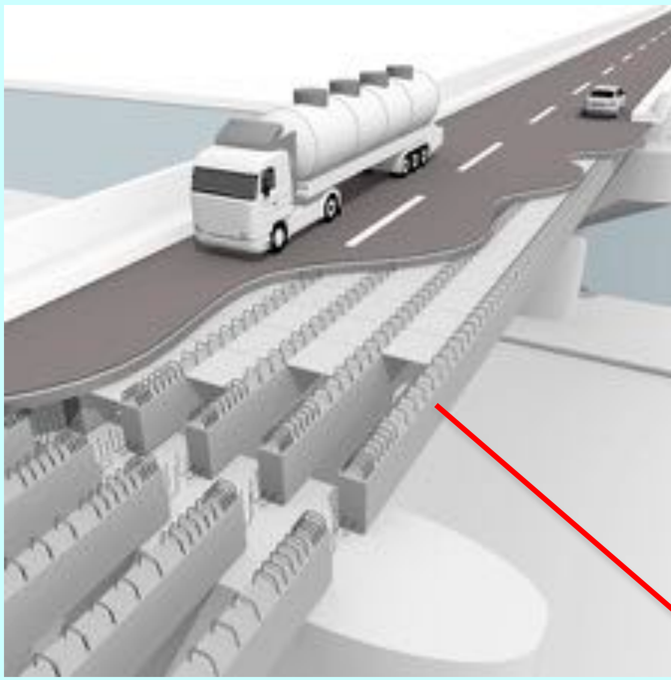
Retta di regressione

La statistica nella ricerca di leggi sperimentali

Cominciamo con un problema importante nella progettazione e realizzazione di ponti, strade, edifici …

Il ruolo delle travi

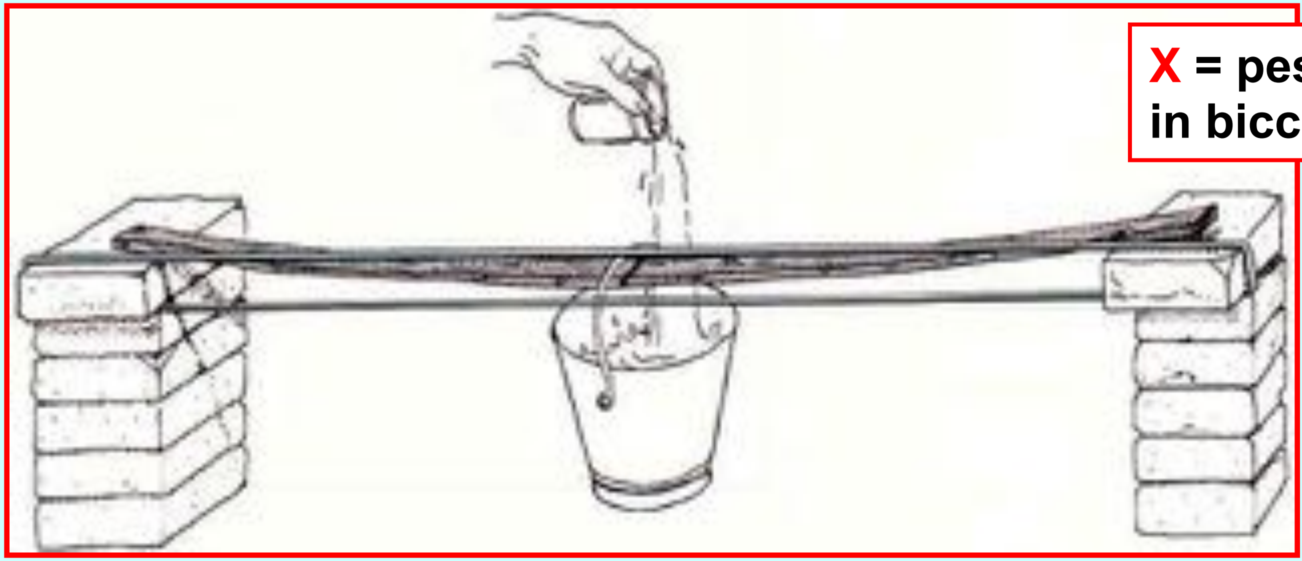
Le travi: oggetto di studi sperimentali nel campo dell'ingegneria edile e dell'architettura



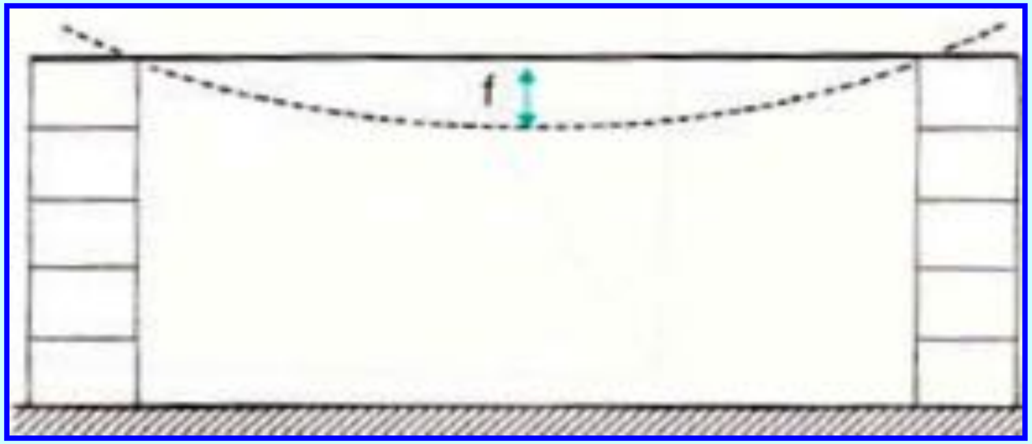
TRAVI

Ricerca di una legge sperimentale

Un esperimento per studiare come si deforma una trave per sostenere il peso di un ponte, di un soffitto, …



X = peso misurato in bicchieri d'acqua

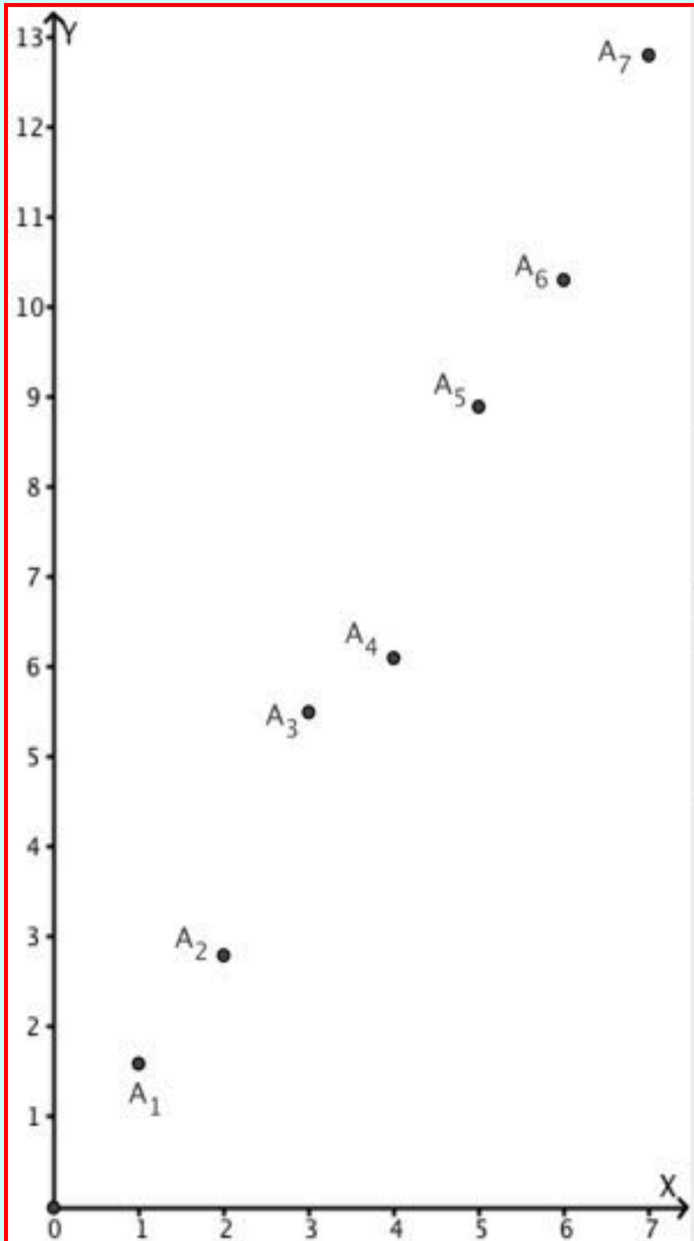


Y = deformazione della trave, misurata con la lunghezza di *f*

Ricerca di una legge sperimentale

Foglio di calcolo con i dati elencati in una tabella e rappresentati su un piano Oxy.

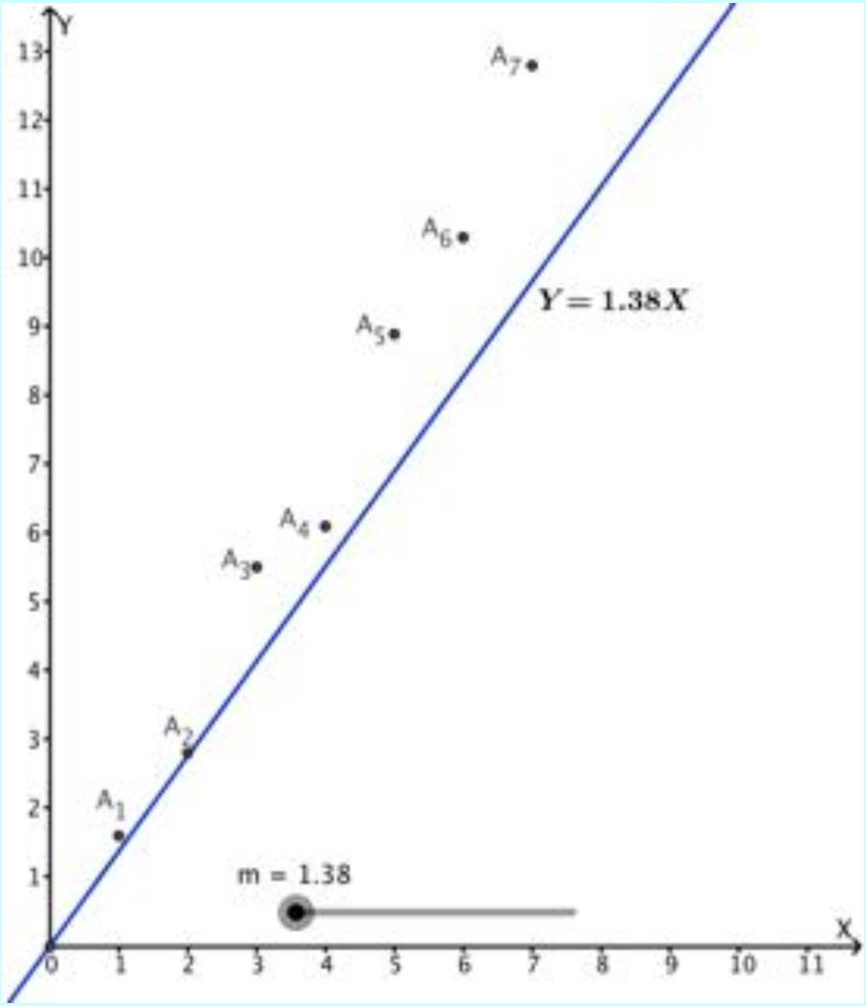
I punti sembrano 'quasi allineati su una retta'.



	A	B
1	X	Y
2	0	0
3	1	1.6
4	2	2.8
5	3	5.5
6	4	6.1
7	5	8.9
8	6	10.3
9	7	12.8
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		

Cerco la retta che raccorda i punti

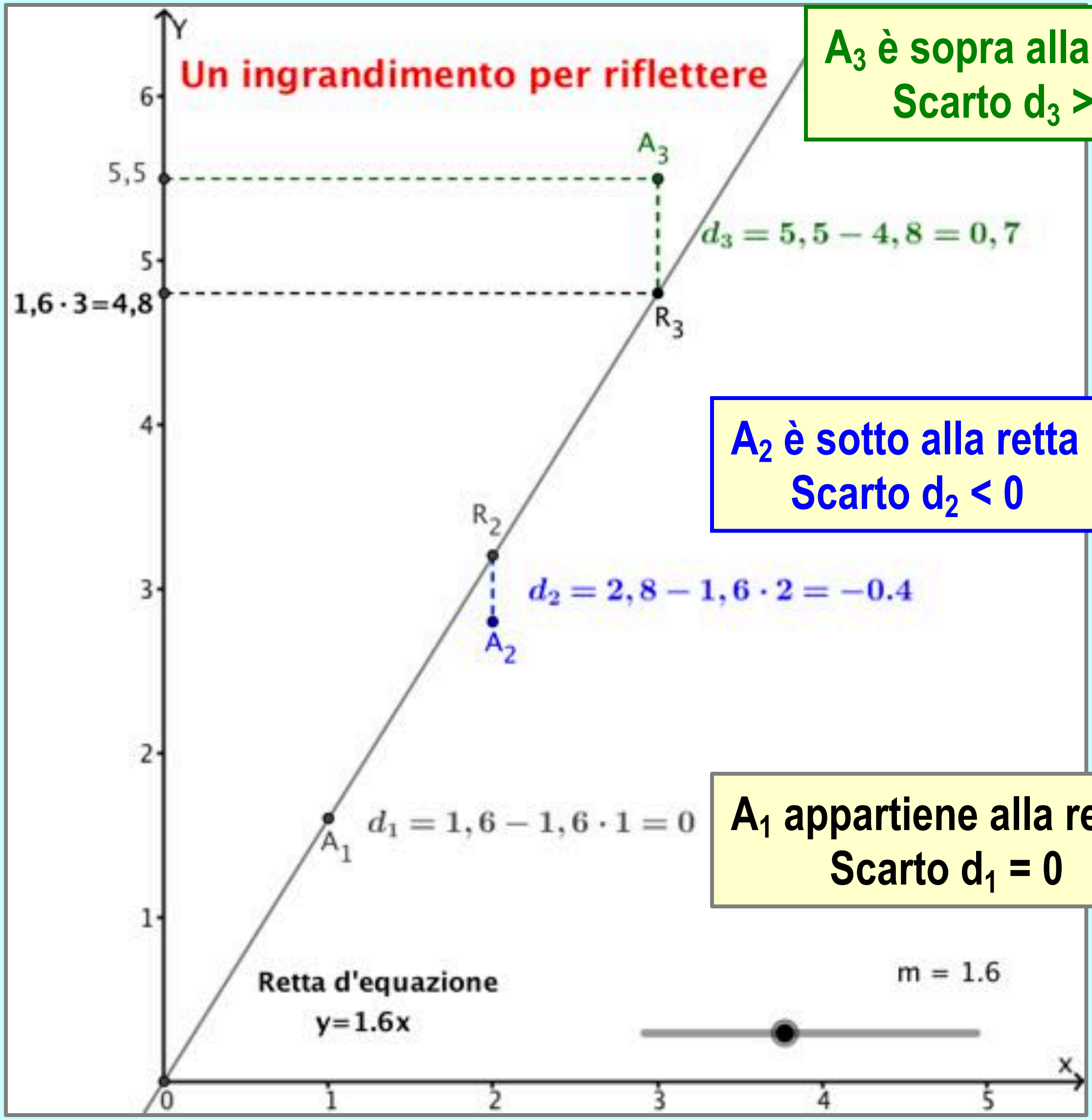
La retta passa per O(0, 0): se il peso X dell'acqua è 0, la trave non si deforma ed è 0 anche Y. L'equazione della retta sarà dunque del tipo $Y = mX$. Come trovo la pendenza m? Tentativi con foglio di calcolo.



Osserva l'animazione

Quanto ‘è vicina’ una retta ai punti sperimentali?

Calcolo gli scarti: alcuni sono positivi, altri negativi



A_3 è sopra alla retta
Scarto $d_3 > 0$

A_2 è sotto alla retta
Scarto $d_2 < 0$

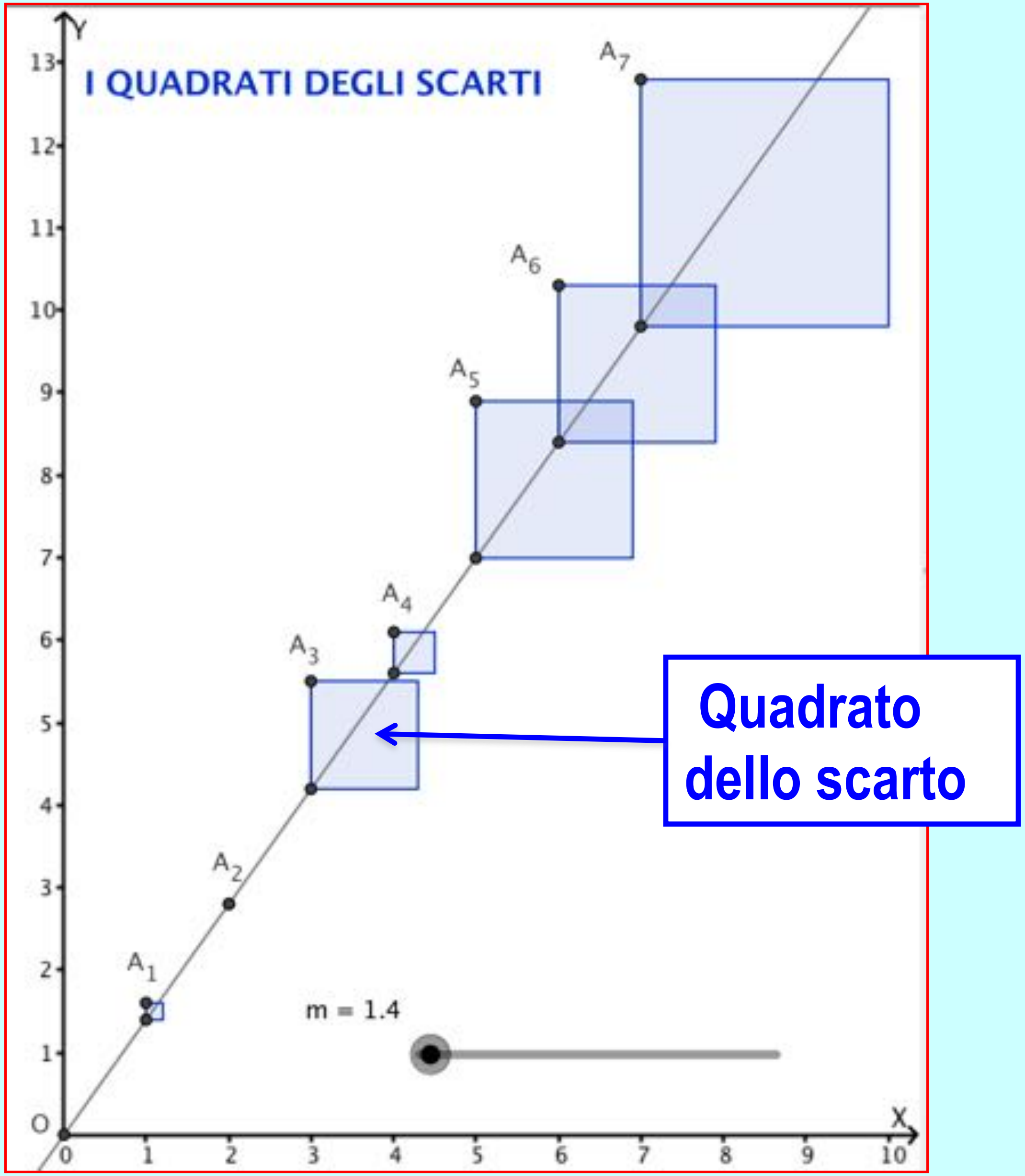
A_1 appartiene alla retta
Scarto $d_1 = 0$

La somma degli scarti non è utile per rispondere.

Conviene valutare la somma dei quadrati degli scarti.

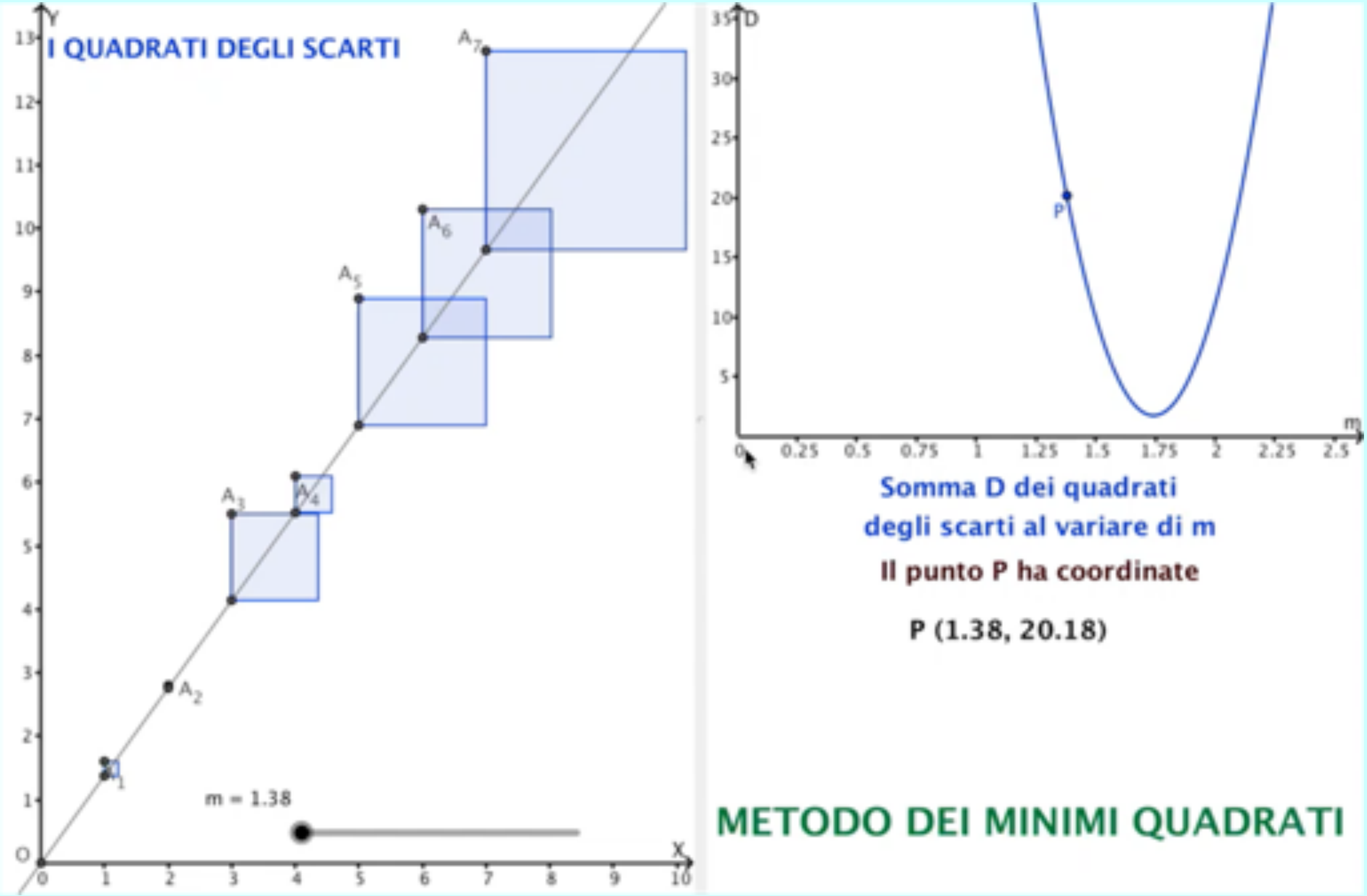
Daniela Valenti, Treccani Scuola 2015

I quadrati degli scarti



Daniela Valenti, Treccani Scuola 2015

Trovo la retta 'più vicina' ai punti sperimentali



VIDEO

La retta ‘dei minimi quadrati’ per O

La retta s_O ‘dei minimi quadrati’ è quella che meglio raccorda gli otto punti sperimentali.
X e Y sono legate dalla legge

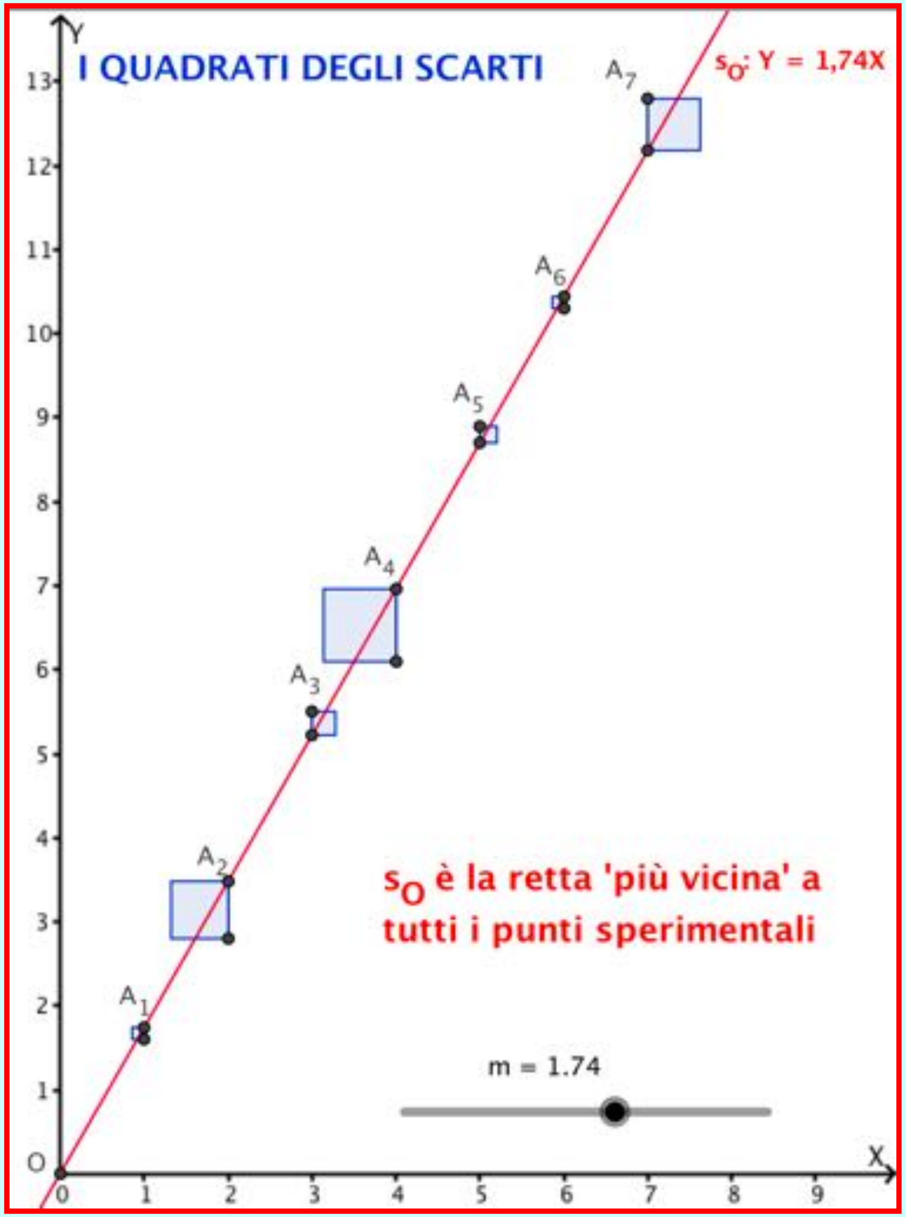
$$Y = 1{,}74X$$

APPLICAZIONI

Prevedere le deformazioni della trave caricata con altri pesi, vicini a quelli sperimentali, senza ripetere l’esperimento; ad esempio:

con X = 8, prevedo

$Y = 1{,}74 \times 8 = 13{,}92$



Attività

Completa la scheda di lavoro per studiare e applicare la ‘retta dei minimi quadrati’.

Che cosa hai ottenuto

La retta 'dei minimi quadrati' per O

Quesito 1

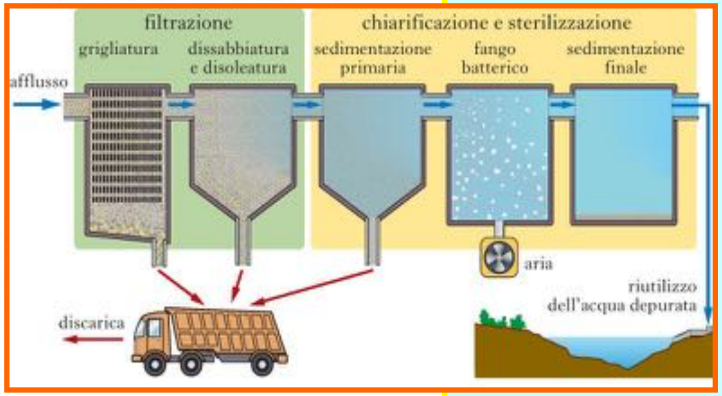
Esempio numerico	In generale
Sono dati i punti sperimentali O(0; 0) , $A_1(1; 1{,}6)$, $A_2(2; 2{,}8)$, …, $A_7(7; 12{,}8)$	Sono dati i punti sperimentali O(0; 0) , $A_1(x_1; y_1)$, $A_2(x_2; y_2)$, …, $A_N(x_N; y_N)$
Esamino la somma D dei quadrati degli scarti fra la retta $y = mx$ e i punti sperimentali, al variare di m $D = (1{,}6 - m\cdot 1)^2 + (2{,}8 - m\cdot 2)^2 + \ldots + (12{,}8 - m\cdot 7)^2$ Sviluppo i quadrati e ottengo una legge del tipo $D = a \cdot m^2 + b \cdot m + c$ con $a = 1^2 + 2^2 + \ldots + 7^2$ $b = -2(1 \cdot 1{,}6 + 2 \cdot 2{,}8 + \ldots + 7 \cdot 12{,}8)$ $c = 1{,}6^2 + 2{,}8^2 + \ldots + 12{,}8^2$ Il grafico della legge è una parabola con la concavità rivolta verso l'alto. Perciò D è minima per $m = -\frac{b}{2a} = -\frac{-2(1\cdot 1{,}6 + 2\cdot 2{,}8 + \ldots + 7\cdot 12{,}8)}{2(1{,}6^2 + 2{,}8^2 + \ldots + 12{,}8^2)}$ Perciò la pendenza m_O che rende minima la somma dei quadrati degli scarti è data da: $m_O = \frac{1\cdot 1{,}6 + 2\cdot 2{,}8 + \ldots + 7\cdot 12{,}8}{1{,}6^2 + 2{,}8^2 + \ldots + 12{,}8^2}$	Esamino la somma D dei quadrati degli scarti fra la retta $y = mx$ e i punti sperimentali, al variare di m $D = (y_1 - m \cdot x_1)^2 + (y_2 - m \cdot x_2)^2 + \ldots + (y_N - m \cdot x_N)^2$ Sviluppo i quadrati e ottengo una legge del tipo $D = a \cdot m^2 + b \cdot m + c$ con $a = x_1^2 + x_2^2 + \ldots + x_N^2$ $b = -2(x_1 \cdot y_1 + x_2 \cdot y_2 + \ldots + x_N \cdot y_N)$ $c = y_1^2 + y_2^2 + \ldots + y_N^2$ Il grafico della legge è una parabola con la concavità rivolta verso l'alto. Perciò D è minima per $m = -\frac{b}{2a} = -\frac{-2(x_1 \cdot y_1 + x_2 \cdot y_2 + \ldots + x_N \cdot y_N)}{2(x_1^2 + x_2^2 + \ldots + x_N^2)}$ Perciò la pendenza m_O che rende minima la somma dei quadrati degli scarti è data da: $m_O = \frac{x_1 \cdot y_1 + x_2 \cdot y_2 + \ldots + x_N \cdot y_N}{x_1^2 + x_2^2 + \ldots + x_N^2} = \frac{\sum_{k=1}^{k=N} x_k \cdot y_k}{\sum_{k=1}^{k=N} x_k^2}$

La retta 'dei minimi quadrati' per O

Quesito 2

2. *Nei grandi acquedotti è importante filtrare l'acqua e trattare i sedimenti ottenuti prima di distribuirla; nella tabella qui sotto* ***X*** *indica la quantità d'acqua filtrata (misurata in metri cubi) e* ***Y*** *la quantità di sedimenti estratti (misurata in kg).*

X	31	33	37	40	45
Y	14,0	17,1	20,4	21,3	27,4



A partire dai dati in tabella risolvi i seguenti quesiti:

a. Per prevedere i futuri sedimenti trovi la retta dei minimi quadrati. Spiega perché la retta deve passare per O(0, 0).

Perché se filtro 0 m^3 d'acqua, ottengo 0 kg di sedimenti.

b. Completa il calcolo della pendenza m_O della retta s_O con l'aiuto di una calcolatrice tascabile

$$m_O = \frac{31 \cdot 14 + 33 \cdot 17{,}1 + 37 \cdot 20{,}4 + 40 \cdot 21{,}3 + 45 \cdot 27{,}4}{31^2 + 33^2 + 37^2 + 40^2 + 45^2} \cong 0{,}54$$

c. Come prevedere la quantità di sedimenti che si otterrà dopo aver filtrato $60 m^3$ d'acqua?

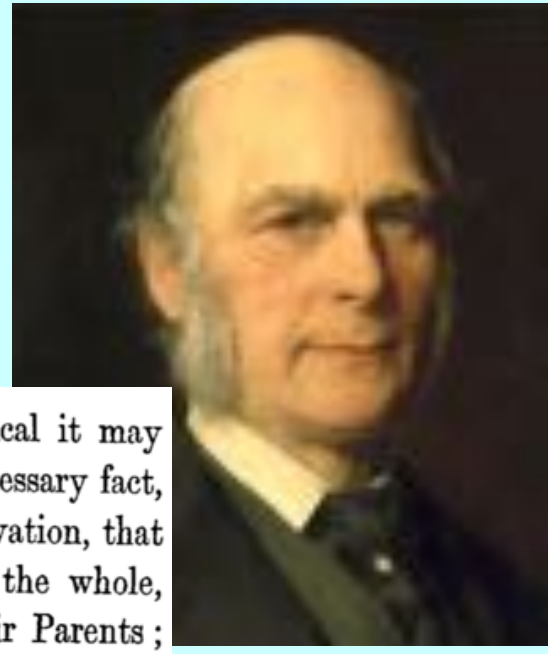
Con la legge Y = 0,54X; per X = 60, ottengo Y = 0,54 · 60 = 32,4

d. Come sapere di quanto aumentano i sedimenti, quando l'acqua aumenta di $1 m^3$?

Con la pendenza $m_O = 0{,}54 = \frac{\Delta Y}{\Delta X}$. Se ΔX = 1, allora ΔY = 0,54.

Perché il nome ‘retta di *regressione*’?

F. Galton, Regno Unito
1822 - 1911



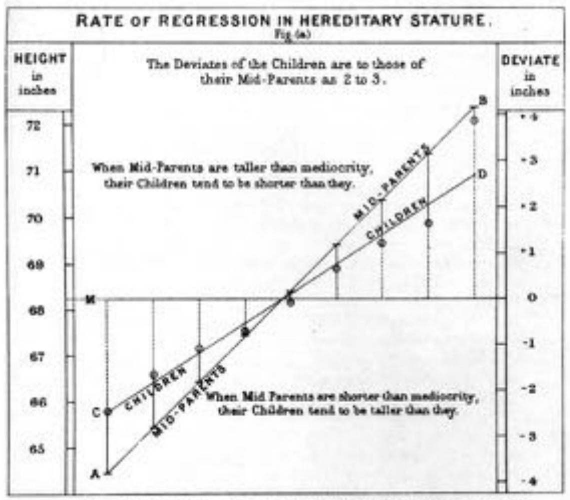


Figure 8.8. Galton's graphical illustration of regression; the circles give the average heights for groups of children whose midparental heights can be read from the line AB. The difference between the line CD (drawn by eye to approximate the circles) and AB represents regression toward mediocrity. (From Galton, 1886a.)

Regression.—a. Filial: However paradoxical it may appear at first sight, it is theoretically a necessary fact, and one that is clearly confirmed by observation, that the Stature of the adult offspring must on the whole, be more *mediocre* than the stature of their Parents; that is to say, more near to the M of the general Population.

Galton mette in relazione l’altezza dei padri e dei figli e trova che:
- padri più alti della media nazionale hanno figli più bassi del padre;
- padri più bassi della media nazionale hanno figli più alti del padre.
Perciò Galton parla di *‘regressione verso la media’.*
La ‘retta di regressione’ nasce per studiare questa singolare relazione.