

Introduzione alla probabilità e all'inferenza statistica

note introduttive

Liceo Pascoli – Firenze, 16 gen 2008

Che cos'è la STATISTICA ?

A che servono gli statistici ?

Liceo Pascoli – Firenze, 16 gen 2008

Che cos'è la statistica ?

per **Statistica** si intendeva in origine la raccolta di dati economico-demografici di vitale interesse per lo Stato. Da allora la statistica si è sviluppata in un metodo scientifico d'analisi, oggi applicato a tutte le scienze sociali e naturali.

Definizione formale:

la **Statistica** è *la disciplina che si occupa dell'elaborazione dei risultati dell'osservazione di uno o più caratteri posseduti dagli elementi di un insieme determinato, con l'intento di esprimere un giudizio su alcuni aspetti di una realtà di interesse che, in quanto riferita ad un insieme e non ai singoli elementi che lo compongono, viene chiamata **fenomeno collettivo**.*

Liceo Pascoli – Firenze, 16 gen 2008

Che cos'è la statistica ?

In altre parole:

La Statistica è Matematica Applicata volta all'analisi dei dati concernenti un particolare fenomeno, dati che possono derivare da:

- studi osservazionali
- studi di sperimentazione

Per **dato osservazionale** si intende il dato rilevato su un fenomeno preconstituito, esistente in natura e generalmente finito.

Per **dato sperimentale** si intende invece un dato creato in circostanze controllate, mediante una replicazione dell'esperimento potenzialmente infinita.

Liceo Pascoli – Firenze, 16 gen 2008

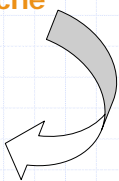
L'essenza della Statistica

La ragion d'essere della statistica è la presenza di un certo livello di variabilità nei dati

(altrimenti, paradossalmente, la comprensione del fenomeno sarebbe possibile ricorrendo ad una sola osservazione)

nello studio dei fenomeni collettivi si è consapevoli che **al variare dell'unità statistica u entro una certa popolazione $P = \{u\}$ variano certe caratteristiche misurate su u**

in altre parole,
per lo studio di un fenomeno caratterizzato da assenza di variabilità **non serve scomodare uno statistico**



Liceo Pascoli – Firenze, 16 gen 2008

L'essenza della Statistica

Alcuni semplici esempi:

- altezza e peso degli studenti di una classe
- reddito dei parlamentari
- votazioni riportate all'esame di Statistica dagli studenti di un certo corso di Laurea
- valutazione dell'efficacia dei titoli di studio universitari
- soddisfazione nei confronti del trasporto pubblico locale
- durata delle lampadine ad alto risparmio energetico
- ecc...

Liceo Pascoli – Firenze, 16 gen 2008

Il problema della misura

Conversione dell'oggetto di studio in **dato statistico**:

problema del tipo di misura che si intende adottare come espressione delle manifestazioni rilevate

- altezza → metro
- soddisfazione → ????

Misura di tutte le unità statistiche
(analisi costo/tempo, impossibilità a procedere...)



Misura sola di una (piccola) parte di esse

Liceo Pascoli – Firenze, 16 gen 2008

Le branche della STATISTICA

Statistica DESCRITTIVA

In questo settore rientrano i metodi per **sintetizzare** con opportune grandezze le caratteristiche salienti dei fenomeni collettivi.

La **descrizione** passa attraverso le fasi di formazione del dato statistico e del trattamento matematico dello stesso.

Per formazione del dato statistico si può intendere:

- l'elaborazione di dati preesistenti in natura (dati anagrafici, indici aziendali di bilancio, dati di produzione industriale ecc.)
- la necessità di procedere **all'effettiva rilevazione delle informazioni** necessarie alla comprensione del fenomeno di interesse

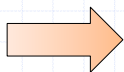
Liceo Pascoli – Firenze, 16 gen 2008

Il processo di rilevazione dei dati

Il **processo di rilevazione dei dati** è generalmente distinto nelle fasi di:

- **definizione del piano di rilevazione,**
- **raccolta delle informazioni,**
- **spoglio e classificazione.**

La fase più delicata è senza dubbio la prima.



- **Rilevazioni complete (censuarie)**
- **Rilevazioni campionarie**

Liceo Pascoli – Firenze, 16 gen 2008

Le rilevazioni campionarie

Fondamentale diviene quindi in statistica il ruolo dell'esperimento campionario.

Si dice **campione** un aggregato di unità statistiche appartenenti ad una certa popolazione e selezionate mediante un esperimento di campionamento (generalmente collegato ad una procedura di randomizzazione).

Si osservi che la strategia campionaria è la sola possibile quando la popolazione è virtualmente infinita; inoltre la popolazione da cui si estrae il campione, detta popolazione campionata, non sempre coincide con la popolazione obiettivo.

Liceo Pascoli – Firenze, 16 gen 2008

Le rilevazioni campionarie

Tra i vari tipi di campionamento una menzione particolare spetta al **campione casuale semplice CCS** che è una forma di campionamento **equiparabile allo schema dell'urna** in cui le palline - che identificano le diverse unità della popolazione - **hanno tutte la stessa probabilità di essere estratte.**

Si distingue fra:

- **CCS con reimmissione**
- **CCS senza reimmissione.**

Si può in generale affermare che un campione non costituisce quasi mai una riproduzione fedele della popolazione su piccola scala.

Liceo Pascoli – Firenze, 16 gen 2008

Le branche della STATISTICA

Statistica INFERENZIALE

Tuttavia se l'estrazione è casuale, i dati possono fornire informazioni sulla variabilità della popolazione e sulla fiducia da accordare a tali informazioni. Questi problemi sono oggetto della Statistica Inferenziale o Induttiva.

Il termine inferenza deriva dal latino e letteralmente significa: argomentare, desumere.

Si effettua inferenza quando si generalizza l'esperimento operando una sorta di estensione dal particolare al generale; le generalizzazioni però non sono certe.

Liceo Pascoli – Firenze, 16 gen 2008

Statistica inferenziale e Probabilità

L'inferenza è quindi un processo d'azzardo e l'incertezza viene misurata in termini probabilistici.



La **PROBABILITÀ** è il fondamento logico per fare inferenza sulla Popolazione oggetto d'indagine.

Ma ...

che cos'è la PROBABILITÀ?

La Probabilità

La **PROBABILITÀ** è un concetto primitivo.

Per definirla occorre introdurre alcuni ingredienti:

- esperimento casuale**
- evento**
- spazio degli eventi B**

es: lancio del dado o di una moneta

La probabilità è una funzione matematica su B con certe proprietà

La Probabilità

Approcci alla probabilità:

- **impostazione classica;**
es: moneta
- **impostazione frequentista;**
es: moneta truccata
- **impostazione soggettiva;**
es: uomo su Marte

- **impostazione Assiomatica**

L'impostazione assiomatica delle Probabilità

- 1) $P(A) \geq 0$
- 2) $P(\Omega) = 1$
- 3) $P(A \cup B) = P(A) + P(B)$
se $A \cap B = \emptyset$

$$4) P(A / B) = \frac{P(A \cap B)}{P(B)}$$

Formalizzazione matematica di concetti intuitivi

Un po' di svago

- **Dado dietro la porta**
- **Gioco del lotto:** prob. che esca 23 al secondo estratto
- **Mazzo di 40 carte:** prob. che esca un K alla seconda estraz.
- **I 4 prigionieri**
- **Le 3 buste**
- **35 studenti su un autobus:** prob. che almeno 2 abbiano stessa data di nascita
- Il valore atteso ed il **Paradosso di San Pietroburgo**
- Come misurare l'**area di un lago**