

CORSO DI STATISTICA PER INGEGNERIA FISICA  
ANNO ACCADEMICO 2020/2021

ESERCITAZIONE 1: STATISTICA DESCRITTIVA

**Esercizio 1.** I gruppi sanguigni di 12 persone sono

$B, B, AB, 0, A, 0, A, A, A, B, A, A$

Si costruisca la tabella di distribuzione di frequenza ed il relativo diagramma a barre.

**Esercizio 2.** Un certo macchinario produce lotti da 100 pezzi ciascuno. I seguenti dati riportano il numero di pezzi difettosi presenti in 25 lotti ispezionati:

1, 5, 3, 1, 3, 2, 2, 1, 2, 5, 3, 0, 1, 4, 3, 7, 1, 3, 1, 7, 2, 1, 2, 4, 8

' Si costruiscano

- (a) una tabella di distribuzione di frequenza,
- (b) il relativo istogramma,
- (c) un boxplot.

Si determinino inoltre:

- (d) media,
- (e) quartili e 80-esimo percentile,
- (f) mode,
- (g) varianza,
- (h) deviazione standard,
- (i) differenza interquartile,
- (l) range dei dati osservati.

**Esercizio 3.** I seguenti dati si riferiscono alla temperatura (in  $^{\circ}C$ ) massima giornaliera raggiunta da un certo apparecchio nell'ultimo mese:

14,7   12,3   18,3   10,2   11,5   15,1   14,2   14,7   13,8   17,3  
25,3   26,4   31   19,4   17,5   17,6   16,8   18   13,8   10,7  
12,6   14,5   17,8   19,6   17,2   13,1   13,9   14,2   13,7   18,1

Si costruisca una tabella di distribuzione di frequenza e si disegni il relativo istogramma.

Si calcolino media, mediana e deviazione standard per i dati assegnati e per le stesse temperature espresse in  $^{\circ}F$  (*si ricorda che se  $t$  ed  $f$  rappresentano la medesima temperatura misurata in gradi Celsius e Fahrenheit rispettivamente, allora  $f = \frac{9}{5}t + 32$* ) [ $\bar{y} = 61.60^{\circ}F$ ,  $s_y = 8.27^{\circ}F$ ,  $q_2 = 58.82^{\circ}F$ ].

**Esercizio 4.** Negli ultimi 5 anni un certo macchinario ha necessitato di 11 interventi di manutenzione. Le cause del malfunzionamento sono state:

- arresto del macchinario per un guasto meccanico (5 volte)
  - arresto del macchinario per un guasto del sistema di controllo (4 volte)
  - produzione di un numero eccessivo di pezzi difettosi (2 volte)
- (1) Costruire la tabella della distribuzione di frequenza
  - (2) Rappresentare i dati graficamente.

**Esercizio 5.** In un'azienda il numero di dipendenti maschi e femmine è ripartito per età nel modo seguente:

Età	Femmine	Maschi
21 – 30	220	284
31 – 40	280	427
41 – 50	295	388
51 – 55	104	146
56 – 60	31	125

- (1) Quante e quali variabili stiamo considerando in questo set di dati? Che tipo di variabili sono? [*Stiamo considerando due variabili: “Età” e “Sesso”. “Età” è una variabile quantitativa e “Sesso” è qualitativa.*]
- (2) Vogliamo studiare la distribuzione dell'età dei dipendenti per entrambe le categorie della variabile “Sesso”. Costruire la tabella della distribuzione di frequenza dell'età per i dipendenti maschi e femmine (includendo: classi, frequenza assoluta, frequenza relativa, densità e frequenza cumulata).
- (3) Costruire l'istogramma delle frequenze relative per le due categorie di dipendenti e commentare il risultato.

**Esercizio 6.** Un operatore misura il pH di una soluzione per otto volte, utilizzando lo stesso strumento, e ottiene i seguenti dati:

7.15   7.20   7.18   7.19   7.21   7.20   7.17   6.50.

- (a) Calcolare media e varianza campionaria, mediana e IQR. [ $\bar{x} = 7.1$ ,  $s^2 = 0.05914$ ,  $m = 7.185$ ,  $IQR = 0.04$ ]

- (b) Ripetere il calcolo, escludendo l'ultima misurazione. Commentare i risultati ottenuti.  $\bar{x} = 7.1857$ ,  $s^2 = 0.00043$ ,  $m = 7.19$ ,  $IQR = 0.03$ . *Quantili e IQR sono indici di posizione e dispersione più robusti rispetto a media e varianza campionaria.*
- (c) Disegnare il boxplot sia dei dati completi, sia di quelli in cui avete escluso l'ultima misurazione.

**Esercizio 7.** Questi dati rappresentano la concentrazione di ferro presente in dieci campioni di sedimento del Lago di Como misurata in  $\mu g/l$ :

5.0, 1.5, 3.3, 5.1, 1.8, 3.2, 3.4, 3.4, 5.0, 2.5

Si calcolino: media, varianza, deviazione standard, mediana, quartili e ottantesimo percentile. Se  $m$  indica la mediana dei dati, per quali valori di  $k$  l'intervallo  $[m - k, m + k]$  contiene esattamente il 60% dei dati?  $\bar{x} = 3.42$ ,  $s^2 = 1.67$ ,  $s = 1.293$ ,  $Q_1 = 2.5$ ,  $Q_2 = 3.35$ ,  $Q_3 = 5$ ,  $q_{0.8} = 5$ ,  $1.55 \leq k < 1.65$

**Esercizio 8.** In una classe di 50 studenti le età dei ragazzi sono così distribuite:

Età	20	21	22	23	24	25	26	27
N.studenti	12	15	6	5	5	3	3	1

Determinare l'età media degli studenti, la deviazione standard dell'età, la mediana e la moda. Si stabilisca inoltre, senza rifare i calcoli, quali dei precedenti indici cambierebbero se avessimo 1 studente di 26 anni e 3 di 27.  $\bar{x} = 22.040$ ,  $s = 1.958$ ,  $Q_2 = 21$ ,  $\text{moda} = 21$ . *Gli indici che cambierebbero sarebbero media e deviazione standard.*

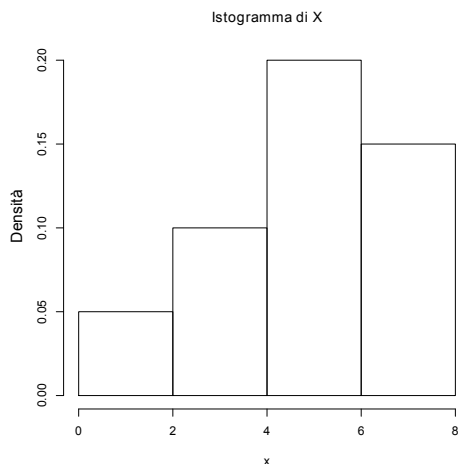
**Esercizio 9.** La seguente tabella sintetizza i dati raccolti nel 1798 dal fisico Henry Cavendish, relativi alla misura della densità della terra espressa come multiplo della densità dell'acqua.

Classi	Freq. Assoluta
(4,4.25]	1
(4.25,4.5]	0
(4.5,4.75]	0
(4.75,5]	1
(5,5.25]	1
(5.25,5.5]	14
(5.5,5.75]	9
(5.75,6]	3

- (a) Si rappresenti tramite istogramma la distribuzione dei dati raccolti.
- (b) Si descriva, sulla base dell'istogramma, la forma della distribuzione dei dati riguardanti la densità della terra. *[Distribuzione unimodale, asimmetria con coda a sinistra (possibile outlier).]*

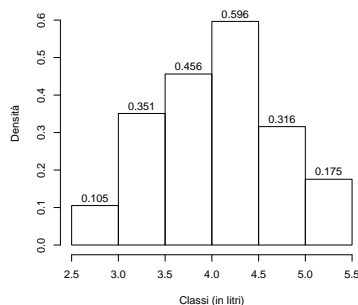
- (c) Si calcoli, in base alla tabella riportata, la media delle misurazioni della densità terrestre.  
 $\bar{x} = 5.435.$

**Esercizio 10.** L'istogramma seguente rappresenta la distribuzione di frequenza di una variabile  $X$ .



- (a) Si costruisca la tabella di distribuzione di frequenza di  $X$ .
- (b) Si determinino le classi contenenti i quartili di  $X$ .  $[Q_1 \in [2, 4), Q_2 \in [4, 6), Q_3 \in [6, 8)]$
- (c) Determinare una stima della media, del primo e terzo quartile.  $\bar{x} \simeq 4.8, Q_1 \simeq 3.5, Q_3 \simeq 6.33.$

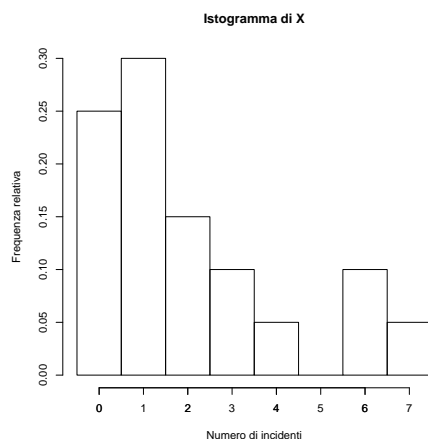
**Esercizio 11.** Il seguente istogramma rappresenta la distribuzione di frequenza del volume  $X$  della forza di espirazione (FEV: Forced Expiratory Volume), misurata in litri per 57 studenti di medicina. In ordinata è riportata la densità delle classi, ovvero il rapporto fra la loro frequenza relativa e la loro ampiezza.



- (a) Calcolare la percentuale di studenti la cui FEV è inferiore a 3.5 litri  $[22.8\%]$ .

- (b) Stabilire a quale classe appartiene la mediana di  $X$   $[4 \leq m \leq 4.5]$ .
- (c) Sulla base dell'istogramma, dire quale relazione ci si aspetta fra la media e la mediana di  $X$ . *L'istogramma non presenta particolari asimmetrie, quindi mi aspetto che media e varianza siano simili.*

**Esercizio 12.** Per ognuno dei 180 autisti di autobus di una azienda di trasporti municipale, è stato osservato il numero di incidenti  $X$  compiuti durante l'anno 2000. I risultati di questa indagine sono riassunti nel seguente istogramma:



Si calcolino:

- (a) la moda, la media e la mediana di  $X$ ;  $[moda = 1, \bar{x} = 2.05, m = 1]$
- (b) la deviazione standard di  $X$ ;  $[s = 2.1148]$
- (c) il primo e il terzo quartile di  $X$ ;  $[Q_1 = 0.5, Q_3 = 3]$
- (d) l'ottantacinquesimo percentile di  $X$ .  $[q_{0.85} = 5]$
- (e) Quanti autisti hanno fatto un numero di incidenti inferiori alla media?  $[126]$