

**La Retta di Regressione: Un Esempio.** Se  $x$  è la temperatura in Celsius e  $y$  quella in Fahrenheit, si ha l'equazione

$$y = (1,8)x + 32.$$

Adesso consideriamo i seguenti dati  $(x_i, y_i)$  (per le temperature in Celsius e in Fahrenheit):

$$(0, 30), \quad (10, 40), \quad (20, 70), \quad (30, 85), \quad (40, 105).$$

Quindi ci sono alcuni errori nei dati Fahrenheit.

| $i$   | $x_i$ | $y_i$ | $x_i^2$ | $x_i y_i$ |
|-------|-------|-------|---------|-----------|
| 1     | 0     | 30    | 0       | 0         |
| 2     | 10    | 40    | 100     | 400       |
| 3     | 20    | 70    | 400     | 1400      |
| 4     | 30    | 85    | 900     | 2550      |
| 5     | 40    | 105   | 1600    | 4200      |
| Somma | 100   | 330   | 3000    | 8550      |
| Media | 20    | 66    | 600     | 1710      |

Quindi

$$\bar{x} = 20, \quad \bar{y} = 66, \quad \overline{x^2} = 600, \quad \overline{xy} = 1710.$$

L'equazione per i coefficienti  $a$  e  $b$  che appaiono nell'equazione  $y = ax + b$  della retta di regressione, sono le soluzioni del seguente sistema di equazioni:

$$\begin{pmatrix} \overline{x^2} & \bar{x} \\ \bar{x} & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \overline{xy} \\ \bar{y} \end{pmatrix}.$$

In questo caso otteniamo il sistema di equazioni

$$\begin{pmatrix} 600 & 20 \\ 20 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1710 \\ 66 \end{pmatrix}.$$

Il determinante della matrice  $2 \times 2$  è  $\sigma^2 = \overline{x^2} - (\bar{x})^2 = 200$ . Quindi risulta la soluzione

$$\begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{200} \begin{pmatrix} 1 & -20 \\ -20 & 600 \end{pmatrix} \begin{pmatrix} 1710 \\ 66 \end{pmatrix} = \begin{pmatrix} 1,85 \\ 27 \end{pmatrix}.$$

Dunque la retta di regressione è

$$y = (1,85)x + 27.$$