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 r

Hanno senso per VARIABILI QUANTITATIVE

MEDIA ARITMETICA

μ

- 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 x la definizione di altre grandezze
- Strumento di SINTESI di fenomeni additivi (ha senso fare la somma)
- Sempre calcolabile nelle var. quantitative

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

valore medio (mu)

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

- Distr. freq. assolute : $(\underset{\text{modalità}}{x_1}, \underset{\text{freq. assoluta}}{n_1}), \dots, (x_K, n_K)$ (coppie) K modalità distinte ciascuna con la sua freq. assoluta

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

- Distr. freq. relative : $(\underset{\text{modalità}}{x_1}, \underset{\text{freq. relative}}{f_1}), \dots (x_K, f_K)$

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

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 x N

PROPRIETÀ M. ARITM.

1. Principio di Cauchy $x_{(1)} \leq \mu \leq x_{(N)}$
2. Principio di Chisini : $f(x_1, \dots x_N) = x_1 + \dots + x_N = \overbrace{\mu + \mu + \dots + \mu}^{N \text{ volte}} = N\mu$
3. La somma degli scarti dalla media è NULLA

$$\sum_{i=1}^{N} \underbrace{(x_i - \mu)}_{\text{SCARTO}} = \underbrace{\sum_{i=1}^{N} x_i}_{N \sum_{i=1}^{N} \frac{x_i}{N}} - \underbrace{\sum_{i=1}^{N} \mu}_{\mu + \mu + \dots + \mu = N\mu} = N \underbrace{\frac{\sum_{i=1}^{N} x_i}{N}}_{= \mu} - N\mu = N\mu - N\mu = 0$$

def. di media aritmetica

4) Minimizzazione di una funzione

⇓

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

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

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

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

$\cancel{-2}\left(\sum_{i=1}^{N} x_i - Na\right) = 0$

quindi

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

↓

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

→ 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

propr. di TRASLATIVITÀ

⇓

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

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

⇓

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

⇓

$\mu_y = a + \mu_x$

propr. di OMOGENEITÀ

⇓ x definire un operatore lineare

trasformata di scala dei dati $y_i = b x_i \quad (b \in \mathbb{R})$

⇓

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

⇓

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

le due propr. inx sono tali che per ogni trasformata lineare $y = a + bx$

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

⇓

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

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

MEDIA ARMONICA → μ_a

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

NB: • NON sempre calcolabile: No se $\in X = 0$

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

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

• Distribuzione freq. assoluta: $(x_1, n_1) \dots (x_K, n_K)$

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

es. km / produzione / ... velocità

• Distribuzione freq. relativa $(x_1, f_1) \dots (x_K, f_K)$

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

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

$$\mu_{a,y} = \frac{N}{\frac{1}{b x_1} + \dots + \frac{1}{b x_N}} = \frac{N}{\sum_{i=1}^{N} \frac{1}{b x_i}} = \frac{bN}{\sum_{i=1}^{N} \frac{1}{x_i}} = b \mu_{a,x}$$

④ Proprietà associative:

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

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

media armonica delle medie armoniche

il testo inverte l'ordinamento

MEDIA GEOMETRICA: • 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

μ_g

• le osservazioni devono essere > 0

↓

se $= 0$ $x_i = 0$ la media è nulla

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

$$\boxed{\mu_g = \sqrt[N]{x_1 \cdots x_N} = \sqrt[N]{\prod_{i=1}^{N} x_i}}$$ → produttoria

• Distribuzione di freq. assoluta: $(x_1, n_1), \dots, (x_K, n_K)$

$$\mu_g = \sqrt[N]{x_1^{n_1} \cdots x_K^{n_K}} = \sqrt[N]{\prod_{K=1}^{K} x_K^{n_K}}$$

• Distribuzione di freq. relativa

$$\mu_g = x_1^{f_1} \cdots x_K^{f_K} = \prod_{K=1}^{K} x_K^{f_K}$$

(-1) per trovarlo il testo

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

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

(ricorda $\sum_{i=1}^{N} a x_i$)

$\prod_{i=1}^{N} a = a^N$ $\quad \prod_{i=1}^{N}\prod_{j=1}^{M} a = \prod_{i=1}^{N} a^M = a^{N \cdot M}$

($\neq \sum$ = a N Ma)

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

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

PROPRIETÀ M. GEOMETRICA

① Principio di Cauchy

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

③ Proprietà di omogeneità

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

$$\boxed{\mu_{g,y} = \sqrt{b x_1 \cdots b x_N} = b \mu_{g,x}}$$

④ Proprietà associativa

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

⇓

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

NB: se i valori di x sono troppo alti:

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

da cui:

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

MEDIE POTENZIATE $\Rightarrow$ famiglie di medie di ordine r, $r \in \mathbb{R}$

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

(espressione generica)

Casi particolari

- $r = 1 \Rightarrow$ media aritmetica μ
- $r = -1 \Rightarrow$ media armonica μ_a
- $r \to 0 \Rightarrow$ media geometrica μ_g
- $r \to -\infty \Rightarrow x_1$
- $r \to +\infty \Rightarrow x_{(N)}$

ordine crescente: $x_{(1)} \le \ldots \le \mu_a \le \mu_g \le \mu \le \ldots \le x_{(N)}$

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

$\mu^{(2)}$

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

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

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

- Distribuzione di freq. assolute: $(x_1 n_1) \dots (x_K n_K)$

$$\mu^2 = \sqrt{\frac{x_1^2 n_1 + \dots + x_K^2 n_K}{N}} = \sqrt{\frac{\sum_{i=1}^{K} x_i^2 n_i}{N}}$$

- 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 PONDERATE / PESATE: ad ogni osservazione x_i è associata una misura di importanza, un peso collegato alla natura dei dati.

$\mu^{(r)}$

⇓

$x_1, \dots, x_K$ con pesi $w_1, \dots, w_K$

⇓

allora:

$$\mu^{(r)} = \sqrt{\frac{\sum_{i=1}^{K} x_i^r w_x}{\sum_{i=1}^{K} w_i}}$$