TEOREMA DI BAYES

$(H_1, H_2, \dots, H_n) \subseteq \mathcal{F}$ è un s.c.a.

$i = 1, \dots, n, \quad \mathbb{P}(H_i) > 0$ ed è nota

$B \in \mathcal{F} : \quad \mathbb{P}(B) > 0$

B si è verificato

$i = 1, \dots, n, \quad \mathbb{P}_B(H_i) = ?$

Risulta

$$\mathbb{P}_{H_i}(B) \cdot \mathbb{P}(H_i) = \mathbb{P}(B \cap H_i) = \mathbb{P}_B(H_i) \cdot \mathbb{P}(B)$$

$$\mathbb{P}_{H_i}(B)\,\mathbb{P}(H_i) = \mathbb{P}(B \cap H_i) = \mathbb{P}_B(H_i)\,\mathbb{P}(B)$$

da cui

$$\mathbb{P}_B(H_i) = \frac{\mathbb{P}_{H_i}(B)\cdot \mathbb{P}(H_i)}{\mathbb{P}(B)}$$

$$= \frac{\mathbb{P}_{H_i}(B)\cdot \mathbb{P}(H_i)}{\sum_{j=1}^{n} \mathbb{P}_{H_j}(B)\cdot \mathbb{P}(H_j)}.$$

ESEMPIO

B : "una banconota da 20 euro falsa è nelle mie tasche"

Ci sono 3 tipografie sospettate di stampare banconote false da 20 euro.

	T_1	T_2	T_3
$\mathbb{P}(T_i)$	0,05	0,20	0,75
$\mathbb{P}_{T_i}(B)$	0,90	0,40	0,30

Inoltre, è stato valutato che

$$\mathbb{P}(T_1) = 0{,}05 \;; \quad \mathbb{P}(T_2) = 0{,}20 \;; \quad \mathbb{P}(T_3) = 0{,}75$$

$$\mathbb{P}_{T_1}(B) = 0{,}90 \;; \quad \mathbb{P}_{T_2}(B) = 0{,}40 \;; \quad \mathbb{P}_{T_3}(B) = 0{,}30$$

Applicando il teorema di Bayes, si ottiene:

$$\mathbb{P}_B(T_1) = \frac{\mathbb{P}_{T_1}(B) \cdot \mathbb{P}(T_1)}{\mathbb{P}_{T_1}(B) \cdot \mathbb{P}(T_1) + \mathbb{P}_{T_2}(B) \cdot \mathbb{P}(T_2) + \mathbb{P}_{T_3}(B) \cdot \mathbb{P}(T_3)}$$

$$= \frac{0,90 \cdot 0,05}{0,90 \cdot 0,05 + 0,40 \cdot 0,20 + 0,30 \cdot 0,75}$$

$$= \frac{0,045}{0,045 + 0,08 + 0,225} = \frac{0,045}{0,35}$$

$$= 0,128.$$

$$\mathbb{P}_B(T_3) = \frac{0,30 \cdot 0,75}{0,35} = \frac{0,225}{0,35} = 0,64$$

$$\mathbb{P}_B(T_2) = 1 - 0,128 - 0,64 = 0,23.$$

$\square$

ESEMPIO

Due tiratori sparano un colpo ciascuno su di un bersaglio. Siano 0,9 e 0,5 le probabilità che essi colpiscano il centro del bersaglio. Calcolare la probabilità che essi bersaglio. Calcolare la probabilità che, essendo stato colpito il centro del bersaglio una sola volta, l'autore di detto centro è il tiratore con la probabilità maggiore.

<u>Svolgimento</u>

Siano G_1 e G_2 i due tiratori con

$$\mathbb{P}(G_1) = 0,9 \quad e \quad \mathbb{P}(G_2) = 0,5.$$

Le alternative sono:

$\mathbb{P}_{H_1}(B) = 0$; H_1 : "due colpi che non fanno centro"

$\mathbb{P}_{H_2}(B) = 0$; H_2 : "due colpi a centro"

$\mathbb{P}_{H_3}(B) = 1$; H_3 : "G_1 centra e G_2 non centra"

$\mathbb{P}_{H_4}(B) = 1$ H_4 : "G_1 non centra e G_2 centra"

Sia B l'evento che solo uno dei due

tiratori abbia fatto centro.

$$\mathbb{P}_B(H_3) = \frac{\mathbb{P}_{H_3}(B)\cdot\mathbb{P}(H_3)}{\mathbb{P}_{H_1}(B)\,\mathbb{P}(H_1)+\mathbb{P}_{H_2}(B)\,\mathbb{P}(H_2)+\mathbb{P}_{H_3}(B)\,\mathbb{P}(H_3)+\mathbb{P}_{H_4}(B)\,\mathbb{P}(H_4)}$$

$$= \frac{1\cdot 0{,}9\cdot 0{,}5}{\underbrace{0\cdot 0{,}1\cdot 0{,}5}_{=0}+\underbrace{0\cdot 0{,}9\cdot 0{,}5}_{=0}+1\cdot 0{,}9\cdot 05+1\cdot 0{,}1\cdot 0{,}5}$$

$$= \frac{0{,}9\cdot \cancel{0{,}5}}{0{,}9\cdot\cancel{0{,}5}+0{,}1\cdot\cancel{0{,}5}} = \frac{0{,}9}{0{,}9+0{,}1} = 0{,}9$$

$$\mathbb{P}_B(H_4) = 1-\mathbb{P}_B(H_3) = 1-0{,}9 = 0{,}1.$$

RIEPILOGO

$\mathbb{P}(H_1) = 0{,}1\cdot 0{,}5 = 0{,}05$; $\mathbb{P}_B(H_1) = 0$

$\mathbb{P}(H_2) = 0{,}9\cdot 0{,}5 = 0{,}45$; $\mathbb{P}_B(H_2) = 0$

$\mathbb{P}(H_3) = 0{,}9\cdot 0{,}5 = 0{,}45$; $\mathbb{P}_B(H_3) = 0{,}90$

$$\mathbb{P}(H_4) = 0{,}1 \cdot 0{,}5 = 0{,}05 ; \quad \mathbb{P}_B(H_4) = 0{,}10 .$$

$\square$

MEDIE DI UNA VARIABILE ALEATORIA

Si consideri una rilevazione dati di un carattere quantitivo Y; $\underline{y} = (y_1, y_2, \dots, y_n)$

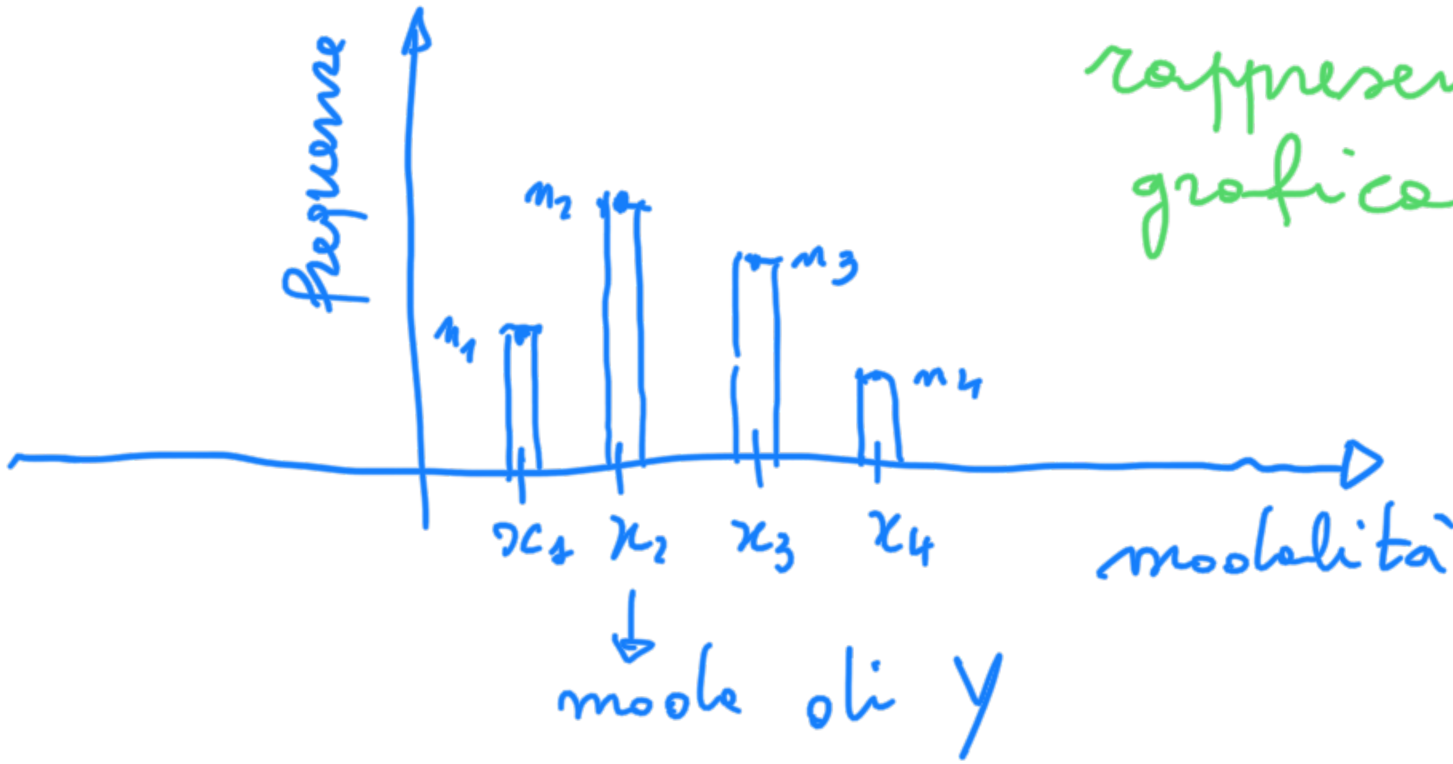
La media aritmetica della rilevazione dati è

$$\overline{y} = y_1 + y_2 + \dots + y_n .$$

Se per ottenere ciascun dato bisogna procedere con un conteggio, allora si potranno individuare $k \leq n$ numeri interi positivi $x_1, x_2, \dots, x_k$ che vengono detti "modalità" del carattere ognuno dei quali si presenta una o più volte nella rilevazione dati. Allora, se si individua il numero n_i che rappresenta quante volte la modalità x_i, per $i = (1, 2, \dots, k)$, le k coppie

$$(x_1, \underbrace{n_1), (x_2, n_2), \dots, (x_k, n_k}_{\text{frequenze}})$$

rappresentano in maniera sintetica la rilevazione dati $\underline{y}$. Se si riferisce la k-pla delle precedenti coppie rispetto ad un sistema di riferimento cartesiano si ottiene il diagramma delle frequenze



rappresentazione grafica di $\underline{y}$

Un altro modo per rappresentare la K-ple delle coppie modalità-frequenze è di tipo tabellare; esso è più conveniente dal punto di vista operazionale:

modalità	frequenze assoluta	relativa	intensità
x_1	n_1	$f_1 = n_1/n$	$i_1 = x_1 \cdot f_1$

rappresen-tazione tabellare di $\underline{y} = (y_1, \dots, y_n)$

x_2	n_2	$f_2 = n_2 / n$	$i_2 = x_2 \cdot f_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	n_k	$f_k = n_k / n$	$i_k = x_k \cdot f_k$
$\sum_{i=1}^{k} \longrightarrow$	n	1	$\bar{y}$

$n \downarrow$ "taglia"

$\bar{y} \downarrow$ media aritmetica

ESEMPIO

$$\underline{y} = (\overset{y_1}{2}, \overset{y_2}{5}, \overset{y_3}{3}, 2, 1, 3, 5, 2, 2, 3, 1, 4, 1, 1, 3, 2, 2, 3, 4, 4)$$

Quindi $n = 20$, $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, $x_4 = 4$, $x_5 = 5$, $K = 5$.

modalità x	frequenze ass. n	frequenze rel. f	intensità i

1	4	4/20	4/20
2	6	6/20	12/20
3	5	5/20	15/20
4	3	3/20	12/20
5	2	2/20	10/20
	20	1	53/20 = 2,65

□

Si osservi che

$$\bar{y} = \frac{y_1 + y_2 + \cdots + y_n}{n} = \frac{n_1 x_1 + n_2 x_2 + \cdots + n_k x_k}{n}$$

$$= \boxed{\frac{n_1}{n}} x_1 + \boxed{\frac{n_2}{n}} x_2 + \cdots + \boxed{\frac{n_k}{n}} x_k .$$

a questo punto è utile ricordare che la frequenza relativa di un evento in un certo numero di prove fornisce una approssimazione alla probabilità dello evento. In tal caso:

$$\frac{n_1}{n} \simeq \mathbb{P}(Y=x_1); \quad \frac{n_2}{n} \simeq \mathbb{P}(Y=x_2); \ldots; \frac{n_k}{n} \simeq \mathbb{P}(Y=x_k).$$

Si consideri ora un numero aleatorio discreto X. Esso è caratterizzato da

$$S_X = \{x_1, \ldots, x_k) \quad \text{spettro di } X$$

(che è analogo all'insieme delle modalità) e dalla attribuzione delle probabilità

$$i = (1, 2, \ldots, k), \qquad p_i := \mathbb{P}(X = x_i)$$

(che è analoga alla determinazione delle frequenze relative).

Si determina in tal modo la k-pla di coppie

$$(x_1, p_1), (x_2, p_2), \dots, (x_k, p_k)$$

che è analoga alla k-pla delle coppie modalità - frequenze.

Pertanto, si definisce la media del numero aleatorio Y ponendo:

$$\mathbb{E}(X) := \sum_{j=1}^{k} x_j \cdot p_j = \sum_{j=1}^{k} x_j \, \mathbb{P}(X = x_j).$$

Più in generale, si definisce la media di ordine r di X ponendo:

$r \in \mathbb{N}$ $$\mathbb{E}(X^r) := \sum^{k} x_j^r \cdot p_j = \sum^{k} x_j^r \, \mathbb{P}(X = x_j).$$

Tutto quello che è stato detto per il caso di un numero aleatorio discreto con spettro finito è valido anche nel caso di numeri aleatori discreti con spettro numerabile. In tal caso la somma è sostituita dalla serie:

$$r \in \mathbb{N}, \qquad \mathbb{E}(Y^r) := \sum_{j=1}^{\infty} x_j^r \cdot p_j = \sum_{j=1}^{\infty} x_j^r \, \mathbb{P}(X = x_j).$$

$\square$