

**Definizione di STATISTICA:** La statistica si può definire come un insieme di metodi scientifici finalizzati alla conoscenza quantitativa e qualitativa dei fenomeni collettivi mediante la raccolta, l'ordinamento, la sintesi e l'analisi dei dati.

I fenomeni collettivi sono quei fenomeni il cui studio richiede una pluralità di osservazioni.

## **Statistica DESCRITTIVA e Statistica INFERENTE**

**Statistica DESCRITTIVA:** l'oggetto dell'analisi è la *popolazione*.

La Statistica DESCRITTIVA ha come finalità la presentazione di un certo fenomeno in forma sintetica.

**Statistica INFERENTE:** l'oggetto dell'analisi è un *campione*.

La Statistica INFERENTE ha come finalità di trarre le caratteristiche di una popolazione sulla base di un numero limitato di osservazioni.

**Le fasi di un'indagine statistica:**

- 1)la formulazione del problema;**
- 2)la rilevazione dei dati;**
- 3)lo spoglio e la classificazione dei dati;**
- 4)l'elaborazione;**
- 5)la presentazione dei risultati.**

## FASE 1) la formulazione del problema

→ definire l'OGGETTO dell'analisi: *si definisce **unità statistica** il caso individuale componente del collettivo statistico.*

→ definire la “popolazione” da analizzare: *si definisce **popolazione o collettivo statistico** l'insieme dei casi individuali presi come riferimento per studiare il fenomeno in esame.*

→ scegliere il carattere (o caratteri) da prendere in considerazione e le sue modalità.

Il carattere è l'aspetto elementare rispetto al quale viene considerata la singola unità statistica.

Le modalità sono i possibili modi di presentarsi del carattere oggetto di analisi.

Il **dato** rappresenta l'unità elementare da cui ha origine ogni indagine statistica: il dato è quel valore osservato che viene rilevato in corrispondenza al carattere scelto di volta in volta per studiare il fenomeno in esame.

# **CLASSIFICAZIONE DEI DATI**

**mutabile (carattere qualitativo):**

viene espressa in forma verbale

**variabile (carattere quantitativo):**

viene individuata mediante dei valori, indicanti la misura dell'intensità con cui quel carattere si manifesta

**MUTABILE: scala nominale - scala ordinale**

## **VARIABILE DISCRETA O CONTINUA**

### **ESERCIZIO**

Stabilire la tipologia dei seguenti caratteri (qualitativo nominale / qualitativo ordinale/ quantitativo discreto/ quantitativo continuo):

|   |  |
|---|--|
| Prezzo di una pizza                     |  |
| Partito politico di appartenenza        |  |
| Reddito della famiglia                  |  |
| Voto all'esame di statistica            |  |
| Tasso di interesse annuale              |  |
| Grado di soddisfazione di un cliente    |  |
| Numero della scarpa                     |  |
| Altezza                                 |  |
| Marca di caffè preferita                |  |
| Regione di provenienza                  |  |
| Tempo giornaliero dedicato alla lettura |  |
| Prodotto interno lordo dell'Italia      |  |

**Variabile discreta: DATI SEMPLICI** (: *distribuzione disaggregata*) **O PONDERATI** (: *distribuzione di frequenze*)

**Variabile continua: dati suddivisi in classi.**

- Definizione: si chiama *distribuzione di frequenze di una variabile suddivisa in classi* lo schema con cui si associa a ciascuna classe la rispettiva frequenza.
- Definizione: si chiama *densità di frequenza della classe* il rapporto fra la frequenza della classe e la sua ampiezza. Il numero che si ottiene indica quante sono, in media, le unità statistiche per ogni unità del carattere.

## **FASE 2)la rilevazione dei dati**

- in base al tempo: rilevazione CONTINUA; rilevazione PERIODICA; rilevazione OCCASIONALE.
- In base all'oggetto dell'indagine statistica: rilevazione TOTALE o rilevazione CAMPIONARIA.

## **FASE 3) spoglio e classificazione dei dati e la loro rappresentazione grafica**

Spoglio manuale o spoglio automatico



TABELLE STATISTICHE o MATRICI dei DATI: tabelle semplici o tabelle a doppia entrata

**Distribuzione semplice di frequenze e Distribuzione congiunta di frequenze (: tabella a doppia entrata)**

**Distribuzione di frequenze assolute, relative e cumulate**



rappresentazione grafica:

Rappresentazione di dati ponderati: *grafico per ordinate*.

Rappresentazione grafica di dati divisi in classi: *istogramma*.

## NOTAZIONE

Frequenza assoluta  $f_i$  (notazione Cicchitelli  $n_i$ )

Frequenza relativa  $f_i^R$  (notazione Cicchitelli  $f_i$ )

Frequenza cumulata  $F_i$  (notazione Cicchitelli  $N_i$ )

### ESERCIZIO

Un'indagine sul fumo ha portato i seguenti risultati

| Classe d'età | Fumatore | Sesso  |         |
|--------------|----------|--------|---------|
|              |          | Maschi | Femmine |
| Giovani      | sì       | 20     | 10      |
|              | no       | 70     | 20      |
| Adulti       | sì       | 30     | 30      |
|              | no       | 60     | 20      |
| Anziani      | sì       | 40     | 50      |
|              | no       | 80     | 70      |

- presentare le distribuzioni di frequenze univariate per sesso, fumo e classe d'età;
- costruire la tabella di frequenze a doppia entrata per classe d'età e sesso.

### SOLUZIONI

a)

| Sesso   | f(sesso) |
|---------|----------|
| maschi  | 300      |
| femmine | 200      |
| tot.    | 500      |

| Fumo | f(fumo) |
|------|---------|
| sì   | 180     |
| no   | 320     |
| tot. | 500     |

| cl.età | f(età) |
|--------|--------|
| giov.  | 120    |
| ad.    | 140    |
| anz.   | 240    |
| tot.   | 500    |



b)

| Classe<br>d'età | sesso  |         |
|-----------------|--------|---------|
|                 | maschi | femmine |
| Giovani         | 90     | 30      |
| Adulti          | 90     | 50      |
| Anziani         | 120    | 120     |
| tot.            | 300    | 200     |

## ESERCIZIO

Il numero di componenti per famiglia  $x$  rilevati in un'indagine statistica condotta su 50 famiglie è risultato il seguente

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 4 | 2 | 2 | 3 | 3 | 5 | 2 | 3 | 1 | 4 | 2 | 6 | 1 | 2 | 6 | 1 | 2 | 2 | 3 | 5 | 2 | 3 | 1 | 4 |
| 2 | 2 | 3 | 3 | 5 | 2 | 3 | 1 | 4 | 2 | 6 | 1 | 2 | 2 | 3 | 1 | 4 | 2 | 2 | 3 | 2 | 2 | 3 | 3 | 5 |

Della variabile  $x$  presentare la distribuzione di frequenze assolute, relative e relative cumulate.

## SOLUZIONI

| $x$ | $f(x)$ | $p(x)$ | $P(x)$ |
|-----|--------|--------|--------|
| 1   | 8      | 0,16   | 0,16   |
| 2   | 18     | 0,36   | 0,52   |
| 3   | 12     | 0,24   | 0,76   |
| 4   | 5      | 0,10   | 0,86   |
| 5   | 4      | 0,08   | 0,94   |
| 6   | 3      | 0,06   | 1      |
|     | 50     | 1      |        |

# FONTI STATISTICHE

## 1) FONTI STATISTICHE UFFICIALI

- ISTAT <http://www.istat.it>  
Riguardo a scuola, università e mondo del lavoro:  
[http://www.demo.istat.it/dat/db\\_siti/](http://www.demo.istat.it/dat/db_siti/)

- Camere di Commercio
- Banca d'Italia
- CNR
- ACI, INPS, INAIL
- Istituti universitari
- Ecc...

A livello internazionale:

- EUROSTAT <http://epp.eurostat.ec.europa.eu>
- ONU <http://unstats.un.org>

## 2) RILEVAZIONI “AD HOC”