

3. DI COSA TRATTA LA STATISTICA: STATISTICA DESCRITTIVA E STATISTICA INFERENZIALE

Un'indagine statistica si occupa di un **"FENOMENO COLLETTIVO"** (cioè di un fenomeno che si presenta in una pluralità di soggetti ... sovente, in *tanti o tantissimi* soggetti).

ESEMPI DI FENOMENI COLLETTIVI

- La conoscenza delle lingue straniere.
- I gusti musicali.
- L'età alla quale ci si sposa.
- La presenza di animali domestici negli appartamenti.
- La lunghezza delle piste ciclabili nelle città.

Se, ad esempio, noi fossimo interessati al fenomeno "utilizzo del telefonino da parte degli studenti di una data scuola", gli studenti di quella scuola sarebbero la nostra *"popolazione statistica"*, ogni singolo studente sarebbe una *"unità statistica"*, e i *"caratteri"* da studiare sarebbero "quanto, quando e come questi ragazzi utilizzano il telefonino".

Un "fenomeno collettivo" viene preso in esame nell'ambito di una data **"POPOLAZIONE STATISTICA"**, della quale si studia una certa caratteristica o **"CARATTERE"** (NOTA) andando ad analizzare quali sono le **"MODALITÀ"** con cui questo carattere si può manifestare e verificando, in ciascun elemento (= **"UNITÀ STATISTICA"**) della "popolazione", quale di tali modalità è presente, per desumere da tutto ciò **conteggi, percentuali, "medie", "indici di dispersione", rappresentazioni grafiche**.
NOTA: sovente si studiano, "in parallelo", sulla stessa popolazione statistica, *più* caratteri di uno stesso fenomeno

ESEMPI

<p>a) Fenomeno collettivo: le caratteristiche fisiche.</p> <p>Possibili "popolazioni statistiche":</p> <ul style="list-style-type: none"> <input type="checkbox"/> l'insieme dei residenti in Italia, <input type="checkbox"/> o l'insieme dei cittadini di nazionalità italiana, <input type="checkbox"/> oppure l'insieme delle donne nate a Stoccolma <p>Un possibile carattere: il colore degli occhi.</p> <p>Possibili scelte per le modalità di questo carattere:</p> <ul style="list-style-type: none"> <input type="checkbox"/> marrone <i>oppure:</i> <input type="checkbox"/> marrone <input type="checkbox"/> azzurro <input type="checkbox"/> non marrone <input type="checkbox"/> verde <input type="checkbox"/> grigio <p>E' una "unità statistica" la singola persona della quale si rileva il colore degli occhi.</p>	<p>b) Fenomeno collettivo: il livello di istruzione.</p> <p>Possibili "popolazioni statistiche":</p> <ul style="list-style-type: none"> <input type="checkbox"/> l'insieme dei cittadini italiani dai 30 anni ai 40 <input type="checkbox"/> oppure l'insieme dei Vigili del Fuoco <p>Un possibile carattere: l'attestato scolastico più alto conseguito.</p> <p>Una possibile scelta per le modalità di questo carattere:</p> <ul style="list-style-type: none"> <input type="checkbox"/> licenza elementare <input type="checkbox"/> licenza media <input type="checkbox"/> diploma di scuola media superiore <input type="checkbox"/> laurea <p>E' una "unità statistica" la singola persona di cui constatiamo il grado di istruzione.</p>
<p>c) Fenomeno collettivo: il costo degli affitti</p> <p>Una possibile "popolazione statistica":</p> <ul style="list-style-type: none"> <input type="checkbox"/> l'insieme degli appartamenti in affitto in una determinata località <p>Due possibili caratteri:</p> <ul style="list-style-type: none"> <input type="checkbox"/> la cifra risultante dal contratto d'affitto <input type="checkbox"/> oppure la cifra complessiva, compresi gli accordi "in nero" 	<p>(problema: come riuscire a rilevare la "vera" cifra?)</p> <p>Una possibile scelta per le modalità del carattere:</p> <ul style="list-style-type: none"> <input type="checkbox"/> meno di 300 euro mensili <input type="checkbox"/> da 300 a 449 euro <input type="checkbox"/> da 450 a 599 euro <input type="checkbox"/> dai 600 euro in su <p>E' una "unità statistica" il singolo appartamento per il quale si va ad annotare il costo del relativo affitto.</p>

Come mostra l'esempio c), una "popolazione statistica" *non* deve essere necessariamente un gruppo di *persone*!
Un'ulteriore situazione: se interessasse lo studio del numero di piccoli che le coniglie di un dato allevamento generano nel corso della loro vita, una "unità statistica" sarebbe una coniglia.

OSSERVAZIONI

Una volta deciso il "carattere" di cui ci vogliamo interessare, la scelta delle sue "modalità" non è univoca! Così come spetta a noi, secondo i nostri interessi o le nostre esigenze, la scelta della "popolazione statistica".

SINONIMI

"Popolazione statistica" = "popolazione" = "COLLETTIVO statistico" = "collettivo" = "universo".

Il numero delle unità statistiche si dice anche **"NUMEROSITÀ"** della popolazione.

LE DIVERSE TIPOLOGIE DI CARATTERI

I “**caratteri**” si distinguono fra

- “**QUALITATIVI**” (= le cui modalità sono espresse da un aggettivo, da un sostantivo o da un avverbio, es. *il colore dei capelli, il grado di soddisfazione rispetto a un prodotto*)
- e “**QUANTITATIVI**” (= modalità espresse da un numero, es. *il peso o il reddito di una persona, oppure il numero di uova prodotte da una gallina in un mese*).

Fra i caratteri “**QUALITATIVI**”, distinguiamo

- quelli “**ORDINATI**”
- e quelli “non ordinati”, o “**SCONNESSI**”.

Ad es., è “ordinato” il carattere “*livello di istruzione*”. E’ invece “sconnesso” il carattere “*colore dei capelli*”.

Fra i caratteri “**QUANTITATIVI**”, ce ne sono di

- “**DISCRETI**”: quelli che sono descritti da **numeri interi**, come ad es. il “*numero dei figli*” di una donna, o il “*numero di esami già superati*” da un universitario
- e “**CONTINUI**”: quelli espressi, almeno in linea di principio, da un **numero reale** ... ma, soprattutto, quelli dei quali interessa *non tanto* il valore preciso, *quanto* il fatto se siano compresi in un dato **intervallo**.

Ad esempio, l’ “*area di una superficie coltivata*”;

ma anche *il peso o l’altezza individuali, il tempo che si impiega per percorrere una certa distanza, la larghezza delle strade*, sono da considerarsi caratteri “continui”.

Rifletti: se chiediamo a una persona di darci la sua altezza in cm, e questa ci risponde “171”,

vuol dire che la sua altezza rientra nella fascia fra 170,5 e 171,5 e che quella persona ha scelto l’intero 171 in quanto ha valutato questo valore “tondo” come il più vicino alla “vera” misura, che però intera non sarà ...

DEFINIZIONI DI “STATISTICA”; SUA IMPORTANZA; STATISTICA DESCRITTIVA E INFERENZIALE

Possiamo dire che la **statistica** è la disciplina che, **innanzitutto, insegna ad esprimere le caratteristiche salienti di un insieme di dati, anche molto vasto, in modo sintetico, con l’aiuto di numeri dotati di valore “riassuntivo” (le “medie”, gli “indici di dispersione”)** e con il supporto di rappresentazioni grafiche svariate.

- La parte più elementare della statistica è la cosiddetta “**STATISTICA DESCRITTIVA**”. La statistica *descrittiva* **analizza TUTTE le unità statistiche della popolazione considerata**.
- La “**STATISTICA INFERENZIALE**” si occupa invece di estrarre dalla popolazione, quando questa è troppo vasta per poter essere studiata nella sua interezza, un sottoinsieme di unità statistiche detto “**campione**”, che verrà esaminato allo scopo di dedurre (= **inferire**) da questa analisi parziale indicazioni che possano valere per la **popolazione intera**, valutando il **grado di attendibilità** di tali indicazioni.

Ballatori, nel 1980, definì la statistica come

“**disciplina che studia i fenomeni collettivi, cioè quei fenomeni per la cui osservazione (descrizione, misura) è necessaria una massa di osservazioni di fenomeni elementari**”.

E Maccacaro, nel 1975, la definì come il “**saper parlare di ciò che non si conosce: o meglio, parlare correttamente di ciò che non si conosce completamente**”.

NEL MONDO CONTEMPORANEO una grandissima mole di dati viene quotidianamente rilevata e messa a disposizione del pubblico. D’altronde l’elaborazione e la rappresentazione grafica dei dati stessi sono al giorno d’oggi enormemente facilitate e rese veloci dall’utilizzo del **computer**.

In questo contesto **LA STATISTICA ASSUME UN’IMPORTANZA COLOSSALE**.

E’ ormai irrinunciabile anche per il profano saper interpretare, per esempio,

l’**attendibilità dei sondaggi**, o i **grafici** e gli **indici** che continuamente giornali e TV ci propongono.

E la **scienza** non potrebbe progredire senza **strumenti statistici** in grado

di **sintetizzare le informazioni** e di effettuare **previsioni e valutazioni di affidabilità!**

Due esempi soltanto fra i tantissimi possibili:

- 1) la **teoria degli errori di misura**;
- 2) nelle sperimentazioni cliniche di un farmaco su un campione di malati, il “**p-value**”, valore che esprime la probabilità che il buon effetto riscontrato sia dovuto al caso, anziché a una reale efficacia del farmaco (di solito vengono considerati significativi per il test di efficacia valori $p < 0,05$).

ORIGINE DEL TERMINE “STATISTICA”

Proviene dalla lingua **italiana** (Ghislini, 1589) e precisamente dalla parola “**Stato**”, con riferimento all’utilizzo dei rudimenti di questa disciplina, nel XVI secolo, per studi amministrativi e politici.