

**TECNICO DELLA PIANIFICAZIONE ECONOMICA  
E AMBIENTALE DELLE AREE PORTUALI**

**LEZIONE 26/10/05**

**STATISTICA**

Antigone Marino

antigone@na.infn.it

# LA DISPERSIONE O VARIAZIONE

L'attitudine di certi dati numerici a disporsi intorno ad un valore medio è detta **dispersione** o variazione dei dati.

Si possono ottenere varie misure della dispersione, le più comuni delle quali sono

- il campo di variazione;
- lo scostamento medio assoluto dalla media aritmetica;
- l'intervallo tra il 10° ed il 90° percentile;
- lo scarto quadratico medio.

# IL CAMPO DI VARIAZIONE

Il **campo di variazione** di un insieme di numeri è la differenza tra il più grande e il più piccolo dei numeri stessi.

ESEMPIO: il campo di variazione dei numeri

2, 3, 3, 5, 8, 7, 5, 8, 10, 12

è dato dalla differenza

$$12 - 2 = 10$$

# LO SCOSTAMENTO MEDIO ASSOLUTO

Lo **scostamento semplice medio assoluto** dalla media aritmetica di un insieme di  $N$  numeri  $X_1, X_2, \dots, X_N$  è definito da

$$\bar{\sigma}_{\text{assoluto}} = \frac{\sum_{j=1}^N |X_j - \bar{X}|}{N} = \frac{\sum |X - \bar{X}|}{N} = \overline{|X - \bar{X}|}$$

dove  $\bar{X}$  è la media aritmetica dei numeri e  $|X_j - \bar{X}|$  è il valore assoluto della deviazione di  $X_j$  da  $\bar{X}$ .

# ESEMPIO

Trovare lo scostamento semplice medio assoluto dei numeri

**2, 3, 6, 8, 11**

La media aritmetica di questi numeri è

$$\bar{X} = \frac{2+3+6+8+11}{5} = 6$$

Quindi lo scostamento semplice medio assoluto dalla media è

$$\begin{aligned}\bar{\sigma} &= \frac{|2-6| + |3-6| + |6-6| + |8-6| + |11-6|}{5} = \\ &= \frac{|-4| + |-3| + |0| + |2| + |5|}{5} = \frac{4+3+0+2+5}{5} = \frac{14}{5} = 2,8\end{aligned}$$

# L'INTERVALLO tra il 10° E 90° PERCENTILE

**L'intervallo tra il 10° e il 90° percentile** di un insieme di dati è definito come.

$$\text{Intervallo tra il } 10^\circ \text{ e il } 90^\circ \text{ percentile} = P_{90} - P_{10}$$

Dove  $P_{10}$  e  $P_{90}$  sono il decimo ed il novantesimo percentile dei dati.

Si può anche usare il semi-intervallo tra il 10° e il 90° percentile, definito come

$$\frac{1}{2}(P_{90} - P_{10})$$

# SCARTO QUADRATICO MEDIO

Lo **scarto quadratico medio** di un insieme di  $N$  numeri  $X_1, X_2, \dots, X_N$  è definito da

$$\sigma = \sqrt{\frac{\sum_{j=1}^N (X_j - \bar{X})^2}{N}} = \sqrt{\frac{\sum (X - \bar{X})^2}{N}} = \sqrt{(X - \bar{X})^2}$$

Se a  $X_1, X_2, \dots, X_K$  sono associate le frequenze  $f_1, f_2, \dots, f_K$  rispettivamente, allora

$$\sigma = \sqrt{\frac{\sum_{j=1}^K f_j (X_j - \bar{X})^2}{N}} = \sqrt{\frac{\sum f (X - \bar{X})^2}{N}} = \sqrt{(X - \bar{X})^2}$$

# LA VARIANZA

La **varianza** di un insieme di dati è definita come il quadrato dello scarto quadratico medio.

$$\sigma^2 = \frac{\sum_{j=1}^N (X_j - \bar{X})^2}{N}$$

$$\sigma^2 = \frac{\sum_{j=1}^K f_j (X_j - \bar{X})^2}{N}$$

# DISPERSIONE ASSOLUTA E RELATIVA

Le dispersioni reali, come vengono determinate dallo scarto quadratico medio o altre misure di dispersione, vengono dette **assolute**.

Ma una dispersione di *10cm* nella misurazione di una distanza pari a *1000m* ha molto meno importanza della stessa dispersione di *10cm* nella misura di una distanza pari a *50m*.

Perciò si ricorre alla dispersione relativa, definita dalla relazione

$$\text{Dispersione relativa} = \frac{\text{Dispersione assoluta}}{\text{Media}}$$

# COEFFICIENTE DI VARIAZIONE

Se la dispersione assoluta è lo scarto quadratico medio e la media è la media aritmetica, allora la dispersione relativa è detta coefficiente di variazione o **coefficiente di dispersione**.

Tale coefficiente è dato dalla relazione,

$$\text{Coefficiente di dispersione} = \frac{\sigma}{\bar{X}}$$

E generalmente viene espresso in percentuale.

Notare che il coefficiente di dispersione è indipendente dall'unità di misura utilizzata, quindi si può utilizzare per confrontare le distribuzioni in cui le unità di misura siano differenti.

# VARIABILE STANDARDIZZATA

La variabile

$$z = \frac{X - \bar{X}}{\sigma}$$

che misura le deviazioni dalla media e dove l'unità di misura è lo scarto quadratico medio, è detta **variabile standardizzata** ed è una quantità non dimensionale (cioè indipendente dall'unità di misura usata).