

CORRELAZIONE

tre voti di matematica e voti di informatica.

MAT.	INF.
X	Y
7	9
8	8
5	6
9	8
4	5

rette di regressione di y su x

$$y = f(x)$$

$$y = mx + q$$

$$\begin{cases} \sum y_i = m \sum x_i + nq \\ \sum x_i y_i = m \sum x_i^2 + q \sum x_i \end{cases}$$

$$\sum = \sum_{i=1}^n$$

x _i	y _i	x _i y _i	x _i ²	y _i ²	y _i ²
7	9	63	49	81	81
8	8	64	64	64	64
5	6	30	25	36	36
9	8	72	81	64	64
4	5	20	16	25	25
\sum 33	36	249	235	36	270

$$\begin{cases} 36 = 33m + 5q \\ 249 = 235m + 33q \end{cases}$$

$$\begin{cases} q = 7,2 - 6,6m \\ 249 = 235m + 237,6 - 217,8m \end{cases}$$

risolvo

$$11,4 = 17,2m \Rightarrow m = \frac{57}{86} \approx 0,66279$$

$$\begin{cases} m = 0,66279 \dots \\ q = 2,82558 \dots \end{cases}$$

$$y = 0,66279x + 2,82558$$

m è il coefficiente di regressione di y su x

Per trovare le rette di regressione di x su y basta scambiare x e y

$$\begin{cases} \sum x_i = m \sum y_i + nq \\ \sum x_i y_i = m \sum y_i^2 + q \sum y_i \end{cases} \quad \begin{cases} 33 = 36m + 5q \\ 249 = 270m + 36q \end{cases}$$

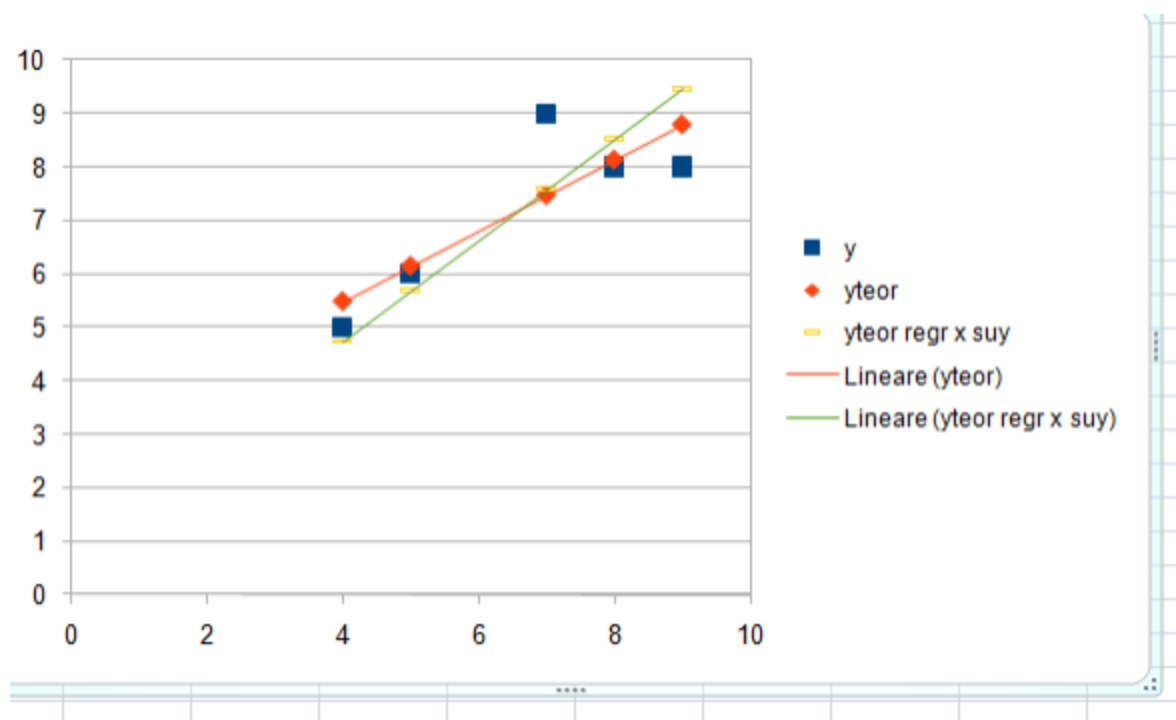
$$\begin{cases} q = 6,6 - 7,2m \\ 249 = 270m + 237,6 - 259,2m \end{cases} \quad \begin{cases} \text{risolvo} \\ m = 1,05 \end{cases} \quad \begin{cases} m = 1,05555 \dots \\ q = -1 \end{cases}$$

$$x = 1,05y - 1$$

è il coefficiente di regressione di x su y

$$m_1 = 1,05 = \frac{95}{90} = \frac{19}{18}$$

cioè $y = 0,947368x + 0,947368$ ← serve per disegnarla



$$-1 \leq r \leq 1$$

Coefficiente di correlazione lineare

$$r = \pm \sqrt{m \cdot m_1}$$

↑ + se m e m₁ sono positivi
- se m e m₁ sono negativi

$$r = + \sqrt{\frac{57}{86} \cdot \frac{19}{18}} = \sqrt{0,6996} = 0,8364$$

$$\sum X_i = 33$$

$$\sum Y_i = 36$$

$$\bar{X} = \frac{\sum X_i}{n} = \frac{33}{5} = 6,6$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{36}{5} = 7,2$$

$$X'_i = \text{scarti di } X_i = X_i - \bar{X}$$

$$Y'_i = \text{scarti di } Y_i = Y_i - \bar{Y}$$

BAR (6,6; 7,2)

le due rette di regressione si incontrano nel baricentro \uparrow

X_i	Y_i	scarti di x X'_i	scarti di y Y'_i	$X'_i Y'_i$	$X_i'^2$	$Y_i'^2$
7	9	0,4	1,8	0,72	0,16	3,24
8	8	1,4	0,8	1,12	1,96	0,64
5	6	-1,6	-1,2	1,92	2,56	1,44
9	8	2,4	0,8	1,92	5,76	0,64
4	5	-2,6	-2,2	5,72	6,76	4,84
		0	0	11,4	17,2	10,8

$$m = \frac{\sum X'_i Y'_i}{\sum X_i'^2}$$

$$m_1 = \frac{\sum X'_i Y'_i}{\sum Y_i'^2}$$

$$m = \frac{11,4}{17,2} = 0,662790697$$

$$m_1 = \frac{11,4}{10,8} = 1,05$$

Il baricentro (6,6; 7,2) è un punto che appartiene ad entrambe le rette di regressione

Quindi per trovare la retta di regressione di y su x

sostituisco le coordinate del BAR nell'equazione $y = 0,66279x + q$

$$7,2 = 0,66279 \cdot 6,6 + q$$

$$q = 7,2 - 4,374414 = 2,82558$$

Quindi la retta di regressione di y su x è: $y = 0,66279x + 2,82558$

per trovare la retta di regressione di x su y

sostituisco le coordinate del BAR (6,6; 7,2) nell'equazione $x = 1,05y + q$

$$6,6 = 1,05 \cdot 7,2 + q \Rightarrow q = 6,6 - 7,6 = -1$$

$$1,05 = \frac{105-10}{90} = \frac{95}{90} = \frac{19}{18} \quad \frac{19}{18} \cdot 7,2 = 0,4 \cdot 19 = 7,6$$

Quindi la retta di regressione di x su y è $x = 1,05y - 1$