

## Definizione di frequenza empirica

**Frequenza assoluta:** La frequenza empirica assoluta di una certa caratteristica è data dal numero di volte che essa si presenta all'interno di un dato campione (operativamente: serie di dati)

$$x_i \rightarrow n_i$$

**Frequenza relativa:** La frequenza empirica relativa è definita come il Rapporto fra la frequenza assoluta di una certa Caratteristica e la numerosità del campione  $n$  (numero di dati totale)

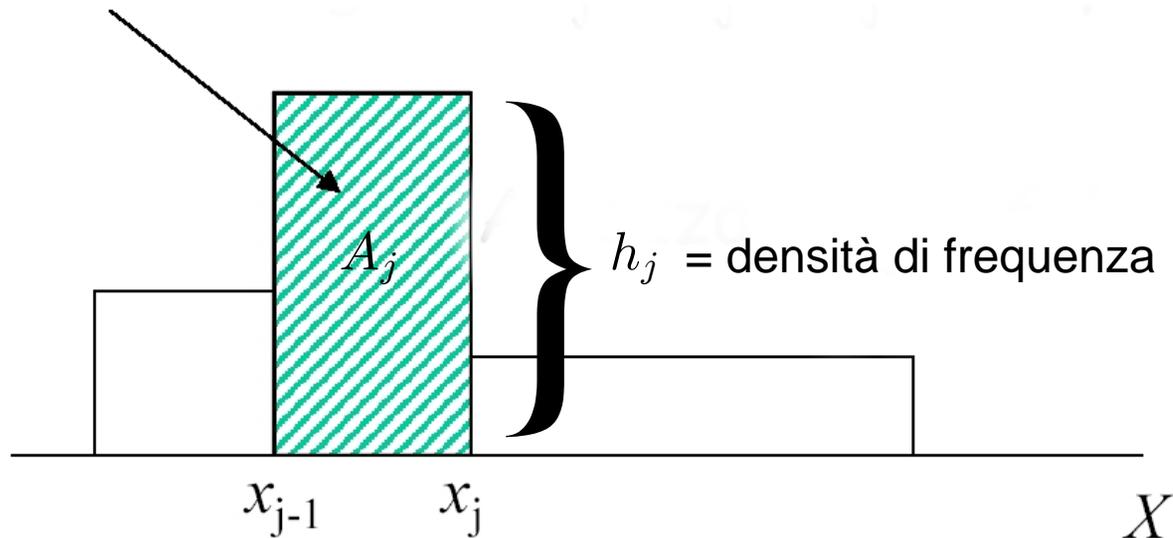
$$f_i \rightarrow \frac{n_i}{n}$$

Per cui si ha ovviamente che:  $\sum_{i=1}^n n_i = n$  e  $\sum_{i=1}^n f_i = 1$

## Metodi puramente descrittivi

### Istogrammi

$$A_j = (x_j - x_{j-1}) \cdot h_j = \Delta x \cdot h_j = \frac{n_j}{n} = f_j$$



Alcune espressioni per la scelta del numero delle Classi  $n_c$

$$n_c = \sqrt{n} \quad (1)$$

$$n_c = 1 + 3.3 \log_{10} n \quad (2)$$

Range dei dati

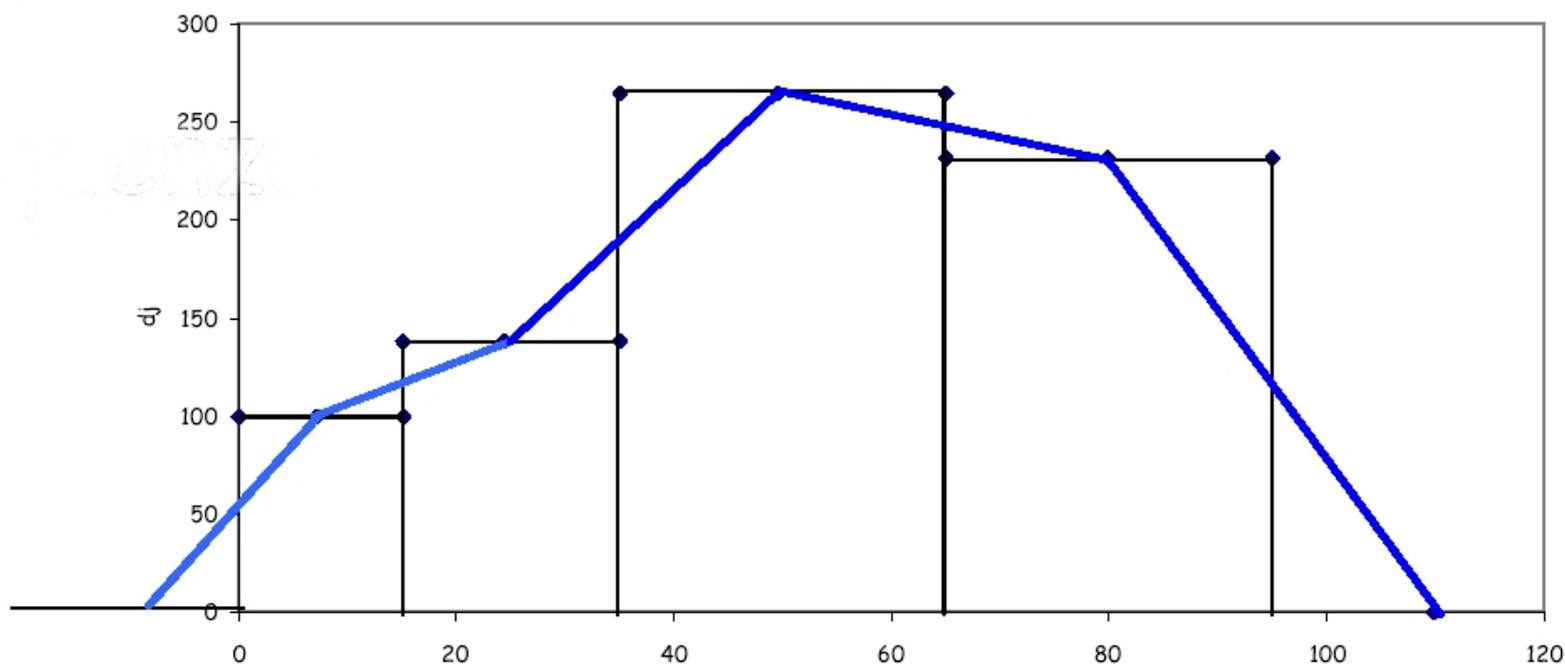
$$n_c = \frac{rn^{1/3}}{2iqr} \quad (3)$$

Distanza inter-quartile

Con:

$$iqr = Q_3 - Q_1$$

## Poligoni di frequenza



## Funzione di ripartizione empirica

Si definisce **funzione di distribuzione cumulata empirica** o **funzione di ripartizione empirica** di una variabile  $X$ , e si indica con  $F_X$  quella applicazione:

$$F_X : \mathbb{R} \rightarrow [0, 1]$$

tale che

$$F_X(x) = P[X \leq x] = P[u : X(u) \leq x] \quad \forall x \in \mathbb{R}$$

## Alcune proprietà della funzione di ripartizione empirica

$$\lim_{x \rightarrow -\infty} F_X(x) = 0 \quad \mathbf{e} \quad \lim_{x \rightarrow +\infty} F_X(x) = 1$$

$F_X$  è monotona non decrescente cioè per  $a < b$   $F_X(a) \leq F_X(b)$

$F_X$  è continua da destra cioè:  $\lim_{h \rightarrow 0^+} F_X(x + h) = F_X(x)$  con  $h > 0$