

Le indagini statistiche

EDITRICE
LA SCUOLA

- 1 - La statistica
- 2 - La raccolta dei dati: le frequenze
- 3 - L'elaborazione dei dati: gli indici statistici
- 4 - La rappresentazione dei dati

La statistica

La **statistica** è la scienza che studia i **fenomeni collettivi** servendosi di metodi matematici specifici.

Qualche esempio....

- I contagi in un epidemia
- I diplomati in un certo Paese in un dato anno
- le migrazioni

Oppure in dimensioni ristrette...

- Promossi e bocciati in un anno scolastico in una data scuola
- I libri letti in un anno in una classe

La statistica

In Italia esiste l'**ISTAT, Istituto Nazionale di Statistica** che si occupa di raccogliere i dati a livello nazionale sui più diversi fenomeni e collabora con altri enti dello Stato



Sul portale ISTAT sono disponibili numerosi repertori di dati e pubblicazioni liberamente consultabili

<https://www.istat.it/it/>

Esiste il **corso di laurea in scienze statistiche** presente in quattro università italiane: *Bologna, Padova, Milano Bicocca, Roma la Sapienza* <https://www.universita.it/laurea-scienze-statistiche/>

indagine statistica, popolazione statistica

Studiando un fenomeno collettivo si effettua un' **indagine statistica**, che consiste in più fasi:

- **la raccolta dei dati**
- **l'elaborazione dei dati**
- **la rappresentazione dei dati e la loro interpretazione**

Si chiama **popolazione statistica** il collettivo che viene studiato. Se si considera solo un sottoinsieme della popolazione statistica, si dice che tale sottoinsieme è un **campione statistico**. Il singolo individuo della popolazione o del campione è detto **unità statistica**.

Dati qualitativi e dati quantitativi

I dati raccolti vengono chiamate **variabili statistiche** e sono di due tipi:

- variabili **qualitative o dati qualitativi** : espresse/i tramite una qualità (ad esempio il colore degli occhi di un gruppo di studenti)
- variabili **quantitative o dati quantitativi**: espresse/i da un numero (per esempio il numero di studenti con occhi azzurri)

La raccolta dei dati

I dati sono raccolti attraverso:

- Questionari, sondaggi, interviste
- Rilevazioni specifiche, come ad esempio per l'epidemia Covid - 19 il numero di tamponi effettuati da un ospedale o la CO₂ rilevata da una centralina

Costruire un questionario o uno strumento di rilevazione è l'operazione più complessa di un'indagine

Anche stabilire il campione statistico è un'operazione molto complessa

Se l'indagine si estende **all'intera popolazione di un Paese, si parla di CENSIMENTO**

Elaborazione dei dati – ordinamento

PARTIAMO DA UN PROBLEMA/INDAGINE

gli sport praticati da 20 alunni di una classe: ecco le risposte

Calcio, calcio, nuoto, judo, calcio, calcio, pallavolo, judo, calcio, basket, basket, pallavolo, basket, pallavolo, pallavolo, calcio, calcio, calcio, calcio

METTIAMO IN ORDINE CRESCENTE

Nuoto - judo, judo -basket, basket, basket - pallavolo, pallavolo, pallavolo, pallavolo - calcio, calcio, calcio, calcio, calcio, calcio, calcio, calcio, calcio, calcio

In totale abbiamo 20 «voci/dati», ma alcuni si ripetono

Elaborazione dei dati – le frequenze

I dati raccolti si possono riordinare in una tabella a doppia entrata e **individuare le frequenze, ovvero quante volte un certo dato si ripete**

FACCIAMO UN ESEMPIO

Organizziamo i dati raccolti circa gli sport praticati da 20 alunni di una classe

La somma delle frequenze
corrisponde al **TOTALE DEGLI
ALUNNI**

Sport	Frequenza
Calcio	10
Pallavolo	4
Basket	3
Judo	2
Nuoto	1
Totale	20

Frequenza assoluta, frequenza relativa, frequenza percentuale

- La **frequenza assoluta** (o semplicemente **frequenza**) di un dato statistico è il numero delle volte che il dato si presenta all'interno di un'indagine statistica: *nel nostro problema, quanti studenti praticano il calcio, quanti la pallavolo, ecc*
- La **frequenza relativa** di un dato statistico è il rapporto tra la frequenza assoluta di quel dato e il numero totale dei dati: *nel nostro problema per il calcio la frequenza relativa è $10/20 = 1/2$*
- Se la frequenza relativa è espressa in percentuale viene detta **frequenza percentuale**. Per il calcio la frequenza percentuale = $10/20 \times 100 = 50\%$

TABELLA DI FREQUENZA COMPLETA

Dato (sport)	Frequenza assoluta	Frequenza relativa	Frequenza percentuale
Calcio	10	10/20	$10/20 \times 100 = 50\%$
Pallavolo	4	4/20	$4/20 \times 100 = 20\%$
Basket	3	3/20	$3/20 \times 100 = 15\%$
Judo	2	2/20	$2/20 \times 100 = 10\%$
Nuoto	1	1/20	$1/20 \times 100 = 5\%$
TOTALE	20	20/20	$50\% + 20\% + 15\% + 10\% + 5\% = 100\%$

CLASSI DI FREQUENZA

In certe situazioni le frequenze assolute possono essere raggruppate in CLASSI DI FREQUENZA, intervalli che comprendono un massimo e un minimo

ESEMPIO

se indagine riguarda l'altezza di un gruppo di studenti, non ha molto senso considerare le frequenze delle singole altezze, ma può essere conveniente stabilire quanti studenti hanno altezza tra 150 cm e 160 cm; quanti tra 161cm e 170cm e così via...

L'elaborazione dei dati: gli indici statistici

Dopo aver raccolto i dati è necessario elaborarli per sintetizzare i dati stessi mediante **indici statistici** significativi: la **moda**, la **media** e la **mediana**.

La **moda** è il dato con la maggiore frequenza assoluta.

ESEMPIO

Nel caso analizzato in precedenza, la moda è il **calcio**, perché è lo sport praticato da più alunni, cioè il dato con la maggiore frequenza assoluta.

L'elaborazione dei dati: gli indici statistici

La **media aritmetica** di più dati statistici **quantitativi** è il quoziente fra la somma dei dati e il numero dei dati.

ESEMPIO

Se una squadra di calcio, nelle ultime 11 partite, ha segnato i seguenti gol: 0, 1, 3, 0, 1, 1, 4, 3, 2, 2, 5, allora la media aritmetica del numero di gol è:

$$\frac{0 + 1 + 3 + 0 + 1 + 1 + 4 + 3 + 2 + 2 + 5}{11} = \frac{22}{11} = 2$$

OSSERVAZIONE

È possibile calcolare la media aritmetica tenendo conto delle frequenze di ogni dato:

$$\frac{0 \cdot 2 + 1 \cdot 3 + 2 \cdot 2 + 3 \cdot 2 + 4 \cdot 1 + 5 \cdot 1}{2 + 3 + 2 + 2 + 1 + 1} = \frac{22}{11} = 2$$

L'elaborazione dei dati: gli indici statistici

La **mediana** è il dato che occupa la **posizione centrale in una sequenza di dati** disposti in ordine crescente (o decrescente).

ESEMPIO

Considerando i gol segnati dalla squadra di calcio nelle ultime **11 partite**:

0, 1, 3, 0, 1, 1, 4, 3, 2, 2, 5,

è necessario anzitutto disporre i dati in ordine crescente:

0, 0, 1, 1, 1, **2**, 2, 3, 3, 4, 5.

La posizione centrale è la sesta, quindi la mediana è 2.

L'elaborazione dei dati: gli indici statistici



ATTENZIONE



Se i **dati sono in numero pari**, dopo averli ordinati in modo crescente si devono considerare i **due dati centrali e calcolarne la media**.

ESEMPIO

Se si **considerano 12 partite** in cui sono stati segnati i seguenti gol, già scritti in ordine crescente:

0, 0, 1, 1, 1, **2, 3**, 3, 4, 4, 4, 5, la mediana è $\frac{2+3}{2} = 2,5$

La rappresentazione dei dati

I dati raccolti in un'indagine possono essere rappresentati IN UN GRAFICO

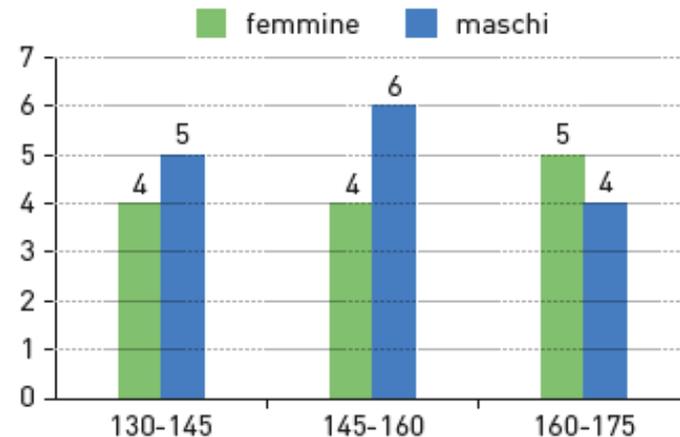
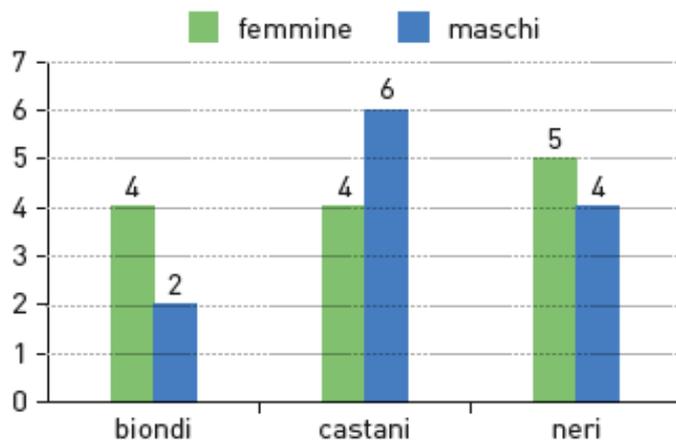
IL GRAFICO A BARRE

Il grafico a barre si costruisce generalmente utilizzando dei rettangoli affiancati.

L'altezza dei rettangoli indica la frequenza (assoluta o percentuale).

ESEMPIO

Nei due grafici a barre seguenti sono riportati il colore dei capelli nel primo caso e le altezze in cm raggruppate in classi di frequenza, nel secondo caso



La rappresentazione dei dati

L'AREOGRAMMA

Negli areogrammi, o grafici a torta, l'ampiezza dei settori è direttamente proporzionale alla frequenza percentuale.

ESEMPIO

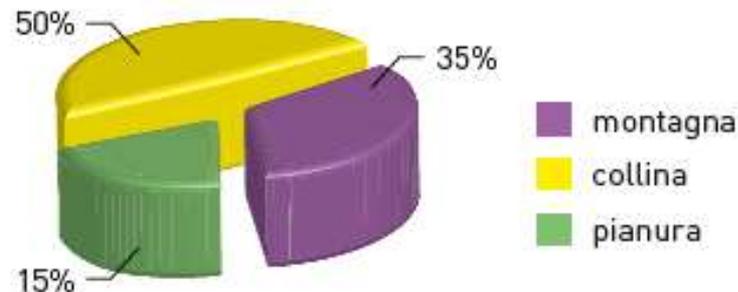
A partire dalla tabella in cui sono rappresentate le percentuali della suddivisione del territorio della Campania in zona montuosa, collinare e pianeggiante, si costruiscono le proporzioni per individuare l'ampiezza dei diversi settori:

Territorio della Campania	
Montagna	35%
Collina	50%
Pianura	15%
Totale	100%

$$35 : 100 = x : 360^\circ \text{ da cui si ricava } x = 35 \cdot 360^\circ : 100 = 126^\circ$$

$$50 : 100 = x : 360^\circ \text{ da cui si ricava } x = 50 \cdot 360^\circ : 100 = 180^\circ$$

$$15 : 100 = x : 360^\circ \text{ da cui si ricava } x = 15 \cdot 360^\circ : 100 = 54^\circ$$

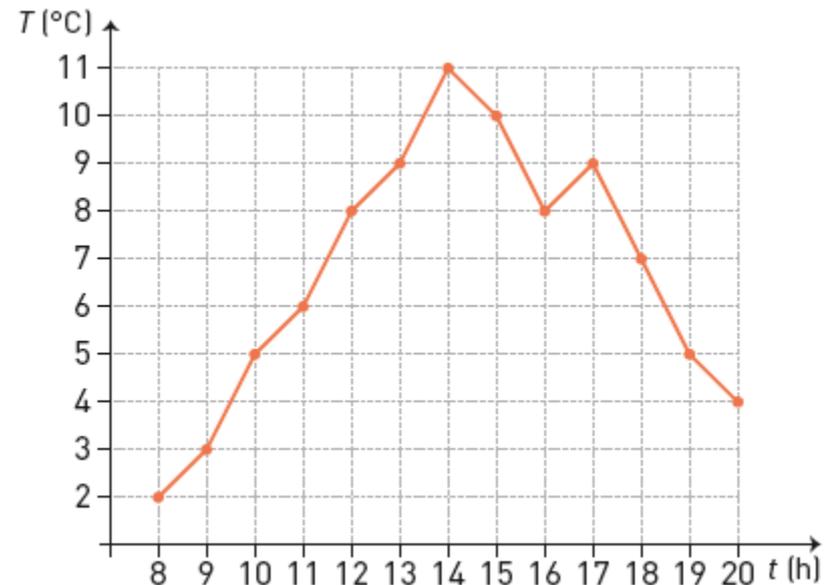


La rappresentazione dei dati

Un **diagramma cartesiano** è un grafico formato da punti nel piano cartesiano, generalmente uniti da una linea spezzata. I diagrammi cartesiani sono utilizzati **per rappresentare le variazioni di un fenomeno in un intervallo di tempo.**

ESEMPIO

Nel grafico a fianco sono riportate le temperature ambientali in un arco temporale di 12 ore.



La rappresentazione dei dati: sintesi

TABELLE

sono ordinati in

I DATI STATISTICI

possono essere visualizzati con

Regione	2005	2006	2007	2008
Piemonte	1,26	1,31	1,35	1,39
Valle d'Aosta	1,34	1,45	1,48	1,56
Lombardia	1,35	1,41	1,44	1,49
Trentino-Alto Adige	1,53	1,54	1,56	1,59
Veneto	1,35	1,39	1,42	1,45
Friuli Venezia Giulia	1,24	1,30	1,35	1,36
Liguria	1,18	1,21	1,25	1,31
Emilia-Romagna	1,33	1,38	1,42	1,45

Numero medio di figli per donna per regione (anni 2005 - 2008).

CARTE TEMATICHE

utilizzano

SIMBOLI, COLORI



IDEOGRAMMI

utilizzano

DISEGNI

Regione	Superficie boschiva (%)
Liguria	10%
Toscana	10%
Calabria	5%
Lazio	5%
Sicilia	1%

Legenda



10% della superficie boschiva



1% della superficie boschiva

ISTOGRAMMI

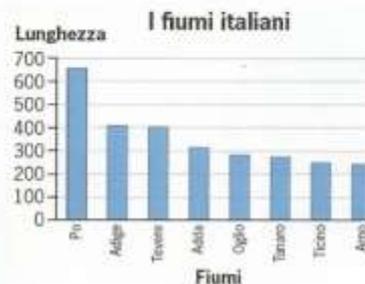
GRAFICI

costituiti da

RETTANGOLI

che rappresentano

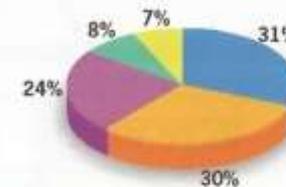
QUANTITÀ e MISURE



GRAFICI A TORTA

visualizzano

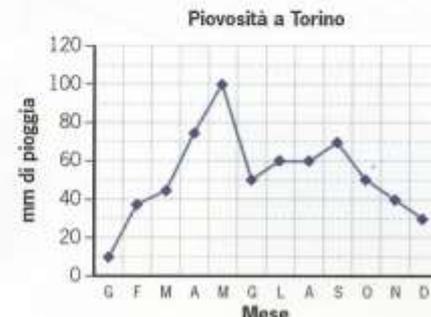
PERCENTUALI di un TOTALE



DIAGRAMMI CARTESIANI

rappresentano

ANDAMENTO di un FENOMENO NEL TEMPO



Gli elementi di un grafico

Un grafico deve sempre avere

- **Un titolo, che indica l'argomento:** esempio la temperatura di Roma a maggio 2015, oppure i libri letti in un anno da 20 studenti, ecc
- **Le unità di misura:** i giorni e i gradi, numero libri e frequenza, ecc
- **La fonte dei dati:** esempio ISTAT, OMS, ecc
L'anno di rilevazione dei dati

Letture ed interpretazione dei dati

Leggere ed interpretare dati statistici è **un'operazione complessa** attraverso la quale si fanno osservazioni, riflessioni, previsioni su un certo fenomeno, si mettono a confronto i dati rilevati con dati storici, si analizza una certa situazione e per fare ciò è spesso necessario avere conoscenze approfondite in vari ambiti.

Rielaborazione
Maria Vittoria Cicinelli
