è la curva che meglio rappresenta questa popolazione di dati? Ovvero la curva per cui il coefficiente di determinazione è massimo<sup>2</sup>?
- Ricavare analiticamente l'equazione della curva in forma esplicita partendo dalla equazione 1 e calcolare nuovamente il tempo di raddoppio.
- Rappresentare i risultati ottenuti in una tabella:

<sup>1</sup>Nota: Considerare il primo giorno (24 febbraio) come il tempo  $t = 0$

<sup>2</sup>**Non** considerare la curva "polinomiale"

Risultati			
Settimane	$m$	$q$	$t_{double}$
I° Settimana			
II° Settimana			
III° Settimana			

## 2.3 Terza Ondata

Utilizzare un nuovo Foglio Google, che chiameremo "Terza Ondata".

Consideriamo la terza ondata della pandemia e prendiamo i dati dal 16 dicembre 2021 <sup>3</sup> al 29 dicembre 2021.

1. Al fine di ottenere un andamento esponenziale quanto più accurato, creare una nuova colonna in cui verrà calcolata la differenza tra la colonna dei "totale\_positivi" ed il valore costante 305653 <sup>4</sup> ("totale\_positivi\_offset")
2. Fare un grafico dell'intero periodo dei "totale\_positivi\_offset" in funzione del tempo discretizzato.
3. Assumendo un andamento esponenziale dei dati, stimare quanti individui erano positivi il primo gennaio 2022 considerando l'espressione per la funzione trovata.

## 2.4 Analisi Statistica

Consideriamo ora 5 gruppi composti da almeno 2 settimane estratte in maniera casuale dalla popolazione covid, costituendo così 5 campioni differenti  $Y_i$ , con  $i = [1, \dots, 5] \in \mathbb{N}$ .

Per ciascun campione  $Y_i$

1. rappresentare graficamente mediante un *istogramma* la frequenza assoluta della colonna "nuovi positivi" e verificare che la somma di tutte le frequenze assolute costituisca la numerosità del campione.
2. (a) calcolare la **media pesata** del "totale\_ospedalizzati" utilizzando come pesi i valori nella colonna "terapia\_intensiva";  
(b) calcolare la **media aritmetica** di ciascun  $y_j$ ;

---

<sup>3</sup>Questa data corrisponderà al nuovo  $t = 0$

<sup>4</sup>Questo valore serve a rimuovere l'offset alla curva e rappresenta il totale dei positivi il 15 dicembre

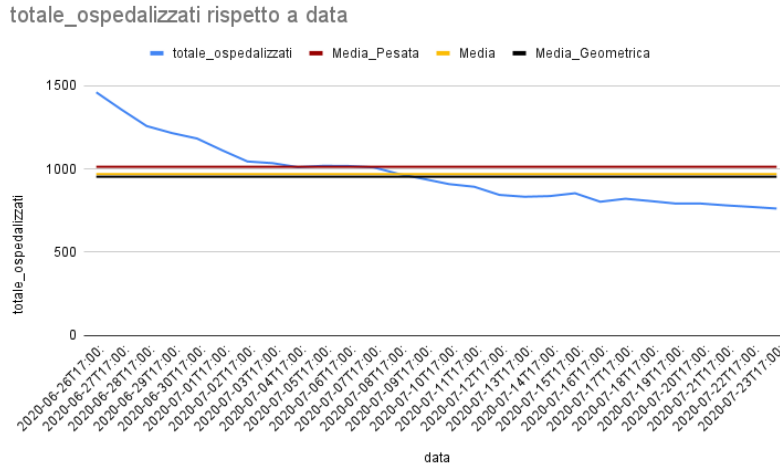


Figure 1: Rappresentazione grafica delle diverse medie per la popolazione  $Y_i$

- (c) infine calcolare la **media geometrica** della colonna "totale\_ospedalizzati" sapendo che la media geometrica si esprime come:

$$\sqrt[N]{y_1 \cdot \dots \cdot y_n} = \sqrt[N]{\prod_j y_j} = (\prod_j y_j)^{\frac{1}{N}} \quad (2)$$

3. calcolare infine la deviazione standard e la varianza.

I valori ottenuti dal calcolo delle medie nel punto (2) sono uguali? Giustificare la risposta rappresentando i risultati in un grafico del tipo in Fig.2.4.

Riportare tutto nella relazione.

## 2.5 Generazione di una Gaussiana

1. Utilizzando la definizione di *distribuzione normale o gaussiana*:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\} \quad (3)$$

partendo da un insieme di numeri  $x$  con cardinalità uguale a 1000 e distribuiti in maniera casuale<sup>5</sup>, tracciare il grafico della curva a campana usando un *grafico a dispersione*.

---

<sup>5</sup>Per generare numeri casuali usare la funzione *CASUALE()*.