

Statistica B – Esercitazione 1

Nel seguito, salvo diversamente specificato, indichiamo con \mathbf{x} una serie statistica di dimensione N e con \mathbf{y} la serie riordinata in senso crescente.

Esercizio 1. Dimostrare che

1. se N è dispari, allora la mediana è $y_{\frac{N+1}{2}}$ ed è unica;
2. se N è pari, allora la mediana non è necessariamente unica;
3. caratterizzare le serie statistiche che hanno una sola mediana e calcolarla.

Esercizio 2. Sia $\ell \in \{1, 2, 3\}$, si dimostri che

1. se $N/(4\ell) \notin \mathbb{N}$, allora il quartile ℓ -esimo è unico ed è uguale a $q_{\ell/4} = y_{\lfloor \frac{N}{4\ell} \rfloor + 1}$
2. se $N/(4\ell) \in \mathbb{N}$, allora $y_{\frac{N}{4\ell}}$ e $y_{\frac{N}{4\ell} + 1}$ sono ℓ -esimi quartili.
3. $q_0 = y_1$ e $q_1 = y_n$.

Esercizio 3. Siano \bar{x} la media aritmetica di \mathbf{x} e $\alpha, \beta \in \mathbb{R}$, dimostrare che

1. la media della serie statistica $z_i := \alpha + \beta x_i$ è $\alpha + \beta \bar{x}$;
2. $\sum_{i=1}^N (x_i - \bar{x}) = 0$;
3. se $m_{\mathbf{x}}(t) = \sum_{i=1}^N (x_i - t)^2$, allora $\sigma_{\mathbf{x}}^2 = \min_{t \in \mathbb{R}} m_{\mathbf{x}}(t)$ in cui $\sigma_{\mathbf{x}}^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$;
4. $\sigma_{\mathbf{x}}^2 = \frac{1}{N} \sum_{i=1}^N x_i^2 - (\bar{x})^2$

Esercizio 4. Data una variabile aleatoria reale X con distribuzione continua e di quadrato integrabile, siano $D_i(t) = \mathbb{E} \left[|X - t|^i \right]$, $i = 1, 2$. Dimostrare che

1. $\text{med}(X) = \operatorname{argmin}_{t \in \mathbb{R}} D_1(t)$ in cui $\text{med}(X)$ è la mediana della distribuzione della v.a. X ;
2. $\mathbb{E}[X] = \operatorname{argmin}_{t \in \mathbb{R}} D_2(t)$;

Esercizio 5. È stata svolta una indagine statistica su una popolazione di N guidatori e un giorno fissato. Per ognuno di essi è stata rilevata la velocità media di percorrenza del tragitto casa-lavoro, indicata con v_i , e il tempo di percorrenza del medesimo tragitto, indicato con t_i , $i = 1, \dots, N$.

Calcolare la velocità media

1. che lascia invariato il tempo totale di percorrenza della popolazione;
2. che lascia invariata la distanza totale percorsa dalla popolazione.

Esercizio 6. Per verificare la precisione di una bilancia è stato pesato 137 volte un oggetto di cui il peso è noto ed è uguale a 32. I risultati sono raccolti nella seguente tabella

Mid-points	29.25	29.75	30.25	30.75	31.25	31.75	32.25	32.75	33.25	33.75	34.25	34.75
n_i	1	2	9	15	17	25	23	23	14	6	1	1
N_i												
F_i												

1. Ricostruire le classi a partire dai mid-points sapendo che sono equispaziate (i mid-points o valori centrali sono i punti centrali delle classi);

2. completare la tabella con le frequenze cumulate assolute (N_i) e relative (F_i);
3. costruire la funzione di ripartizione empirica e tracciarne il grafico;
4. calcolare il quantili sulla base delle classi, con l'approssimazione continua (sfruttando quindi la fdr empirica) e confrontare i risultati;
5. calcolare media e varianza utilizzando i valori centrali.

Esercizio 7. In una popolazione sono state effettuate le seguenti osservazioni

Oss	1	2	3	4	5	6	7	8	9	10	11	12
x_i	0.754	3.828	15.957	6.429	8.124	5.525	1.078	3.459	1.014	15.743	7.098	4.898

1. Determinare quali tra i seguenti valori è la media della serie: (a) 6.159; (b) 16.001; (c) 0.712.
2. Individuare 3 classi e costruire la tabella delle frequenze.
3. Calcolare media e varianza sulla base della tabella delle frequenze e confrontare con il risultato ottenuto al punto 1.
4. Si rappresenti graficamente la tabella delle frequenze con il grafico che si ritiene più idoneo.

Esercizio 8. In 8 stazioni meteorologiche alla stessa ora sono state rilevate le seguenti temperature in gradi *Celsius* e in gradi *Fahrenheit*:

	1	2	3	4	5	6	7	8
°C	23	22	24	24	21	24	24	21
°F	73	72	75	75	70	75	75	70

1. Calcolare la media aritmetica e geometrica della temperatura con le due scale diverse;
2. Calcolare la varianza con le due diverse scale e confrontare i risultati.

Esercizio 9. In una popolazione sono stati rilevati i seguenti valori per un carattere statistico discreto:

19	20	21	22	23	24	25	26
2	12	5	4	4	4	2	2

1. Calcolare il valor medio e la varianza del carattere nella popolazione.
2. Se al posto delle prime tre modalità si fosse sostituito il loro valor medio con la medesima numerosità, ottenendo la seguente serie delle frequenze

20.16	22	23	24	25	26
19.00	4.00	4.00	4.00	2.00	2.00

come cambiano le risposte date alla precedente domanda?

Esercizio 10. 1. Data una serie statistica \mathbf{x} con N osservazioni, dimostrare che per ogni costante C

$$\bar{x} = C + \sum_{i=1}^N (x_i - C)$$

2. Calcolare la media della seguente serie statistica:

1252 1265 1278 1291 1304 1317 1330 1343 1356 1369 1382

3. Dimostrare che una distribuzione di frequenze simmetrica rispetto ad una modalità C ha media C .

Esercizio 11. È stata osservata la seguente serie statistica: $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5) = (10, 22, 13, 12, 60)$.

1. Calcolare la media e la mediana della serie e stabilire (motivandolo) quale si preferisce.
2. Supposto che $x_5 = 23$, si risponda alla precedente domanda.

Esercizio 12. Su un gruppo di lavoratori sono stati osservati i seguenti dati riguardanti la variabile reddito:

R_i	15000	18000	22000
n_i	n_1	n_2	n_3

Determinare per quali distribuzioni di frequenze:

1. la media aritmetica è 19000 oppure 22100;
2. la media aritmetica è 19200 e la varianza della serie è minima;
3. la media geometrica è 17000.

Esercizio 13. È stato rilevato su di un gruppo di lavoratori dipendenti il reddito da lavoro x_i e quello da capitale y_i e la serie osservata è la seguente

	1	2	3	4	5	6	7	8	9	10	11	12	13
x	15337	38003	21963	31312	19151	35591	25059	29779	24954	17087	18665	33601	16830
y	3312	1611	2377	2517	4845	2406	3022	2113	1827	2525	2079	2802	2089
r													

Sia $r_i = x_i + y_i$ il reddito totale dell' i -esimo lavoratore.

1. Calcolare i redditi medi da lavoro, capitale e totale.
2. Dimostrare che $\bar{r} = \bar{x} + \bar{y}$.
3. Verificare se la precedente affermazione è vera per medie di potenze e/o per la media geometrica.