Indici di posizione e di variabilità

Medie

Gli indici di posizione, o medie sintetizzano la posizione di una distribuzione di frequenza

Valore reale rappresentativo

Medie analitiche e posizionali

- Medie analitiche: operazioni su tutti i valori del carattere (media aritmetica, media quadratica, …)

- Medie posizionali: coinvolge solo alcuni valori del carattere (moda e mediana)

Concetto di media

- Media secondo Cauchy: la media x_M di una variabile X è un qualsiasi valore compreso tra il minimo ed il massimo

$$x_1 \leq x_M \leq x_n$$

- Media secondo Chisini: valore compreso tra il minimo ed il massimo tale che

$$f(x_1, x_2, ..., x_n) = f(x_M, x_M, ..., x_M)$$

La media aritmetica

<u>Media aritmetica</u>: esprime la posizione globale di una distribuzione di frequenza.

Sia

$$f(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} x_i$$

e μ la media aritmetica di una variabile X, per il criterio di Chisini:

$$f(x_1, x_2, ..., x_n) = f(\mu, \mu, ..., \mu)$$

La media aritmetica (II)

Ossia

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \mu = n\mu$$

Da cui

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

Media aritmetica ponderata

Date le modalità x1,x2,…,xk che si presentano con frequenze assolute n1,n2,…,nk, si definisce <u>media aritmetica ponderata</u> con pesi n1,n2,…,nk:

$$\mu = \frac{x_1 n_1 + x_2 n_2 + ... + x_k n_k}{n_1 + n_2 + ... + n_k} = \frac{\sum_{i=1}^{k} x_i n_i}{\sum_{i=1}^{k} n_i}$$

Difetti media aritmetica

Scarsissima resistenza ai valori eccezionali.

Un solo valore atipico può far variare la media aritmetica in misura elevatissima

Perdita di significato

Proprietà della media aritmetica

1. La somma algebrica degli scarti è sempre nulla

$$\sum_{i=1}^{n}\left(x_i - \mu\right) = 0$$

<u>Dimostrazione</u>

$$\sum_{i=1}^{n}\left(x_i - \mu\right) = \sum_{i=1}^{n} x_i - n\mu = \sum_{i=1}^{n} x_i - n\frac{\sum_{i=1}^{n} x_i}{n} = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i = 0$$

Esempio

Classi di età		Iscritti
3	15	115
15	25	156
25	40	130
40	50	110
50	60	90
60	70	38
Totale		639

x_i	x_{i+1}	$/x_i$	n_i	$/x_i n_i$
3	15	9	115	1035
15	25	20	156	3120
25	40	32,5	130	4225
40	50	45	110	4950
50	60	55	90	4950
60	70	65	38	2470
			639	20750

$$\overline{x_i} = \frac{x_i + x_{i+1}}{2}$$

$$\mu = \frac{20750}{639} = 32,47$$

Media geometrica

La media geometrica è la quantità G che, sostituita a ciascuna modalità, non ne altera il prodotto.

Sia

$$f(x_1, x_2, ..., x_n) = \prod_{i=1}^{n} x_i$$

Per il criterio di Chisini si ha:

$$f(x_1, x_2, ..., x_n) = f(G, G, ..., G)$$

quindi:

$$\prod_{i=1}^{n} x_i = G^n \Rightarrow G = \left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}} = \left(x_1 \cdot x_2 \cdot ... \cdot x_n \right)^{\frac{1}{n}} \Rightarrow G = \sqrt[n]{x_1 \cdot x_2 \cdot ... \cdot x_n}$$

Media quadratica

La <u>media quadratica</u> è la quantità Q che, sostituita a ciascuna modalità, non ne altera la somma dei quadrati

Sia f

$$f(x_1, x_2, ..., x_n) = \sum_{i=1}^{n} x_i^{\,2}$$

Per il criterio di Chisini:

$$f(x_1, x_2, ..., x_n) = f(Q, Q, ..., Q)$$

$$\sum_{i=1}^{n} x_i^{\,2} = \sum_{i=1}^{n} Q^2 = nQ^2 \Rightarrow Q = \sqrt{\frac{\sum_{i=1}^{n} x_i^{\,2}}{n}}$$

Moda

Moda: indice di posizione cui corrisponde la massima frequenza

Esempio: sia data la serie

13,13,15,16,16,17,18,18,19,22,23,23,23,26,26,45

La moda è il dato 23

Classi di modalità di differente ampiezza

La moda cade nella classe con maggiore densità di frequenza

$$M_0 = L_{M_0} + \left(\frac{\Delta_1}{\Delta_1 + \Delta_2} \right) c$$

L_{M0} confine inferiore della classe modale
Δ_1 eccesso della f modale sulla f della classe immediatamente inferiore
Δ_2 eccesso della f modale sulla f della classe immediatamente superiore
c ampiezza della classe modale

Esempio

Classi di età		Fumatori
30	33	2
34	37	3
38	41	9
42	45	19
46	49	29
50	53	17
54	57	10
58	61	7
62	65	4
Totale		100

$L_{M0}=45.5$
$\Delta_1=10$
$\Delta_2=12$
$c=4$

$$M_0 = 45.5 + \left(\frac{10}{10+12}\right)4 = 47.32$$

Mediana per caratteri discreti

Mediana: bipartisce la distribuzione ordinata in senso non decrescente delle modalità

Per caratteri discreti con modalità ordinate:

$$x_1 \leq x_2 \leq \ldots \leq x_{n-1} \leq x_n$$

Indice pari o dispari

1. n dispari → posto centrale

$$M_e = x_{\frac{n+1}{2}}$$

2. n pari → semisomma delle intensità dei posti centrali

$$M_e = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

Esempio n pari

Sia data la serie

13,15,15,16,16,17,18,18,19,22,23,23,23,26,26,45

n=16 pari

$$M_e = \frac{x_8 + x_9}{2} = \frac{18+19}{2} = 18.5$$

Esempio n dispari

- Sia data la serie

56,56,59,60,63,63,65,67,69,72,145

n=11 dispari

$$M_e = x_{\frac{n+1}{2}} = x_{\frac{11+1}{2}} = x_6 = 63$$