

Riassunti numerici dei dati

Misure di tendenza centrale

- media
- moda
- medianac

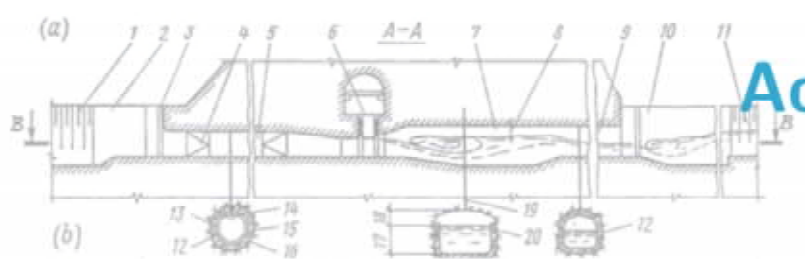
Misure di dispersione

+

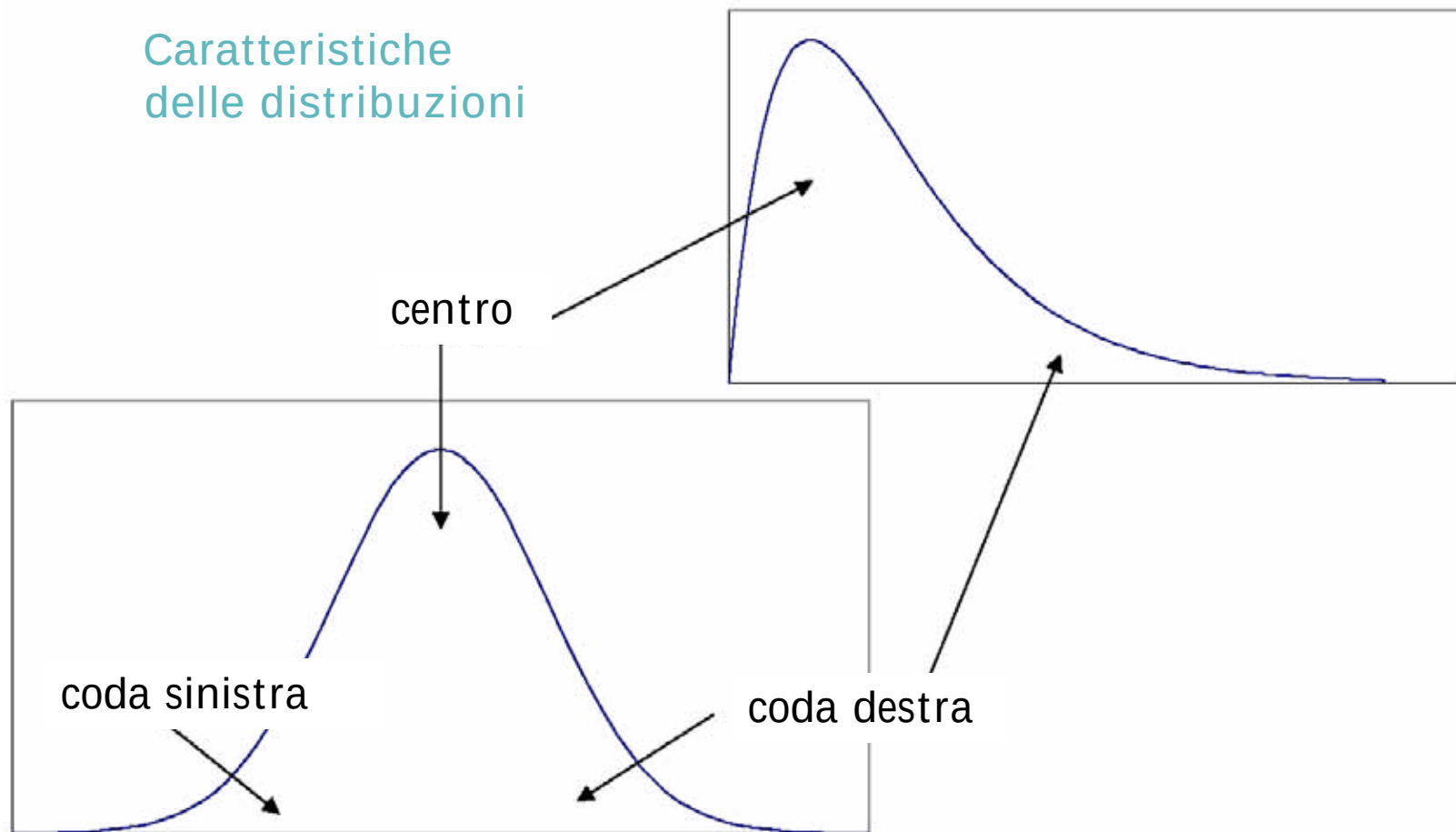
Coefficiente di curtosi:
altezza relativa del picco
richiede un campione ampio
per distribuzioni simmetriche

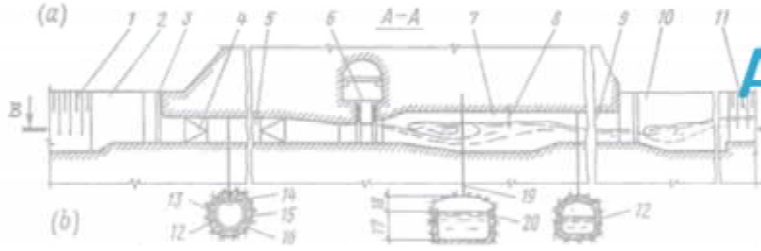
Misure di asimmetria

Coefficiente
di asimmetria

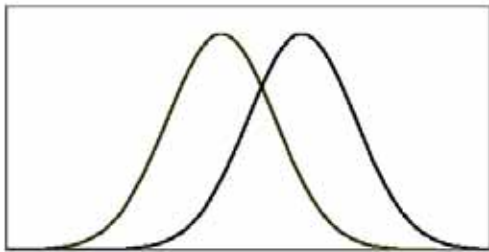


Caratteristiche delle distribuzioni

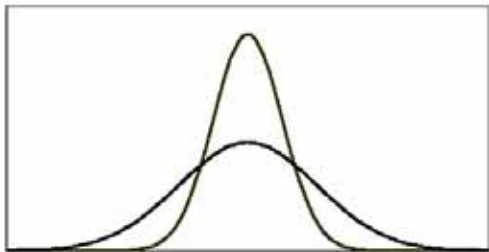




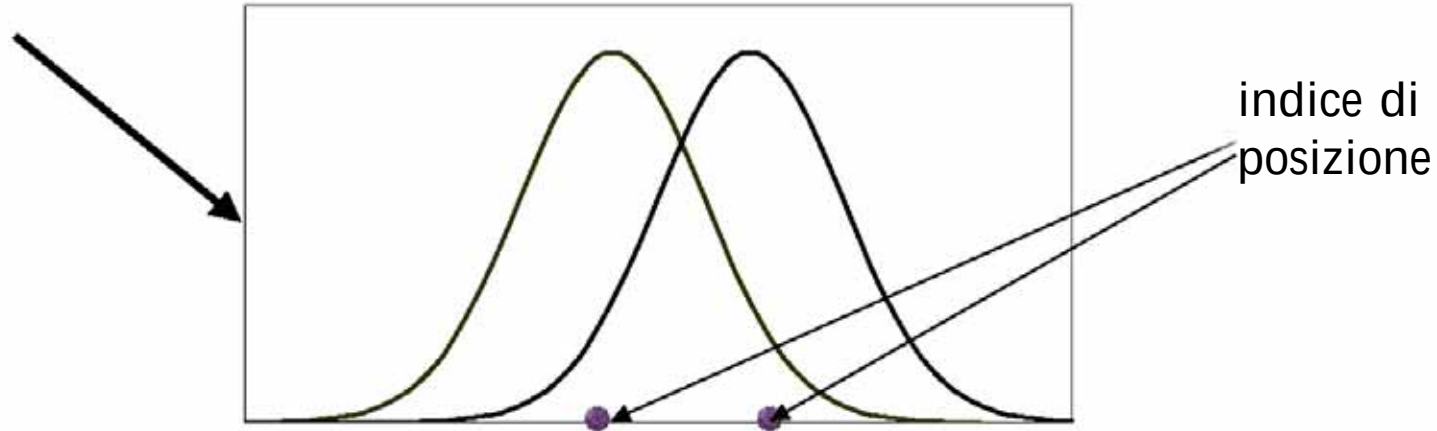
posizione

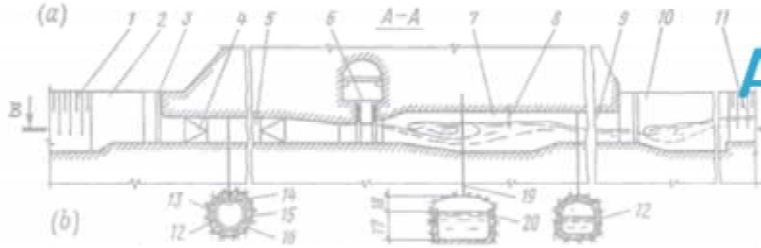


variabilità



simmetria





Misure di tendenza centrale

Media empirica

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Media Spuntata (trimmed mean)

Calcolata considerando solo il 90% centrale dei Dati (cioè compresi tra il 5% ed il 95% dei dati ordinati)

moda

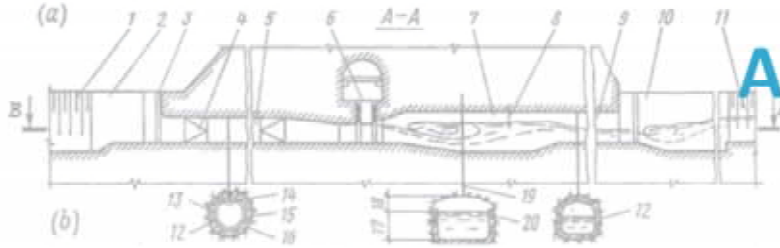
Valore/i con frequenza massima

mediana

E' il valore dell'osservazione per cui nel campione ci sono il 50% delle osservazioni minori o uguali a questa

La media empirica, essendo il baricentro dei dati, risente molto della posizione Dei valori estremi

La mediana invece, non è assolutamente influenzata dagli estremi



Misure di dispersione

Varianza empirica:

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

Uno stimatore più robusto della varianza è:

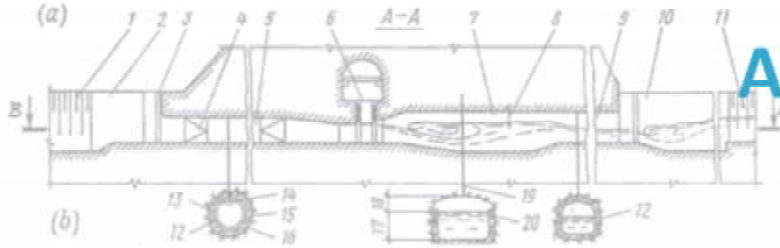
$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Scarto: $\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \approx \tilde{\sigma}_{\frac{1}{2}}$

Scarto spuntato...

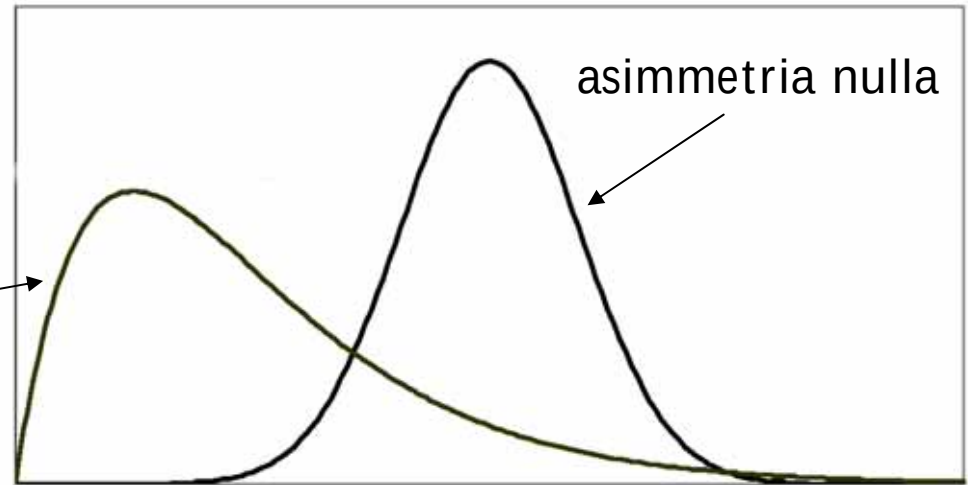
Range: (valore massimo-valore minimo)

Coefficiente di variazione: $cv = \frac{s}{\bar{x}}$ con $\bar{x} \neq 0$



Misure di Asimmetria

asimmetria positiva



Coefficiente di asimmetria
o skewness

$$g_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{ns^3}$$

misura l'asimmetria rispetto alla media

Coefficiente di curtosi
o peakedness

$$g_2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{ns^4}$$

misura il "peso delle code"