

# Glossario

## ❑ CAMPO DI VARIAZIONE

rappresenta la differenza tra la più grande e la più piccola modalità osservate all'interno della distribuzione; è una misura assoluta di variabilità; risente molto della presenza di "valori anomali" all'interno della distribuzione.

## ❑ DIFFERENZA INTERQUARTILE

è data dalla differenza tra il terzo e il primo quartile della distribuzione; è una misura di variabilità assoluta; è preferibile rispetto al campo di variazione in quanto risente meno della presenza di "valori anomali".

## ❑ VARIABILE SCARTO

è data dalla differenza tra una singola modalità e la media aritmetica:  $(x_i - \bar{x})$ .

## ❑ DEVIANZA

è data dalla somma degli scarti elevati al quadrato:  $\sum_{i=1}^n (x_i - \bar{x})^2$ ; in sostanza, si tratta del numeratore della varianza.

## ❑ VARIANZA

è la principale misura di variabilità che abbiamo studiato. Viene calcolata come rapporto tra la sommatoria di tutti gli scarti elevati al quadrato (ossia, la devianza) e la numerosità

$$\text{totale: } \sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

## ❑ SCARTO QUADRATICO MEDIO (o DEVIAZIONE STANDARD)

è dato dalla radice quadrata della varianza:  $\sigma = \sqrt{\sigma^2}$ ; viene calcolato per riportare i dati all'unità di misura originaria.

#### ❑ **MASSIMA VARIABILITA' ( $\sigma_{\max}$ )**

è il massimo valore teorico raggiungibile da  $\sigma$  in una determinata distribuzione. È dato dal prodotto tra la media aritmetica e la radice quadrata di  $(n-1)$ . Ci è utile per giudicare la rappresentatività della media aritmetica.

#### ❑ **COEFFICIENTE DI VARIAZIONE**

è dato dal rapporto tra lo s.q.m. e la media aritmetica, moltiplicato 100:  $CV = \frac{\sigma}{x} 100$ ; viene calcolato per confrontare la variabilità di differenti distribuzioni. Più è elevato, maggiore è la variabilità della distribuzione.

#### ❑ **CARATTERE TRASFERIBILE**

un carattere è detto trasferibile quando può essere ceduto, in tutto o in parte, da un'unità statistica ad un'altra

#### ❑ **CARATTERE EQUIDISTRIBUITO**

quando tutte le unità statistiche possiedono lo stesso ammontare del carattere

#### ❑ **CONCENTRAZIONE**

un carattere (trasferibile) si dice concentrato quando una gran parte dell'ammontare complessivo è nelle "mani" di pochi soggetti (al limite, di uno solo)

#### ❑ **MINIMA CONCENTRAZIONE**

quando tutte le unità statistiche possiedono lo stesso ammontare del carattere (ossia, il carattere è EQUIDISTRIBUITO)

#### ❑ **MASSIMA CONCENTRAZIONE**

quando l'ammontare complessivo del carattere è detenuto da una sola unità statistica

### ❑ RAPPORTO DI CONCENTRAZIONE DI GINI (R)

È l'indice che misura la concentrazione di un carattere; è dato dalla formula:

$$R = \frac{\sum_{i=1}^{n-1} (F_i - Q_i)}{\sum_{i=1}^{n-1} F_i};$$

tale indice può variare tra 0 (nel caso di minima concentrazione) ed 1

(nel caso di massima concentrazione)

### ❑ STANDARDIZZAZIONE

È una trasformazione lineare dei dati che rende le informazioni tratte da diverse

distribuzioni tutte confrontabili. La formula è la seguente:  $Z = \frac{x_i - \bar{x}}{\sigma}$  e l'elemento

fondamentale è che una distribuzione (qualsiasi essa sia) una volta che i dati sono stati standardizzati si trova ad avere media pari a zero e varianza pari a uno.

### ❑ GRAFICO BOX PLOT

È una rappresentazione grafica che mostra da un punto di vista visivo tutte le caratteristiche fondamentali di una distribuzione (minimo, massimo, valore centrale e variabilità)