

La rilevazione e la rappresentazione dei fenomeni statistici

Obiettivi dell'unità didattica:

- Prendere confidenza con alcuni termini tipici del linguaggio statistico
- Rappresentare le informazioni statistiche di base riassumendole all'interno di una tabella di frequenze
- Rappresentare attraverso un grafico le informazioni disponibili

Premessa e contenuti

Obiettivo dell'analisi statistica è quello di rilevare, sintetizzare e descrivere la realtà investigata. Nella vita lavorativa quotidiana tutte le persone sono a contatto con le più svariate fonti di informazione, molte delle quali di carattere quantitativo; la razionalizzazione e la sintesi di tali informazioni statistiche rappresentano alcuni tra gli strumenti più importanti al fine di mettere gli individui in condizione di prendere delle decisioni che non siano solamente il frutto delle sensazioni personali e/o dell'istinto, ma che abbiano delle solide basi oggettive.

Nel nostro percorso didattico cercheremo, dove non strettamente necessario al fine di comprendere le tematiche trattate, di non soffermarci eccessivamente sugli aspetti formali; viceversa, verrà dedicata maggiore attenzione all'utilizzo pratico delle misure introdotte, e all'interpretazione dei risultati ottenuti, in modo da ottenere una struttura più snella all'interno del percorso formativo che sia maggiormente in grado di avvicinare lo studente alle problematiche reali.

Nella presente unità didattica lo studente imparerà a "familiarizzare" con la terminologia statistica: verranno, infatti, preliminarmente enunciati quei termini di base che consentiranno di costruire un "linguaggio comune" statistico, necessario al fine di comprendere correttamente in che modo trattare le informazioni disponibili.

Quindi, verrà analizzato in che modo sia possibile passare da un insieme di dati di base, apparentemente confusi e non facilmente comprensibili per l'utente (come può capitare, ad esempio, con i risultati di un'indagine statistica), ad un corpus omogeneo e più chiaro di informazioni che, opportunamente sintetizzate, consentiranno all'utilizzatore di descrivere la realtà investigata e, eventualmente, di procedere ad ulteriori processi di sintesi.

Infine, il panorama descrittivo verrà completato attraverso il ricorso al trattamento grafico dell'informazione, il quale, vista la sua immediatezza e la semplicità interpretativa, costituisce spesso uno strumento esplicativo di grande impatto al fine di raggiungere convenientemente l'utente finale.

Nell'ambito delle lezioni impartite verranno introdotti i concetti e gli strumenti essenziali della statistica, descrittiva ed inferenziale, con particolare attenzione alla loro utilizzazione nelle scienze economiche e sociali. Lo scopo del corso è quello di fornire agli studenti tutti quegli strumenti di base utili a leggere, sintetizzare, analizzare ed interpretare i fenomeni osservati, con un'ottica di tipo quantitativo. Nella seconda parte del corso, inoltre, verranno introdotti le principali nozioni della statistica inferenziale (in particolare la stima dei parametri) in previsione di possibili approfondimenti ed applicazioni in campo sociale e aziendale (sondaggi, analisi di mercato).

Distribuzione di un carattere e sua rappresentazione

Statistica 1.

- **Statistica vs Matematica**

1- Matematica:

-È una scienza formale che studia le strutture, le quantità, le forme e le relazioni tra di esse attraverso l'uso di simboli e formule.

-Comprende vari rami, tra cui algebra, geometria, analisi e teoria dei numeri.

-La matematica è caratterizzata da un approccio rigoroso e deduttivo, basato su teoremi e dimostrazioni.

2- Statistica:

-È una disciplina che si occupa della raccolta, analisi, interpretazione e presentazione di dati.

-Utilizza strumenti matematici per effettuare analisi sui dati e fare inferenze sulle popolazioni a partire da campioni.

-La statistica è spesso utilizzata in una varietà di campi, come la scienza, l'economia, la sociologia e la medicina, per prendere decisioni basate su evidenze empiriche.

In sintesi, mentre la matematica si concentra su concetti astratti e teorie, la statistica applica strumenti matematici per affrontare problemi pratici legati ai dati. Entrambe le discipline sono interconnesse, poiché la statistica utilizza le tecniche matematiche, ma hanno obiettivi e metodi distinti.

- **Approccio diretto, semplice, poco formale**

- **Capire e non studiare (a memoria ...)**

- **Necessità di spiegare le “funzioni” della statistica (per capire bene cosa andremo a fare ...)**

- **Necessità di imparare un linguaggio comune (essenziale non solo per “capirci” ma anche, come vedremo, per utilizzare le opportune metodologie quantitative)**

La Statistica 2.

La statistica deve fornirci le necessarie metodologie quantitative che ci consentano di:

- **Raccogliere le informazioni**

- **Descrivere le informazioni**

- **Sintetizzare**

- **Fornire indicazioni (tendenze) quantitative (e, dunque, oggettive)**

Statistica Descrittiva e Inferenziale

Statistica Descrittiva

La statistica descrittiva si occupa di descrivere e sintetizzare (con tabelle, grafici, indici sintetici, ecc..) le informazioni (qualitative e quantitative) relative ad un certo gruppo di “soggetti” al fine di far emergere caratteristiche, andamento e le eventuali relazioni che si verificano all’interno del fenomeno analizzato.

Statistica Inferenziale

La statistica inferenziale utilizza le informazioni raccolte da un’indagine campionaria e le “manipola” (grazie all’utilizzo di opportuni “modelli”) in modo da poter generalizzare le indicazioni tratte dal gruppo investigato a tutta la popolazione dal quale tale gruppo è stato estratto.

Popolazione e Campione

Popolazione

Per popolazione si intende l’insieme di tutte le unità che compongono un collettivo. Sono esempi di popolazione: → Tutti i cittadini italiani → Tutti gli studenti dell’Università di Teramo → Tutte le imprese del settore dei servizi → Tutti gli occupati residenti nel comune di Giulianova → Tutte le imprese che esportano all’estero → Tutte le donne sposate con almeno un figlio

Campione

Una parte della popolazione (Deve avere “certe” caratteristiche ...)

Indagini censuarie e campionarie – 1

Indagine censuaria (o Totale) Effettuata su TUTTI i componenti della popolazione	Indagine campionaria Effettuata solo su ALCUNI componenti della popolazione
Pregi e difetti	
<ul style="list-style-type: none">- Esaustività- Ricchezza informazioni- Maggior dettaglio- Costi elevati- Lunghi tempi elaborazione- Difficoltà a raggiungere tutti- Qualità informazioni (?)	<ul style="list-style-type: none">- Economicità- Possibilità di ripetere le indagini continuamente nel tempo- Indagini mirate in base ad esigenze anche improvvise- Errore campionario- Minor livello di dettaglio

Le Distribuzioni Statistiche

Distribuzione Unitaria

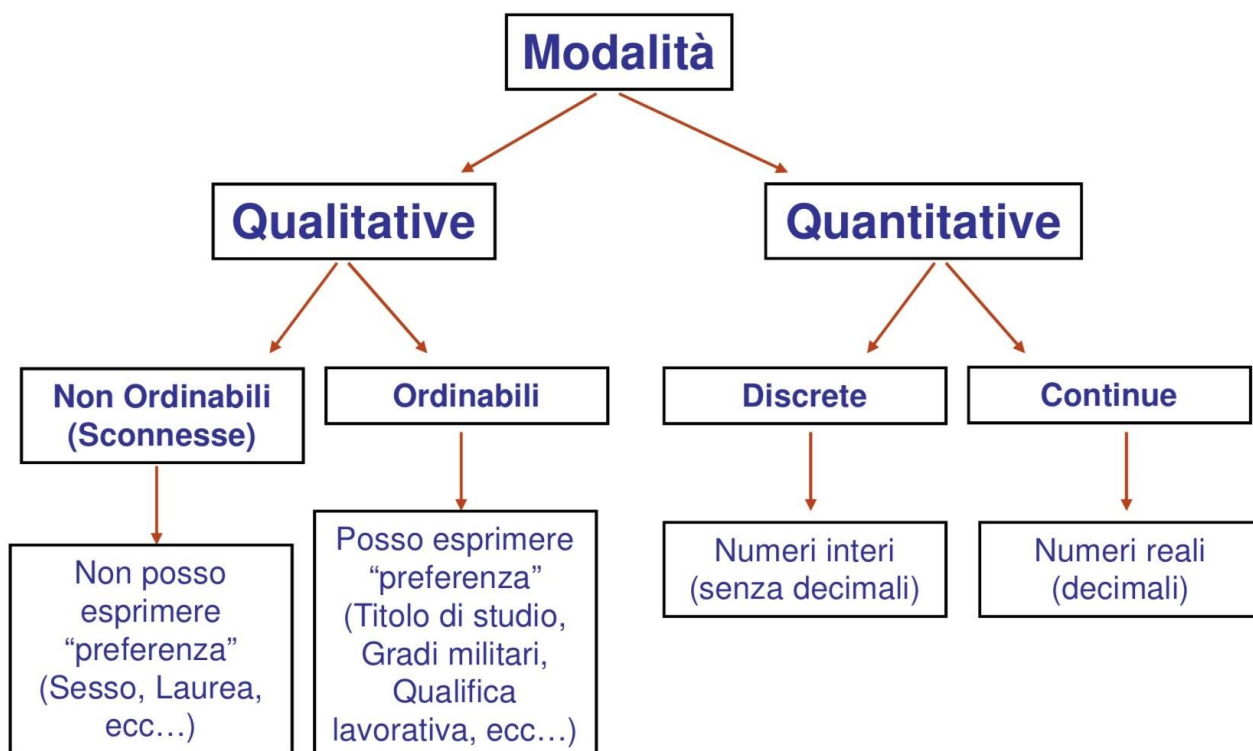
Unità	Sesso	Età	Laurea	Voto	Componenti	Reddito
1	M	24	Economia	109	2	25.100
2	F	25	Economia	110	5	26.500
3	F	24	Matematica	107	3	17.500
4	M	27	Statistica	103	4	35.600
5	F	29	Economia	95	3	76.200
6	F	24	Statistica	107	5	14.700
7	M	25	Matematica	105	4	19.350
8	M	27	Statistica	105	3	22.300
9	M	26	Economia	110	2	18.640
10	F	26	Statistica	110	4	45.020

Unità
statistiche

Caratteri

Modalità

Le modalità del carattere



Le frequenze

DISTRIBUZIONE UNITARIA SEMPLICE: è rappresentata dall'elencazione delle modalità osservate su ogni unità statistica in relazione ad un carattere considerato.

DISTRIBUZIONE UNITARIA MULTIPLA: è rappresentata dall'elencazione delle modalità osservate su ogni unità statistica in relazione ad ogni carattere considerato, quando l'elencazione si riferisce a più di un carattere.

DISTRIBUZIONE DI FREQUENZE: è una tabella nella quale vengono elencate tutte le modalità con le quali si presenta il carattere; ad ogni modalità viene associata la rispettiva frequenza assoluta.

FREQUENZE ASSOLUTE: rappresentano il numero di volte in cui una determinata modalità del carattere viene osservata all'interno del collettivo statistico che stiamo studiando.

X_i con questo simbolo viene indicata la j -esima modalità del carattere.

n_i con questo simbolo indichiamo le frequenze assolute associate alla j -esima modalità del carattere.

FREQUENZE RELATIVE (f_i) sono date dal rapporto tra le frequenze assolute e la numerosità totale: $f_i = n_i/n$.

FREQUENZE PERCENTUALI (p_i) sono date dal prodotto tra le frequenze relative e 100: $p_i = f_i \times 100$.

Esempio 1. **Sondaggio sulle preferenze musicali:**

- 20 persone amano il rock.
- 30 persone amano il pop.
- 10 persone amano il jazz.
- Totale: 60 persone.

2. **Distribuzione dei voti in un esame:**

- 15 studenti hanno preso A.
- 25 studenti hanno preso B.
- 5 studenti hanno preso C.
- 5 studenti hanno preso D.
- Totale: 50 studenti.

In ciascun esempio, le frequenze relative forniscono un'idea chiara di come i vari eventi (o categorie) si confrontano rispetto al totale delle osservazioni.

FREQUENZE CUMULATE (N_i) rappresentano la frequenza con cui si presentano modalità di ordine inferiore o uguale ad una certa modalità

Esempio 1. Altezze di un gruppo di persone

Immagina di avere i seguenti dati sulle altezze di 10 persone (in cm): 170, 165, 180, 175, 160, 185, 169, 172, 174, 178.

2. Considera i seguenti voti di un esame (su una scala da 0 a 30): 23, 28, 29, 20, 15, 28, 30, 25, 22, 27.

FREQUENZE RELATIVE CUMULATE (F_i) rappresentano la frequenza relativa con cui si presentano modalità di ordine inferiore o uguale ad una certa modalità

Esempio 1. Immagina di avere un insieme di dati sulla votazione di un esame:

Voto	Frequenza Assoluta
18	3
19	5
20	7
21	10
22	3

Per calcolare le frequenze relative assolute, possiamo sommare le frequenze e calcolare la percentuale di ciascuna categoria sul totale.

Le Distribuzioni Statistiche

Le distribuzioni statistiche descrivono il modo in cui uno o più caratteri si manifestano (distribuiscono) in un dato collettivo.

- distribuzioni semplici (un solo carattere)
- distribuzioni doppie (due caratteri)
- distribuzioni multiple (molti caratteri)

Si costruiscono associando ad ogni modalità il numero di volte in cui questa si è presentata

Esempio 1

Unità	Sesso
1	M
2	F
3	F
4	M
5	F
6	F
7	M
8	M
9	M
10	F

X_i	n_i
Maschi	170
Femmine	130
	300

Esempio 2

Unità	Laurea	Componenti
1	Economia	2
2	Economia	5
3	Matematica	3
4	Statistica	4
5	Economia	3
6	Statistica	1
7	Matematica	4
8	Statistica	3
9	Economia	2
10	Statistica	4

$$f_i = \frac{n_i}{n}$$

$$N_1 = n_1$$

$$N_2 = N_1 + n_2$$

$$N_3 = N_2 + n_3$$

X_i	n_i	f_i
Economia	120	0,400
Statistica	100	0,333
Matematica	80	0,267
	300	1,000

X_i	n_i	N_i
1	50	50
2	80	130
3	90	220
4	60	280
5	20	300
	300	