

Statistica

negli esperimenti reali si effettuano sempre un numero *finito* di misure, (spesso molto limitato) ➤ l'insieme delle misure effettuate costituisce il *campione* statistico

Statistica descrittiva e statistica inferenziale o "predittiva"

- la statistica descrittiva studia i criteri di rilevazione, di classificazione e di sintesi delle informazioni relative ad una "popolazione" oggetto di studio
- la statistica predittiva si prefigge di "*stimare*" caratteristiche della "popolazione" incognita basandosi sulle informazioni fornite da un campione limitato di misure ripetute

in questo contesto si ipotizza che il comportamento della popolazione rispetto ad una variabile aleatoria X venga descritto attraverso una funzione parametrica di probabilit 

$p_X(x | \theta)$ o di densit  di probabilit  $f_X(x | \theta)$ dove θ   *incognito* o pi  in generale $f_X(x | \theta_1, \theta_2, \dots, \theta_n)$ dove θ_i sono i parametri *incogniti*

si devono fare previsioni quando per qualche motivo non si possono conoscere i dati relativi a tutta la popolazione parente, ma solo quelli relativi ad un campione rappresentativo di n misurazioni

$X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$ e attraverso la conoscenza del campione si cerca di stimare la validit  di una determinata congettura sul parametro incognito

per stimare il **valor medio** della popolazione parente di solito si fa uso dello "stimatore" media aritmetica
$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

altri stimatori di centralit  sono la

media geometrica
$$\bar{x}_g = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} = \left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}}$$
 ma con la limitazione che tutti gli x_i devono essere positivi

significato geometrico: la media geometrica di due numeri   la lunghezza del lato di un quadrato equivalente (= stessa area) ad un rettangolo che abbia i lati pari al modulo dei due numeri

media armonica $\overline{x}_a^{-1} = \frac{1}{n} \left(\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n} \right)$ ma con la limitazione che tutti gli x_i non devono essere nulli si utilizza quando entra in gioco la somma di reciproci

media quadratica $\overline{x}_q = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$ ila media quadratica e' sempre maggiore della media aritmetica

mediana campionaria = 50-esimo percentile

moda compionaria = valore piu' probabile

➤ la media aritmetica e' lo stimatore piu' usato per la stima del valor medio per le sue proprieta' di correttezza, consistenza ed efficienza

Backup Slides