INTERPOLAZIONE STATISTICA
metodo dei minimi quadrati

Mentre l'interpolazione matematica deve passare per tutti i punti, l'interpolazione statistica passa attraverso i punti osservati. Ciò allo scopo di fornire un "andamento" del fenomeno.

Si ha l'interesse acciocché la funzione interpolante possa passare il più vicino possibile ai valori interpolati.

Ci sono vari metodi per attuare ciò, ma il più usato è quello dei **minimi quadrati**, utilizzato per *andamenti lineari o parabolici*.

Il principio fu introdotto da C. F. Gauss, ed è:

$$\mathbf{F(d) = \sum_i (y_i^* - y_i)^2}$$

da minimizzare

In cui:
d è lo scarto fra punto interpolato e punto osservato
y* è il punto interpolato
y è il punto osservato.

In pratica, la funzione F(d) vale la sommatoria degli scarti al quadrato e va minimizzata.

Caso delle retta

Per determinare la *retta dei minimi quadrati*, vanno impostate e risolte **le equazioni normali:**

$$\begin{cases} na + b\sum x = \sum y \\ a\sum x + b\sum x^2 = \sum xy \end{cases}$$

Per esempio:

Determinare la retta interpolatrice dei seguenti dati:

x	1	3	4	6	8	9
y	1	2	4	5	5	7

Giacché
$n = 6$
$\sum x = 31$
$\sum x^2 = 207$
$\sum y = 24$
$\sum xy = 156$

Il sistema risulterà

$$\begin{cases} 6a + b31 = 24 \\ a31 + b207 = 156 \end{cases}$$

Da questo sistema si otterrà:
a = 0,4697
b = 0,6832

La retta dei minimi quadrati sarà:
y = 0,47 + 0,68x

Caso della parabola

Le **equazioni normali** della parabola dei minimi quadrati, sono:

$$\begin{cases} na + b\sum x + c\sum x^2 = \sum y \\ a\sum x + b\sum x^2 + c\sum x^3 = \sum xy \\ a\sum x^2 + b\sum x^3 + c\sum x^4 = \sum x^2 y \end{cases}$$

Verificare che i seguenti dati ammettano un'interpolazione parabolica:

x	0,1	0,3	0,8	1,4	2,7
y	2,5	5	9	12	16

Impostiamo intanto una tabella di calcolo:

	x	y	x^2	x^3	x^4	xy	x^2y
	0,1	2,5	0,01	0,001	0,0001	0,25	0,025
	0,3	5	0,09	0,027	0,0081	0,9	0,45
	0,8	9	0,64	0,512	0,4096	7,2	5,76
	1,4	12	1,96	2,744	3,8416	16,8	20,28
	2,7	16	7,29	19,683	53,1441	43,2	116,64
$\sum$	5,3	44,5	9,99	22,967	57,4035	68,95	146,395

Quindi, sostituiamo i valori nel sistema:

$$\begin{cases} 5a + b5{,}3 + c9{,}99 = 44{,}5 \\ a5{,}3 + b9{,}99 + c22{,}967 = 68{,}95 \\ a9{,}99 + b22{,}967 + c57{,}4035 = 146{,}395 \end{cases}$$

Da cui si ottiene:

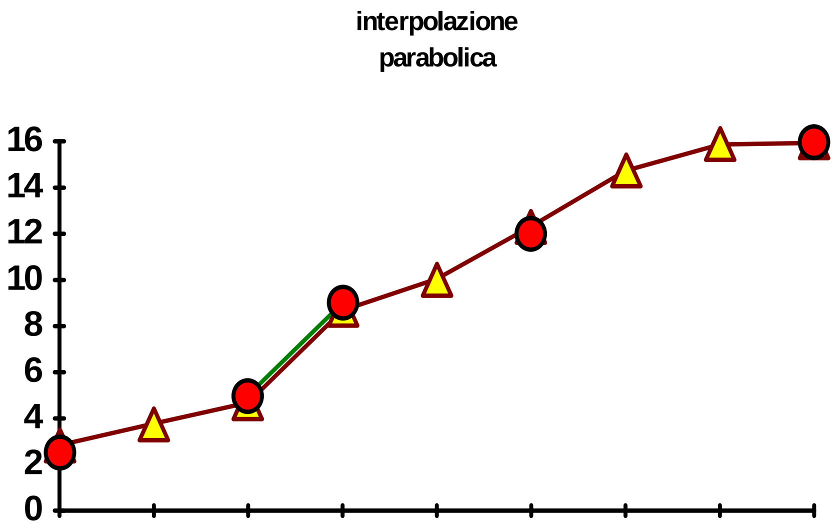
a = 1,8744
b = 9,9072
c = -1,7397

Da cui l'equazione della parabola:

$y = 1{,}8744 + 9{,}9072x - 1{,}7397x^2$

attraverso la quale sarà possibile interpolare i punti desiderati:

x	y interp.	y osserv.
0,1	2,84	2,5
0,2	3,78	
0,3	4,69	5
0,8	8,68	9
1	10,04	
1,4	12,33	12
2	14,73	
2,6	15,87	
2,7	15,94	16



Indici di scostamento

A questo punto, ci si chiede: qual è la bontà dell'approssimazione? E' accettabile lo scostamento?
Il grado di scostamento fra i valori osservati y e quelli interpolati y' è accettabile se gli indici di scostamento <u>non superano</u> il valore **0,1**
Gli indici di scostamento sono di sue tipi:

INDICE LINEARE RELATIVO $I_1 = \frac{\sum |y - y'|}{\sum y'}$

INDICE QUADRATICO RELATIVO $I_2 = \frac{\sqrt{\frac{\sum (y - y')^2}{n}}}{\frac{\sum y'}{n}}$ (preferibile)

Poi ci sono <u>indici di scostamento "medio"</u>, che indicano se la funzione interpolatrice può rappresentare il fenomeno. Se l'interpolazione è perfetta (i punti coincidono), scostamento medio e scostamento relativo saranno uguali a 0.

SCOSTAMENTO MEDIO (media quadratica degli scostamenti)

$$sm = \sqrt{\frac{\sum (y' - y)^2}{n}}$$

SCOSTAMENTO RELATIVO $sr = \sqrt{\frac{\sum (y' - y)^2}{\sum y^2}}$

Nell'ultimo esempio, abbiamo:
$I_1 = 0{,}030$
$I_2 = 0{,}032$
Sm = 0,29
Sr = 0,02