

## **STATISTICA DESCRITTIVA: SIMBOLOGIA E DEFINIZIONI**

### **(A) VARIABILI STATISTICHE (v.s.) DISCRETE**

$$X = \begin{cases} x_i & i = 1 \dots k \\ p_i = \frac{n_i}{N} \end{cases}$$

dove:

$x_i$  = modalità  $i$  – esima della v.s.  $X$  con le  $x_i$  in ordine crescente ovvero:  $x_i < x_{i+1}$

$n_i = n(x_i)$  = frequenza assoluta della modalità  $x_i$

$N = \sum_{i=1}^k n_i$  = numero delle *unità statistiche* considerate = numero dei dati grezzi

$p_i = p(x_i) = \frac{n_i}{N} = Fr\{X = x_i\}$  = frequenza relativa della modalità  $x_i$

$p_i \% = p_i \times 100$  = frequenza relativa percentuale della modalità  $x_i$

calcolo di  $Fr\{X \leq x\}$ ,  $Fr\{a < X < b\}$ , ecc. (= soma delle  $p_i$  delle  $x_i$  che verificano la condizione)

$mod(X)$  = *moda* della v.s.  $X$  = è la modalità della v.s.  $X$  con la  $p_i$  più alta ovvero con la massima  $p_i$  (caso uni-modale ovvero di una sola moda)

$q_\alpha$  = quantile di ordine  $\alpha$ : è la più piccola modalità della v.s.  $X$  che verifica la condizione:

$$\boxed{Fr\{X \leq q_\alpha\} \geq \alpha}$$

con  $\alpha = 0.25$ :  $q_{0.25} = Q_1$  = primo quartile

con  $\alpha = 0.5$ :  $q_{0.50} = Q_2$  = med( $X$ ) = secondo quartile = mediana

con  $\alpha = 0.75$ :  $q_{0.75} = Q_3$  = terzo quartile

$M(X) = \mu$  = media = valore medio = momento primo della v.s.  $X = \sum_{i=1}^k x_i p_i$

$M(X^2)$  = momento secondo della v.s.  $X = \sum_{i=1}^k x_i^2 p_i$

$V(X) = \sigma^2$  = varianza della v.s.  $X = \sum_{i=1}^k [x_i - M(X)]^2 p_i = M(X^2) - M(X)^2$

$s.q.m.(X) = \sigma$  = scarto quadratico medio della v.s.  $X = \sqrt{V(X)}$

**Forma della distribuzione di una v.s. discreta** (sulla base del diagramma ad aste, solo caso uni-modale):

- (asimmetrica) obliqua a destra  $\Rightarrow mod(X) \leq med(X) < M(X)$

- (asimmetrica) obliqua a sinistra  $\Rightarrow M(X) < med(X) \leq mod(X)$

- simmetrica (ovvero le modalità equi-distanti dalla media sono equi-frequenti)

$\Rightarrow M(X) = med(X) = mod(X)$

**(B) VARIABILI STATISTICHE (v.s.) CONTINUE (PER INTERVALLI)**

$$X = \begin{cases} [x_i, x_{i+1}) & i = 1 \dots k \\ c_i & \end{cases}$$

dove:

$$c_i = \frac{p_i}{\Delta_i} = \frac{p_i}{\delta_i} = \text{densità di frequenza dell'intervallo } [x_i, x_{i+1})$$

$$p_i = \frac{n_i}{N} = Fr\{x_i \leq X < x_{i+1}\} = \text{frequenza relativa dell'intervallo } [x_i, x_{i+1}) = \Delta_i \times c_i =$$

$$= \text{area del rettangolo con base l'intervallo } [x_i, x_{i+1}) \text{ ed altezza } c_i$$

$$\Delta_i = \delta_i = x_{i+1} - x_i = \text{lunghezza dell'intervallo } [x_i, x_{i+1})$$

calcolo di  $Fr\{X \leq x\}$ ,  $Fr\{a < X < b\}$ , ecc. (da calcolare a partire dall'istogramma con la regola dell'area: lunghezza di base  $\times$  altezza = lunghezza di base  $\times$  densità  $c_i$ )

*intervallo modale*: è l'intervallo con la massima  $c_i$  (caso uni-modale)

$$\text{mod}(X) = \text{moda} = \text{valore centrale dell'intervallo modale} = \frac{x_i + x_{i+1}}{2}$$

$q_\alpha$  = quantile di ordine  $\alpha$ : è il numero reale che verifica la condizione:

$$\boxed{Fr\{X \leq q_\alpha\} = \alpha} \text{ dove}$$

$Fr\{X \leq q_\alpha\} = p_1 + p_2 + \dots + p_{i-1} + (q_\alpha - x_i)c_i$  (avendo scelto l'intervallo  $[x_i, x_{i+1})$  appropriato a partire dall'istogramma) da cui si ottiene:

$$q_\alpha = x_i + \frac{\alpha - (p_1 + p_2 + \dots + p_{i-1})}{c_i} = \text{quantile di ordine } \alpha$$

con  $\alpha = 0.25$ :  $q_{0.25} = Q_1$  = primo quartile

con  $\alpha = 0.5$ :  $q_{0.50} = Q_2$  =  $\text{med}(X)$  = secondo quartile = mediana

con  $\alpha = 0.75$ :  $q_{0.75} = Q_3$  = terzo quartile

$M(X)$  = media della v.s. continua  $X = M(X')$  dove  $X'$  è la vs. discreta ottenuta discretizzando la v.s. continua  $X$ . I valori  $x'_i$  ( $i = 1, 2, \dots, k$ ) della v.s.  $X'$  sono dati dai valori centrali degli intervalli ovvero  $x'_i = \frac{x_i + x_{i+1}}{2}$  ( $i = 1, 2, \dots, k$ ). La frequenza relativa di ciascun  $x'_i$  è la frequenza relativa del corrispondente intervallo

$$M(X^2) = \text{momento secondo della v.s. continua } X = M(X'^2)$$

$$V(X) = \text{varianza della v.s. continua } X = V(X') = M(X'^2) - M(X')^2$$

$$s.q.m.(X) = \text{scarto quadratico medio della v.s. continua } X = \sqrt{V(X')}$$