

MEDIE ANALITICHE (2° gruppo di indici di posizione)

medie che derivano da una manipolazione dei dati (serie di operazioni effettuate sui dati)

- MEDIA ARITMETICA
- MEDIA ARMONICA
- MEDIA GEOMETRICA
- MEDIA QUADRATICA
- MEDIA DI ORDINE k

Hanno senso per VARIABILI QUANTITATIVE

MEDIA ARITMETICA

Strumento fondamentale della statistica



costante di sintesi che non è solo utile come sintesi dell'informazione come indice di posizione riferita a una variabile, MA è necessaria e utile anche alla definizione di altre grandezze

- Strumento di SINTESI di fenomeni additivi (ha senso per la ^{suma} Σ)
- Sempre calcolabile sulle var. quantitative

• Distr. disaggregata: $x_1, \dots, x_N$

valore
media
(mu)

$$\mu = \frac{x_1 + \dots + x_N}{N} = \frac{1}{N} \sum_{i=1}^N x_i$$

• Distr. freq. assoluta: $(x_1, n_1), \dots, (x_k, n_k)$ (coppie) K modalità distinte ciascuna con la sua freq. assoluta

$$\mu = \frac{x_1 \cdot n_1 + \dots + x_k \cdot n_k}{N} = \frac{1}{N} \sum_{k=1}^K x_k \cdot n_k$$

• Distr. freq. relativa: $(x_1, f_1), \dots, (x_k, f_k)$

$$\mu = x_1 \cdot f_1 + \dots + x_k \cdot f_k = \sum_{k=1}^K x_k \cdot f_k$$

NB Nelle classi si fa la media di ogni classe e poi si moltiplica al rispettivo n_i , si fa la somma e si divide $\times N$

PROPRIETA' M. ARITH.

① Principio di Cauchy: $\alpha_{(1)} \leq \mu \leq \alpha_{(n)}$

② Principio di Chisini: $f(\alpha_1, \dots, \alpha_N) = \alpha_1 + \dots + \alpha_N = \overbrace{\mu + \mu + \dots + \mu}^{N \text{ volte}} = N\mu$

③ La somma degli scarti dalla media è NULLA:

$$\sum_{i=1}^N (\underbrace{\alpha_i - \mu}_{\text{SCARTO}}) = \sum_{i=1}^N \alpha_i - \sum_{i=1}^N \mu = N \cdot \underbrace{\frac{\sum_{i=1}^N \alpha_i}{N}}_{\substack{\text{def. di media aritmetica} \\ = \mu}} - N\mu = N\mu - N\mu = 0$$

④ Minimizzazione di una funzione



la MEDIA ARITMETICA è quel valore a che minimizza la somma dei quadrati degli scarti

$$g(a) = \sum_{i=1}^N (u_i - a)^2$$

Minimizzando $g(a)$ rispetto ad a si ottiene la derivata

$$\frac{\partial g(a)}{\partial a} = -2 \sum_{i=1}^N (u_i - a) = -2 \left(\sum_{i=1}^N u_i - Na \right) = \sum_{i=1}^N u_i - Na = 0$$

$\Rightarrow \sum_{i=1}^N u_i - Na = 0$

quindi:

$$\sum_{i=1}^N u_i - Na = 0 \quad \text{Da cui: } a = \mu$$

$$Na = \sum_{i=1}^N u_i \Rightarrow a = \frac{\sum_{i=1}^N u_i}{N} = \mu$$

→ per essere precisi si dovrebbe calcolare la derivata seconda per trovare il punto di minimo

⑤ la media aritmetica è un operatore lineare

2 caratteristiche

proprietà TRASLATIVITÀ



trasformata di posizione dei dati $y_i = a + u_i$ $a \in \mathbb{R}$

$$\mu_y = \frac{\sum_{i=1}^N (a + u_i)}{N} = \frac{Na + \sum_{i=1}^N u_i}{N} = a + \mu_x$$

quando si trasforma una variabile: $y = a + x$

$$\mu_y = a + \mu_x$$

proprietà di OMogeneITÀ

↓ x definire un operatore lineare

trasformata di scala dei dati $y_i = b u_i$ $b \in \mathbb{R}$

$$\mu_y = \frac{\sum_{i=1}^N (b u_i)}{N} = \frac{b \sum_{i=1}^N u_i}{N} = b \mu_x$$

$$y = b x \Rightarrow \mu_y = b \mu_x$$

le due propr. in x sono tali che per ogni trasformato lineare $y = a + b x$

$$\mu_y = a + b \mu_x$$



la media aritm. è un operatore lineare

⑥ Proprietà associativa

permette di mettere insieme informazioni che provengono da + gruppi di osservazioni (gruppi disgiunti)

Consideriamo L gruppi disgiunti di osservazioni

la media aritmetica è la media aritmetica delle medie aritmetiche dei singoli gruppi μ_ℓ , tenendo conto delle numerosità N_ℓ di ogni gruppo

$$\mu = \frac{\mu_1 N_1 + \dots + \mu_L N_L}{N_1 + \dots + N_L} \quad \ell = 1, \dots, L$$

MEDIA ARMONICA

μ_a

• media che si utilizza quando i fenomeni sono interpretabili come il reciproco di fenomeni di natura additiva

③ • NON sempre calcolabile: No se $\exists x=0$

• Distribuzione di soggegrata: $x_1, \dots, x_N$

$$\mu_a = \frac{N \text{ (tot. osservazioni)}}{\frac{1}{x_1} + \dots + \frac{1}{x_N}} = \frac{N}{\sum_{i=1}^N \frac{1}{x_i}}$$

• Distribuzione freq. assoluta: $(x_1, n_1), \dots, (x_k, n_k)$

$$\mu_a = \frac{N}{\frac{n_1}{x_1} + \dots + \frac{n_k}{x_k}} = \frac{N}{\sum_{k=1}^k \frac{n_k}{x_k}}$$

• Distribuzione freq. relativa: $(x_1, f_1), \dots, (x_k, f_k)$

$$\mu_a = \frac{1}{\frac{f_1}{x_1} + \dots + \frac{f_k}{x_k}} = \frac{1}{\sum_{k=1}^k \frac{f_k}{x_k}}$$

es. km/produzione / ...
velocità

Proprietà m. armonica

① Principio di Cauchy

② Principio di Chisini $f(x_1, \dots, x_N) = \frac{1}{x_1} + \dots + \frac{1}{x_N}$

③ Principio di omogeneità

↳ trasformati di scala dei dati $y_i = b x_i$ $b \in \mathbb{R}$

$$\mu_a y = \frac{N}{\frac{1}{b x_1} + \dots + \frac{1}{b x_N}} = \frac{N}{\sum_{i=1}^N \frac{1}{b x_i}} = \frac{b N}{\sum_{i=1}^N \frac{1}{x_i}} = b \mu_a x$$

④ Proprietà associative:

$\mu_a^{(e)}$ media armonica della popolazione l -esima, di dimensione $N^{(e)}$, $l=1, \dots, L$

$$\mu_a = \frac{N^{(1)} + \dots + N^{(L)}}{\frac{N^{(1)}}{\mu_a^{(1)}} + \dots + \frac{N^{(L)}}{\mu_a^{(L)}}}$$

media armonica delle medie armoniche

MEDIA GEOMETRICA:

μ_g

• si usa per fenomeni di natura moltiplicativa

fenomeni tali per cui le quantità che utilizziamo x il calcolo della media non sono indipendenti, sono una dipendente dall'altra

• le osservazioni devono essere > 0

se $= 0$ $\mu_i = 0$ la media è nulla

• Distribuzione disaggregata: $x_1, \dots, x_N$

$$\mu_g = \sqrt[N]{x_1 \cdot \dots \cdot x_N} = \sqrt[N]{\prod_{i=1}^N x_i}$$

predittoria

• Distribuzione di freq. assoluta: $(x_k, n_k), \dots, (x_K, n_K)$

$$\mu_g = \sqrt[n_1]{x_1^{n_1} \cdot \dots \cdot x_K^{n_K}} = \sqrt[n_1]{\prod_{k=1}^K x_k^{n_k}}$$

• Distribuzione di freq. relativa

$$\mu_g = x_1^{g_1} \cdot \dots \cdot x_K^{g_K} = \prod_{k=1}^K x_k^{g_K}$$

PROPRIETÀ M. GEOMETRICA

① Principio di Cauchy

② Principio di Chisini, $f(x_1, \dots, x_N) = x_1 \cdot \dots \cdot x_N$

③ Proprietà di omogeneità

↳ Trasformato di scalari dei dati $y_i = b x_i$, $b \in \mathbb{R}$

$$\mu_g y = \sqrt[n_1]{b x_1^{n_1} \cdot \dots \cdot b x_N^{n_N}} = b \mu_g x$$

④ Proprietà associativa

$\mu_g^{(e)}$ media geometrica della popolazione l -esima di dimensione $N^{(e)}$, $l=1, \dots, L$

$$\prod_{i=1}^N x_i = x_1 \cdot x_2 \cdot \dots \cdot x_N$$

$$\prod_{i=1}^N a x_i = a x_1 \cdot a x_2 \cdot \dots \cdot a x_N = a^N \cdot x_1 \cdot x_2 \cdot \dots \cdot x_N = a^N \prod_{i=1}^N x_i$$

(ricorda $\sum a x_i$)

$$\prod_{i=1}^N a = a^N \prod_{i=1}^N 1 = a^N \prod_{i=1}^N 1 = a^N$$

($\sum = a N$)

$\prod_{i=1}^N (a + x_i)$ NON si spezza!

$$\prod_{i=1}^N (x_i \cdot y_i) = \prod_{i=1}^N x_i \cdot \prod_{i=1}^N y_i$$

ATTENZIONE!

$$\prod_{i=1}^N (x_i + y_i) \neq \prod_{i=1}^N x_i + \prod_{i=1}^N y_i$$

$$\mu_g = \sqrt[N]{\mu_g^{(1)} \cdots \mu_g^{(N)}}$$

NB: se i valori di x sono troppo alti:

$$\ln \mu_g = \frac{1}{N} \sum_{i=1}^N \ln(x_i)$$

da cui:

$$\mu_g = \exp \{ \ln \sqrt[N]{x_1 \cdots x_N} \} = \exp \left\{ \frac{1}{N} \sum_{i=1}^N \ln(x_i) \right\}$$

MEDE POTENZIATE $\Rightarrow$ famiglia di medie di ordine r , $r \in \mathbb{R}$

$$\mu^{(r)} = \sqrt[r]{\frac{1}{N} \sum_{i=1}^N x_i^r} \quad (\text{espressione generale})$$

Casi particolari

- $r=1 \Rightarrow$ media aritmetica μ
- $r=-1 \Rightarrow$ media armonica μ_a
- $r \rightarrow 0 \Rightarrow$ media geometrica μ_g
- $r \rightarrow -\infty \Rightarrow \mu_{(1)}$
- $r \rightarrow +\infty \Rightarrow \mu_{(N)}$

$\hookrightarrow$ ordine
decrescente:

$$\mu_{(1)} \leq \dots \leq \mu_a \leq \mu_g \leq \mu \leq \dots \leq \mu_{(N)}$$

NB: $\mu^{(r)}$ è una funzione NON decrescente di r

al crescere di r il valore della media potenziata cresce o al più resta uguale

MEDIA QUADRATICA: è una media potenziata di ordine 2

$\mu^{(2)}$

$$\mu^2 = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}$$

• Distribuzione disaggregata: $x_1, \dots, x_N$

$$\mu^2 = \sqrt{\frac{x_1^2 + \dots + x_N^2}{N}}$$

• Distribuzione di freq. assoluta: $(x_1, n_1), \dots, (x_k, n_k)$

$$\mu^2 = \sqrt{\frac{x_1^2 n_1 + \dots + x_k^2 n_k}{N}} = \sqrt{\frac{\sum_{i=1}^k x_i^2 n_i}{N}}$$

• Distribuzione di freq. relative: $(x_1, f_1), \dots, (x_k, f_k)$

$$\mu^2 = \sqrt{x_1^2 f_1 + \dots + x_k^2 f_k} = \sqrt{\sum_{i=1}^k x_i^2 f_i}$$

(Rientra nelle def. di indici che vedremo + avanti.)

MEDIE POTENZIATE PESATE

$\mu^{(r)}$

ad ogni osservazione x_i è associata una misura di importanza, un peso collegato alla natura dei dati.

↓

$x_1, \dots, x_k$ con pesi $w_1, \dots, w_k$

allora:

$$\mu^{(r)} = \sqrt[r]{\frac{\sum_{i=1}^k x_i^r w_i}{\sum_{i=1}^k w_i}}$$