CAPITOLO 1

Caratteri, unità statistiche e collettivo

1 Caratteri, unità statistiche e collettivo

2 Classificazione dei caratteri statistici

3 Suddivisione in classi di un carattere quantitativo

4 I diversi tipi di rilevazione

5 Rilevazione totale e rilevazione campionaria

6 Il questionario

Che cosa è la Statistica

La Statistica è un insieme di tecniche che hanno come scopo la **conoscenza quantitativa dei fenomeni collettivi.**

Operazioni tipiche delle analisi statistiche



Conteggio



Classificazione



Misurazione



Sintesi

tramite modelli esplicativi dei fenomeni reali

Caratteri, unità statistiche e collettivo

Qualche definizione

Unità statistica

il singolo elemento su cui si effettua l'osservazione

Collettivo (popolazione)

l'insieme delle unità statistiche oggetto di studio

Carattere

l'aspetto osservato sulle unità (es. età, sesso, reddito)

Modalità

i diversi modi in cui un carattere si presenta

Un esempio di dati osservati

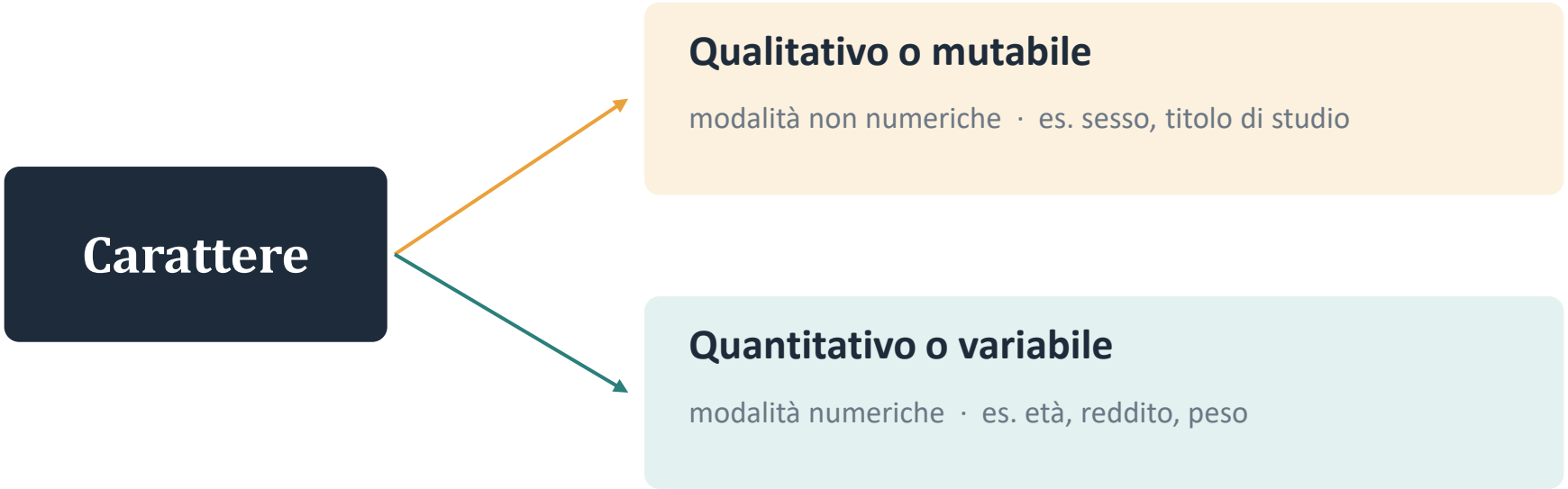
Nome	Età	Sesso	Titolo di studio	Attività	Peso (kg)	Punteggio esercizi
Rossi M.	32	M	laurea	occupato	72	65
Bianchi G.	39	F	laurea	occupato	55	55
Nicoletti C.	46	M	diploma	disoccupato	79	53
Marcelli F.	28	M	diploma	studente	63	78
Petrone A.	51	F	diploma	casalinga	64	21

Ogni riga è un'unità statistica

Ogni colonna è un carattere

Ogni cella è una modalità

Classificazione dei caratteri statistici



N.B. *Le modalità devono essere esaustive e non sovrapposte.*

Caratteri qualitativi (mutabili)

Sconnesso o nominale

Le modalità non hanno un ordine naturale.

- Religione
- Luogo di nascita

Ordinato

Es. grado di soddisfazione, posizione in graduatoria

Caratteri quantitativi (variabili)

Si distinguono secondo tre criteri

Tipo di scala

Scala di intervalli

es. temperatura

Scala di rapporti

es. reddito, altezza

Natura dei valori

Continui

es. peso, altezza

Discreti

es. n. di figli, età in anni

Trasferibilità

Trasferibili

es. reddito, n. di dipendenti

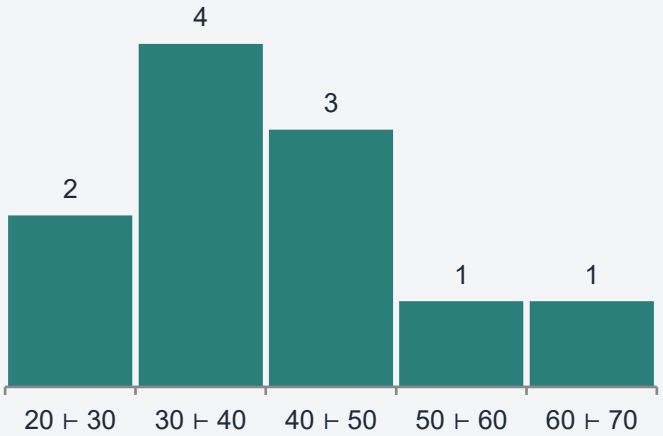
Non trasferibili

es. peso, età

Suddivisione in classi

1. Numero di classi adeguato al problema
2. Classi disgiunte
3. Includere tutte le modalità del carattere
4. Quale ampiezza delle classi?
5. Intervalli aperti a destra o a sinistra

Esempio: età in anni



⊢ : estremo sinistro incluso, destro escluso

Indagine statistica

Indagine totale | Indagine campionaria

Le fasi dell'indagine

1 Definizione degli obiettivi

unità e variabili da rilevare; periodo di riferimento

2 Popolazione e lista

individuazione della popolazione e della lista delle unità statistiche

3 Piano di campionamento

definizione del disegno campionario

4 Raccolta dei dati

tecnica di rilevazione; questionario e pretest; rilevazione sul campo

5 Registrazione dei dati

registrazione su supporto informatico; controllo e correzione

6 Elaborazione e analisi

dei dati raccolti

La rilevazione dei dati: completa o parziale

Completa

Si esaminano tutte le unità statistiche della popolazione oggetto di studio.

Pregi

- Accuratezza delle stime, anche a livelli territoriali molto dettagliati
- Ricchezza delle informazioni raccolte
- Esaustività

Difetti

- Costo elevato
- Tempi di elaborazione molto lunghi
- Qualità dei dati non elevata

Parziale (campionaria)

Si studia un sottoinsieme, detto campione, dell'insieme di riferimento.

Pregi

- Continuità della rilevazione
- Economicità
- Indagini più mirate e approfondite

Difetti

- Dettaglio territoriale limitato
- Variabilità campionaria

Un esempio

Si vogliono studiare le caratteristiche demografiche e sociali (***caratteri***) della popolazione del Lazio al 1° gennaio 1998 (***collettivo***).

Un semplice questionario potrebbe raccogliere le informazioni riportate a destra.

Terza intervista – dati estratti dal questionario

Codice identificativo intervista	**3**
Sesso	**M**
Età in anni compiuti	**38**
Stato civile	**Coniugato**
Provincia di residenza	**LT**
Titolo di studio	**Laurea**
Professione	**Impiegato**
N. componenti della famiglia (oltre l'intervistato)	**4**
Reddito lordo annuo (euro)	**40.000**

Dal questionario alla base dati

Le 20 interviste effettuate con lo stesso questionario formano una base dati (matrice dei dati).

Codice	Carattere 1	Carattere 2	Carattere 3	…
1				
2				
3				
4				
…				

La prima colonna

contiene il codice che identifica ciascuna intervista (unità statistica).

Le colonne successive

riportano le informazioni rilevate col questionario: le variabili o caratteri.

Tabella dei dati

Cod.	Sesso	Età	Stato civile	Prov.	Titolo di studio	Professione	N. comp. fam.	Reddito lordo annuo (€)
1	F	29	Nubile	RM	Laurea	Impiegato	1	35.000
2	F	40	Nubile	FR	Diploma	Disoccupato	3	10.000
3	**M**	**38**	**Coniugato**	**LT**	**Laurea**	**Impiegato**	**4**	**40.000**
4	F	42	Coniugata	FR	Laurea	Libero professionista	3	90.000
5	M	48	Celibe	RI	Laurea	Libero professionista	1	90.000
6	M	59	Coniugata	RM	Diploma	Casalinga	4	15.000
7	M	70	Celibe	RM	Lic. media	Pensionato	5	30.000
8	F	25	Nubile	FR	Laurea	Disoccupato	3	20.000
9	F	35	Coniugata	FR	Laurea	Libero professionista	3	50.000
10	M	46	Celibe	RM	Laurea	Pensionato	2	70.000

Prime 10 interviste. In evidenza la terza intervista, vista nella slide precedente.

Qualche osservazione

La stessa modalità può presentarsi ripetuta, anche molte volte, nella stessa colonna.

Ogni colonna corrisponde a quella che più avanti definiremo distribuzione unitaria di un carattere.

Scelto un carattere, dalla tabella si individua subito la modalità assunta da ogni unità statistica del collettivo.

Viceversa, scelta un'unità, si leggono le modalità che essa presenta per ogni carattere considerato.