

Matematica e statistica

statistica quantitativa ^{su n individui} con modello

$x_1, \dots, x_k$ e frequenze assolute $n_1, \dots, n_k$

$$\mu(x) = \sum_{j=1}^k \frac{n_j}{n} \cdot x_j$$

Misura della dispersione:

Varianza . $\sigma^2(x) = \frac{1}{n} \sum_{j=1}^K m_j (x_j - \mu(x))^2$

deviazione standard o
scarto quadratico medio

$$\sigma(x) = \sqrt{\sigma^2(x)}$$

Fatto : $\sigma^2(x) = \mu(x^2) - \mu(x)^2$

Es : statistiche con valori 1, 3, 4, 7, 8 tutti
con frequenza 1;

$$\mu(x) = \frac{1}{5} (1 + 3 + 4 + 7 + 8) = \frac{21}{5}$$

$$\mu(x^2) = \frac{1}{5} (1 + 9 + 16 + 49 + 64) = \frac{139}{5}$$

$$\sigma^2(x) = \frac{139}{5} - \left(\frac{21}{5}\right)^2$$

Spiegazione (con tutte frequenze = 1) :
per semplicità

$$\begin{aligned} \sigma^2(x) &= \frac{1}{n} \sum_{j=1}^n (x_j - \mu(x))^2 = \frac{1}{n} \sum_{j=1}^n (x_j^2 - 2x_j\mu(x) + \mu(x)^2) \\ &= \frac{1}{n} \sum_{j=1}^n x_j^2 - 2\mu(x) \frac{1}{n} \sum_{j=1}^n x_j + \mu(x)^2 \cdot \frac{1}{n} \sum_{j=1}^n 1 \end{aligned}$$

$$\mu(x^2)$$

$$\mu(x)$$

$$\frac{1}{\mu(x)^2}$$

$$-2\mu(x)^2$$

$$-\mu(x)^2$$

Possibile legame tra due statistiche
popolazione = noi

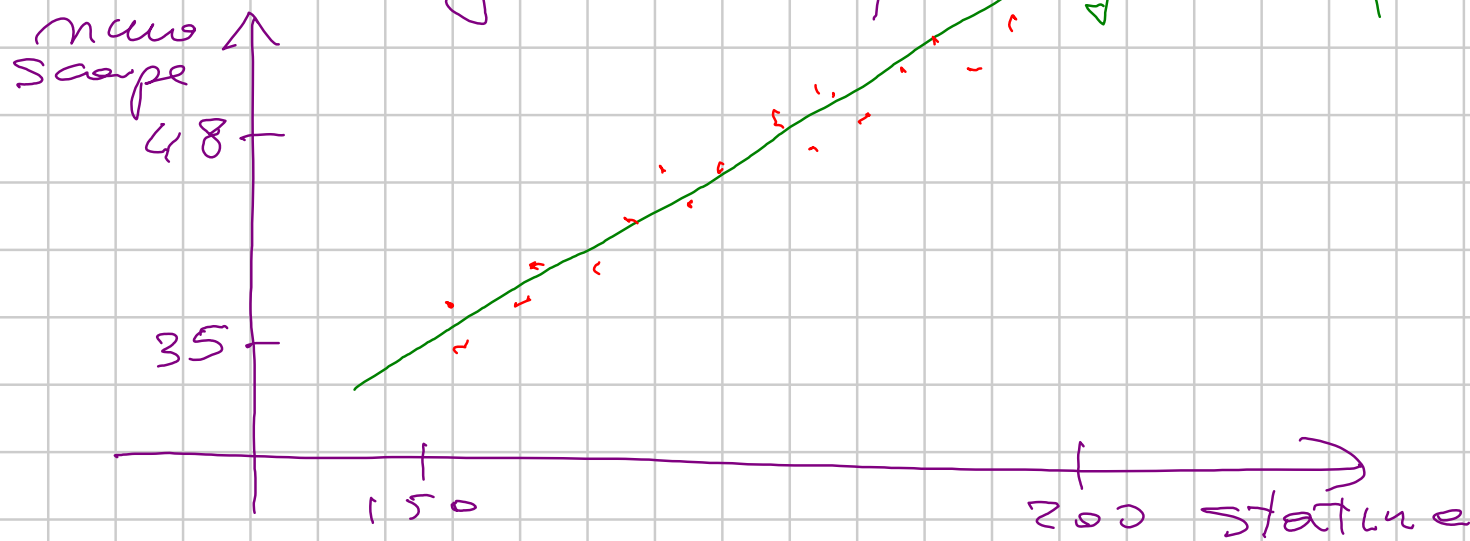
x = stature

y = numero di scarpe

Si può ipotizzare un legame del tipo

$$y = mx + q$$

$$y = mx + q$$



(Suppongo tutti gli $x_1, \dots, x_n$ ripetuti con le loro frequenze)

Covarianza:
$$\sigma(x, y) = \frac{1}{n} \sum_{j=1}^n (x_j - \mu(x)) (y_j - \mu(y))$$

Correlazione:
$$\rho(x, y) = \frac{\sigma(x, y)}{\sigma(x) \cdot \sigma(y)}$$

Calcoliamo nel caso $y = mx + q$ cioè

$$y_j = mx_j + q \quad ; \quad \mu(y) = \frac{1}{n} \sum_{j=1}^n y_j = \frac{1}{n} \sum_{j=1}^n (mx_j + q)$$

$$= m \cdot \underbrace{\frac{1}{m} \sum_{j=1}^m x_j}_{\mu(x)} + q \cdot \underbrace{\frac{1}{m} \sum_{j=1}^m 1}_{1}$$

$$= m \cdot \mu(x) + q$$

$$\sigma(x, y) = \frac{1}{m} \sum_{j=1}^m (x_j - \mu(x))(y_j - \mu(y))$$

$$= \frac{1}{m} \sum_{j=1}^m (x_j - \mu(x))(mx_j + q - (m\mu(x) + q))$$

$$= m \cdot \frac{1}{n} \sum_{j=1}^n (x_j - \mu(x))^2 = m \cdot \sigma^2(x)$$

$$\sigma^2(y) = \frac{1}{n} \sum_{j=1}^n (y_j - \mu(y))^2 = \frac{1}{n} \sum_{j=1}^n (m x_j + q - (m \mu(x) + q))^2$$

$$= m^2 \cdot \frac{1}{n} \sum_{j=1}^n (x_j - \mu(x))^2 = m^2 \cdot \sigma^2(x)$$

$$\Rightarrow \sigma(y) = |m| \cdot \sigma(x)$$

$$r(x, y) = \frac{\sigma(x, y)}{\sigma(x) \cdot \sigma(y)} = \frac{m \cdot \sigma^2(x)}{\sigma(x) \cdot |m| \cdot \sigma(x)} = \frac{m}{|m|}$$

Scoperto: se vale il legame $y = mx + q$
 allora $r(x, y) = \pm 1$

Fatto: $r(x, y)$ è sempre tra -1 e 1 ;
 se è molto vicino a ± 1 allora la

legge $y = mx + q$ è una buona approx
di y . In tal caso:

$$m = \frac{\sigma(x, y)}{\sigma^2(x)}$$

$$q = \mu(y) - m \cdot \mu(x)$$

coeff. ang. e ord. all'origine della retta di regressione

In pratica date due statistiche sulla stessa popolazione

$$x_1, \dots, x_n$$

$$y_1, \dots, y_n$$

- calcolo $f(x, y)$; se $\bar{e}$ "lontano" da ± 1 niente ; se $\bar{e}$ "vicino" a ± 1
- calcolo m, q e li uso per fare previsioni su nuovo valore di y che corrisponde a nuovo valore di x .

Qss: in qualche caso può esserci tra due statistiche x e y un legame del tipo
 $y = m \cdot f(x) + q$ dove f è una funzione.

ES: popolazione = cellule di un tessuto
 x = diametro max
 y = massa

ipotesi semplice $y = m \cdot x^3 + q$

Per scoprire se $y = m \cdot f(x) + q$ è ragionevole
si calcola $\int (f(x), y)$ e si usano $f(x), y$
per calcolare m, q .

Calcolo combinatorio e probabilità

Calcolo combinatorio: Conto di quanti sono i modi in cui può realizzarsi un certo fenomeno (o una scelta).

Es: • possibili uscite lancio dado = 6
• possibili piazzi estratti a tombola = 90

Fatto generale: se ho due fenomeni indipendenti (non correlati) il numero di esiti complessivi dei due

è il prodotto dei numeri di esiti separati.

Disposizione di n oggetti:

un modo di disporre in ordine n oggetti
distinti associati

Quante sono? $n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 3 \cdot 2 \cdot 1 = n!$
(fattoriale)

Convenzione: $0! = 1$

Combinazione di k oggetti tra n

un modo di scegliere k oggetti non ordinati
tra n oggetti distinti e separati.

Es: $\{a, b, c, d\}$ $n=4$ $k=2$

$\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}$

6 modi

La grande cosa:

$$\binom{n}{k}$$

"n binomiale k"

coefficiente binomiale

$$= \frac{n \cdot (n-1) \cdot (n-2) \cdots (n-k+1)}{k!}$$

numero di scelte di k oggetti ordinati

numero di ordinamenti di k oggetti

$$= \frac{n!}{k! \cdot (n-k)!}$$

numero di ordinamenti di n

$(\text{numero di ordinamenti di } k) \cdot (\text{numero di ord. di } (n-k))$

$\binom{n}{k}$ si calcola con triangolo di Tartaglia

3
4
5
6
7

				1					
			1		1				
		1		2		1			
		1	3		3		1		
	1		4	6		4		1	
	1	5		10		10	5		1
	1	6	15		20		15	6	1
1	7	21	35	35	21	7	1		

$\binom{4}{2}$

$\binom{7}{5}$

Fatto: $(a+b)^m = \sum_{k=0}^m \binom{m}{k} a^k \cdot b^{m-k}$

Esempio: $m=5$

$$(a+b)(a+b)(a+b)(a+b)(a+b) = 32 \text{ addendi}$$

raccogliendo: $a^3 b^2$ c'è varie volte

devo scegliere 3 su 5 dei binomi da cui prendere a (e prendo b doppi' altri 2)

$\Rightarrow a^3 b^2$ si trova $\binom{5}{3}$ volte

Probabilità elementari

Esperimento o fenomeno o scelta

con un numero finito possibile di esiti

che hanno tutti la stessa possibilità di

realizzarsi.

Es: estrazione di una carta da un mazzo da 40

evento = uno dei possibili esiti
elementare

Ω = spazio degli eventi = tutti gli eventi elem.

evento = un sottoinsieme di Ω .

Es: un evento elem è re di coppe

$\Omega = \{ \text{le 40 carte} \}$

un evento: una figura

Chiamo probabilità di un evento E

$$P(E) = \frac{\text{numero di el. di } E}{\text{numero di el. di } \Omega}$$

Es: prob. di estrarre una figura da un mazzo di 40

$$P(\text{figura}) = \frac{12}{40} = \frac{3}{10}$$

Oss: se ho due esperimenti con spazi eventi Ω, Λ disgiunti e li espro entrambi; E evento del primo, F evento del secondo

$$p(E \times F) = p(E) \cdot p(F)$$

$$E \times F = \{ (x, y) : x \in E, y \in F \} \subset \Omega \times \Lambda$$

$$\text{Guthrie: } p(E \times F) = \frac{\#(E \times F)}{\#(\Omega \times \Lambda)} = \frac{\#(E) \cdot \#(F)}{\#(\Omega) \cdot \#(\Lambda)}$$

$$= p(E) \cdot p(F) .$$

(# = numero d' elementi)

Eg: lancio un dado e si estraggo numero tumbola:

$$p((\text{dado} \leq 4) \text{ e } (\text{tumbola} > 71))$$

$$= \frac{4}{6} \cdot \frac{18}{90} = \dots$$

Es: fenomeno = estrazione di due carte
da un mazzo di 52.

calcoliamo probabilità di estrarre due figure.

$$1) \frac{\begin{array}{l} \# \text{ possibili} \\ \text{estrazioni} \\ \text{di 2 figure} \end{array}}{\begin{array}{l} \# \text{ possibili} \\ \text{estrazioni di} \\ \text{2 carte} \end{array}} = \frac{\binom{12}{2}}{\binom{52}{2}} = \frac{12 \cdot 11}{2} \cdot \frac{2}{52 \cdot 51} = \dots$$

2) estrazione di 2 figure da 52

= estrazione di 1 figura da 52
e poi di un'altra da 51

$$\text{prob} = \frac{12}{52} \cdot \frac{11}{51}$$

Idea intuitiva di prob. è che se ripetiamo
un esperimento $N \gg 0$ volte un
evento E si realizzerà "circa" $p(E) \cdot N$ volte.

Scommessa equa:

Scommetto contro il banco che avvenga un evento con probabilità p ; punto source A ; se perdo lo perdo, se vinco il banco mi dà V .

Scommessa è equa se giocando $N \gg 0$ volte finisco pari: vinco $p \cdot N$ volte e in tal caso guadagno $V - A$; perdo $(1-p) \cdot N$ volte

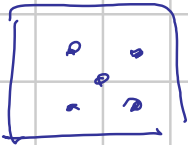
e in tal caso pago δ (bilancio $-\delta$):
bilancio totale:

$$p \cdot N \cdot (v - \delta) + (1 - p) \cdot N \cdot (-\delta)$$

deve fare 0:

$$\cancel{N} \cdot (p \cdot v - \cancel{p \cdot \delta} - \delta + \cancel{p \cdot \delta}) = 0$$

$$V = \frac{A}{P}$$

Es: scommetto $1 \in$ che tirando un dado esce  i.e. dopo n il banco paga $6 \in$ quando esce davvero.

Esercizio: riflettere sull'equità di alcune giocate alla roulette (rosso, ...) e lo

steno pu il cotto -