MEDIE ANALITICHE

Sia $\underline{y} = (y_1, y_2, \dots, y_N)$. "Qual è la media dei dati?" è una domanda non formulata completamente. Bisogna fornire anche un criterio di "mediazione": il criterio di mediazione è detto "funzione di circostanza" oppure "circostanza".

La funzione di circostanza ha come argomento la rilevazione dati; essa si designa, di solito, con

$$C(\underline{y}).$$

Come si ottiene la media rispetto alla

circostanza assegnata? Sia y la media da calcolare e sia

$$\underline{y}^* = (y_1, y_2, \dots, y_N).$$

Allora y è soluzione dell'"equazione di circostanza":

$$C(\underline{y}) = C(\underline{y}^*).$$

ESEMPIO 1

Assegnata $\underline{y} = (y_1, y_2, \dots, y_N)$ e assegnata la funzione di circostanza

$$C(\underline{y}) = \sum_{i=1}^{N} y_i ,$$

la media analitica corrispondente

è la media aritmetica.

Infatti,

$$C(\underline{y}) = C(\underline{y}^*) \iff \sum_{i=1}^{N} y_i = \sum_{i=1}^{N} y \qquad (1)$$

$$\iff \sum_{i=1}^{N} y_i = N y \qquad \Rightarrow$$

da cui discende

$$y = \frac{1}{N} \sum_{i=1}^{N} y_i \equiv \bar{y},$$

ovvero la media aritmetica.

$\square$

<u>ESEMPIO 2</u>

assegnata $\underline{y} = (y_1, y_2, \ldots, y_N)$ ottenute da un carattere positivo e assegnata la funzione di circostanza

$$C(\underline{y}) = \prod_{i=1}^{N} y_i ,$$

la media analitica corrispondente è la "media geometrica".

Infatti

$$C(\underline{y}) = C(\underline{y}^*) \iff \prod_{i=1}^{N} y_i = \prod_{i=1}^{N} y \qquad (2)$$

$$\iff \prod_{i=1}^{N} y_i = y^N$$

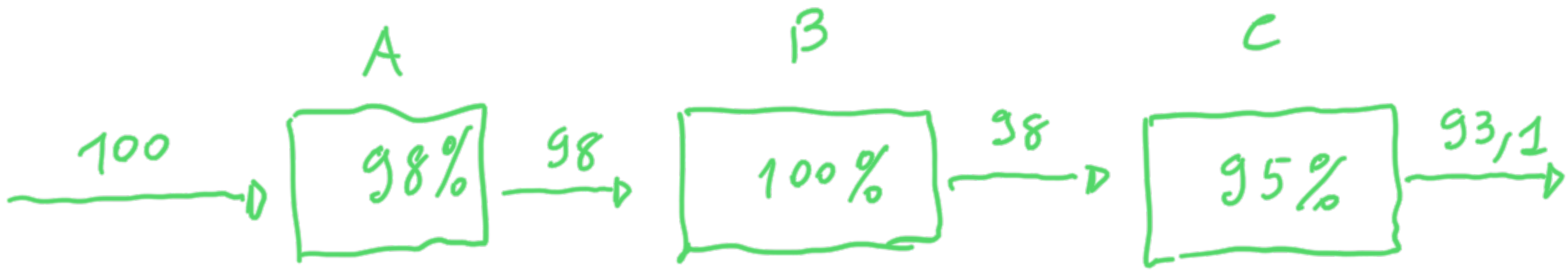
da cui discende

$$y = \sqrt[N]{\prod_{i=1} y_i} \equiv M_g ,$$

ovvero la media geometrica. $\square$

RENDIMENTO MEDIO DI DISPOSITIVI

DISPOSITIVO REALE



$$\underline{y} = (0{,}98,\ 1{,}00,\ 0{,}95)$$

$$C(\underline{y}) = 0{,}98 \cdot 1 \cdot 0{,}95 = 0{,}931$$

$$M_g(\underline{y}) = \sqrt[3]{0{,}98 \cdot 1 \cdot 0{,}95} = 0{,}97645\ldots$$

$$C(\underline{y}^*) = (0,97645\ldots)(0,97645\ldots)(0,97645\ldots) = 0,931.$$

DISPOSITIVO "MEDIO"

$$100 \longrightarrow \overset{A}{\boxed{0,97645\ldots}} \xrightarrow{97,645} \overset{B}{\boxed{0,97645\ldots}} \xrightarrow{97,645} \overset{C}{\boxed{0,97645\ldots}} \xrightarrow{0,931}$$

$\square$

ESEMPIO 3

Assegnata $\underline{y} = (y_1, y_2, \ldots, y_N)$ ottenuta da un carattere non nullo e assegnata la funzione di circostanza

$$C(\underline{y}) = \sum_{i=1}^{N} \frac{1}{y_i},$$

la media analitica corrispondente è la "media armonica".

Infatti:

$$C(\underline{y}) = C(\underline{y}^*) \iff \sum_{i=1}^{N} \frac{1}{y_i} = \sum_{i=1}^{N} \frac{1}{y} \qquad (3)$$

$$\iff \sum_{i=1}^{N} \frac{1}{y_i} = N \frac{1}{y}$$

$$\iff \frac{1}{y} = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{y_i}$$

da cui

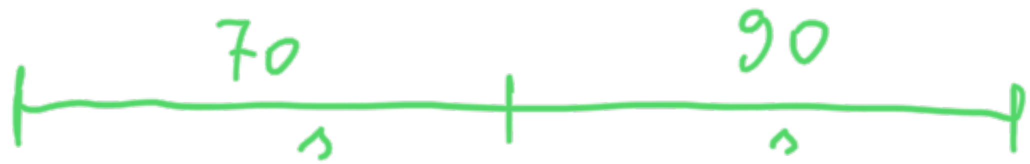
$$y = \left(\frac{1}{N} \sum_{i=1}^{N} \frac{1}{y_i} \right)^{-1} \equiv M_a,$$

che coincide con la media armonica:

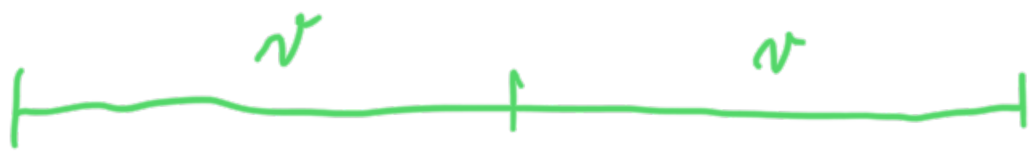
il reciproco della media aritmetica dei reciproci. $\square$

MEDIA DI DUE VELOCITÀ

$v_1 = 70$ Km/h, $v_2 = 90$ Km



dopo t_1 (unità orarie) si arriva a destinazione



dopo t_2 (unità orarie) si arriva a destinazione

In questo contesto la circostanza è

rappresentata dal tempo impiegato per percorrere i due tratti uguali a s:

$$C(v_1, v_2) = t.$$

Ma,

$$v = \frac{s}{t} \Rightarrow t = \frac{s}{v}$$

e quindi

$$t_1 = \frac{s}{v_1} + \frac{s}{v_2} \quad e \quad t_2 = \frac{s}{v} + \frac{s}{v}.$$

In definitiva,

$$t_1 = t_2 \Longleftrightarrow \frac{s}{v_1} + \frac{s}{v_2} = \frac{2s}{v} \Longleftrightarrow s\left(\frac{1}{v_1} + \frac{1}{v_2}\right) = s\,\frac{2}{v}$$

$$\Longleftrightarrow \frac{v}{2} = \left(\frac{1}{v_1} + \frac{1}{v_2}\right)^{-1} \Longleftrightarrow v = 2\left(\frac{1}{v_1} + \frac{1}{v_2}\right)^{-1}$$

$$\Longleftrightarrow \quad v = \left[\frac{1}{2}\left(\frac{1}{v_1} + \frac{1}{v_2}\right)\right]^{-1}$$

$$\Longleftrightarrow \quad v = \left[\frac{1}{2}\left(\frac{1}{70} + \frac{1}{90}\right)\right]^{-1} = 78{,}75 \text{ Km/h}. \qquad \square$$

Altri esempi di funzioni di circostanza sono

- la somma dei quadrati dei dati
- la somma dei quadrati dei dati rapportata alla somma dei dati
- la somma dei dati trasformati tramite la funzione esponenziale

tramite la funzione esponenziale

$$e^{y_1}, e^{y_2}, \dots, e^{y_N}. \qquad \square$$

STATISTICA INFERENZIALE

Supponendo che esistano finiti i valori medi considerati nel seguito, si ricordi che:

1) $\mathbb{E}(aX+b) = a\,\mathbb{E}(X) + b$;

2) $\mathbb{E}(X_1 + X_2) = \mathbb{E}(X_1) + \mathbb{E}(X_2)$;

3) $n \in \mathbb{N}$, $\mathbb{E}(X_1 + X_2 + \dots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \dots + \mathbb{E}(X_n)$

3) $$\mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i);$$

4) $\mathbb{D}^2(aX+b) = a^2\,\mathbb{D}^2(X)$.

DEFINIZIONE

DEFINIZIONE

La quantità

$$\mathrm{Cov}(X_1, X_2) = \mathbb{E}\left[(X_1 - \mu_{X_1})\cdot(X_2 - \mu_{X_2})\right].$$

$\square$

Se $X_2 = X_1$ allore

$$\mathrm{Cov}(X_1, X_1) = \mathbb{E}\left[(X_1 - \mu_{X_1})^2\right] = \mathbb{D}^2(X_1)$$

TEOREMA

Risulta:

5) $$\mathbb{D}^2(X_1 + X_2) = \mathbb{D}^2(X_1) + \mathbb{D}^2(X_2) + 2\,\mathrm{Cov}(X_1, X_2).$$

DIM

$$\mathbb{D}^2(X_1 + X_2) = \mathbb{E}\left\{\left[(X_1 + X_2) - (\mu_{X_1} + \mu_{X_2})\right]^2\right\}$$

$$= \mathbb{E}\left\{\left[(X_1 - \mu_{X_1}) + (X_2 - \mu_{X_2})\right]^2\right\}$$

$$= \mathbb{E}\left[(X_1 - \mu_{X_1})^2 + (X_2 - \mu_{X_2})^2 + 2(X_1 - \mu_{X_1})(X_2 - \mu_{X_2})\right]$$

$$= \mathbb{E}\left[(X_1 - \mu_{X_1})^2\right] + \mathbb{E}\left[(X_2 - \mu_{X_2})^2\right] + 2\,\mathbb{E}\left[(X_1 - \mu_{X_1})(X_2 - \mu_{X_2})\right]$$

$$= \mathbb{D}^2(X_1) + \mathbb{D}^2(X_2) + 2\,\mathrm{Cov}(X_1, X_2).$$

$\square$

DEFINIZIONE

X_1 e X_2 sono indipendenti se ogni evento generato da X_1 è indipendente da ogni evento generato da X_2 $\square$

PROPOSIZIONE

X_1 e X_2 sono indipendenti se e solo se

$$x_1, x_2 \in \mathbb{R}, \quad \mathbb{P}(X_1 \le x_1, X_2 \le x_2) = \mathbb{P}(X_1 \le x_1) \cdot \mathbb{P}(X_2 \le x_2).$$

$\square$

PROPOSIZIONE

Se X_1 e X_2 sono indipendenti allora

$$\mathrm{Cov}(X_1, X_2) = 0.$$

$\square$

COROLLARIO

Se X_1 e X_2 sono indipendenti, allora

6) $$\mathbb{D}^2(X_1 + X_2) = \mathbb{D}^2(X_1) + \mathbb{D}^2(X_2).$$

$\square$

Argomenti di probabilità e statistica
Rita Giuliano
Springer

STATISTICA DESCRITTIVA	STATISTICA INFERENZIALE PARAMETRICA
Ω : popolazione	Ω : insieme degli enti elementari
Y : carattere	X : genitrice
$\underline{y} = (y_1, y_2, \dots, y_N)$	$(X_1, X_2, \dots, X_n)$ campione

n-pla di variabili somiglianti a X e indipendenti | casuale semplice

$x_1, x_2, \dots, x_n$ | dati

$$\left(\Omega,\ \mathcal{F}=\sigma(\mathcal{G}),\ \mathbb{P}_\theta = ?\right)$$

L'unica cosa che si può dire è che

$\theta \in \Theta$ e Θ si chiama regione parametrica.

La statistica inferenziale parametrica mira a determinare delle formule per ottenere "stime" del parametro, o più in generale, di una funzione del parametro.

ESEMPIO

Si sa che una certa moneta è truccata con trucco incognito.

Allora X (la genitrice) ha legge $B(1, p)$ con

$$\mathbb{P}(T) = p \in (0, 1) \text{ incognita.}$$

La regione parametrica è l'intervallo $(0,1)$:

$$\Theta = (0, 1).$$

Si eseguono n lanci della moneta per

ottenere n repliche di X.

$$X_1, X_2, \ldots, X_{10} : \quad X_1 \sim X_2 \sim \cdots \sim X_{10} \sim X \sim B(1, p)$$

$t(X_1, X_2, \ldots, X_{10})$ è una v.a da determinare in maniera tale che

$$t(x_1, x_2, \ldots, x_{10})$$

è un numero che approssima la probabilità di T per la moneta considerata.

$\square$