

Formule - Statistica sociale

Media aritmetica

$$\begin{aligned}\bar{x} &= \sum_{i=1}^n \frac{x_i}{n} = \frac{x_1 + x_2 + \dots + x_i + \dots + x_n}{n} \\ \bar{x} &= \sum_{k=1}^K \frac{x_k * \text{numerosità}_k}{n} \\ \bar{x} &= \sum_{k=1}^K x_k * \text{freq rel}_k \\ \bar{x} &\approx \sum_{k=1}^K \frac{c_k * \text{numerosità}_k}{n} \\ \bar{x} &\approx \sum_{k=1}^K c_k * f_k \\ \bar{x}_w &= \frac{\sum x_i w_i}{\sum w_i} = \frac{x_1 * w_1 + x_2 * w_2 + \dots + x_i * w_i + \dots + x_n * w_n}{w_1 + w_2 + \dots + w_i + \dots + w_n}\end{aligned}$$

Mediana

$$\text{Me} = \begin{cases} \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} & \text{se } n \text{ pari} \\ x_{(\frac{n+1}{2})} & \text{se } n \text{ dispari} \end{cases}$$

Varianza, deviazione standard e coefficiente di variazione

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} [(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2]$$

Formula indiretta per calcolo varianza (nel caso di dati da popolazione)

$$s^2 = \sum_{i=1}^n x_i^2 / n - [\bar{x}]^2 \quad s^2 \text{ è diviso per } n$$

$$s = \sqrt{s^2}$$

$$CV = \frac{s}{\bar{x}} * 100$$

Campo di variazione e differenza (scarto) interquartile

$$R = x_{max} - x_{min}$$

$$IQR = Q_3 - Q_1$$

$$\text{Baffi: } [Q_1 - 1,5 * IQR; Q_3 + 1,5 * IQR]$$

Medie condizionate

$$\bar{y}_{X=x_i} = \frac{1}{n_i} \sum_{j=1}^K y_j n_{ij}$$

Frequenze teoriche di indipendenza

$$n'_{ij} = \frac{n_i \cdot n_j}{n}$$

Chi-quadrato e p-value

$$\chi^2 = \sum_{i=1}^K \sum_{j=1}^H \frac{(n_{ij} - n'_{ij})^2}{n'_{ij}}$$

Valori di riferimento del p-value: 0.1,0.05,0.01

V di Cramer

$$V = \sqrt{\frac{\chi^2}{n * \min (r - 1, c - 1)}}$$

r=righe, c=colonne

Coefficiente di correlazione

$$r_{XY} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$