

STATISTICA: Studio di una serie di dati

Sia data la seguente serie di dati di variabili X e Y relative, rispettivamente, ad età ed altezza di un campione di dieci individui.

X [anni]	16	17	16	18	19	15	20	15	17	18
Y [cm]	165	178	171	180	177	166	172	161	175	173

Effettuare un'analisi statistica, calcolando per entrambe le variabili:

- Media aritmetica
- Varianza campionaria
- Deviazione standard campionaria
- Covarianza

Studiare la correlazione tra le due variabili, interpretandone il risultato.

Ricavare la retta di regressione lineare e riportarla in un grafico insieme alla serie di dati.

SOLUZIONE

Numero elementi: $n = 10$

Media aritmetica

$$\bar{X} = \frac{\sum X_i}{n} = \frac{16 + 17 + 16 + 18 + 19 + 15 + 20 + 15 + 17 + 18}{10} = 17.1$$

$$\bar{Y} = \frac{\sum Y_i}{n} = \frac{165 + 178 + 171 + 180 + 177 + 166 + 172 + 161 + 175 + 173}{10} = 171.8$$

Indicatori di dispersione

$$S_X^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1} = \frac{24.93}{9} = 2.77$$

$$S_Y^2 = \frac{\sum (Y_i - \bar{Y})^2}{n - 1} = \frac{341.6}{9} = 37.96$$

$$S_X = \sqrt{S_X^2} = 1.66$$

$$S_Y = \sqrt{S_Y^2} = 6.16$$

$$S_{XY} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{n - 1} = 6.80$$

Coefficiente di correlazione (Bravais-Pearson)

$$\rho = \frac{S_{XY}}{S_X S_Y} = 0.66$$

→ Variabili direttamente correlate con correlazione moderata

Equazione della retta di regressione lineare

$$Y = 2.46 X + 129.77$$

