



# Capitolo 1

## La rilevazione dei fenomeni statistici

# Capitolo 1

- Caratteri, unità statistiche e collettivo
- Classificazione dei caratteri statistici
- Suddivisione in classi di un carattere quantitativo
- I diversi tipi di rilevazione
- Rilevazione totale e rilevazione campionaria
- Il questionario

# Che cosa è la Statistica

- ◆ La **Statistica** può essere definita come un insieme di tecniche che hanno come scopo la conoscenza quantitativa dei **fenomeni collettivi**
- ◆ Operazioni tipiche delle analisi statistiche sono:
  - ◆ il **conteggio**
  - ◆ la **classificazione**
  - ◆ la **misurazione**
  - ◆ la **sintesi** tramite modelli esplicativi dei fenomeni reali

# Caratteri, unità statistiche e collettivo

Qualche definizione:

- ◆ unità statistica
- ◆ collettivo o popolazione
- ◆ carattere
- ◆ modalità del carattere

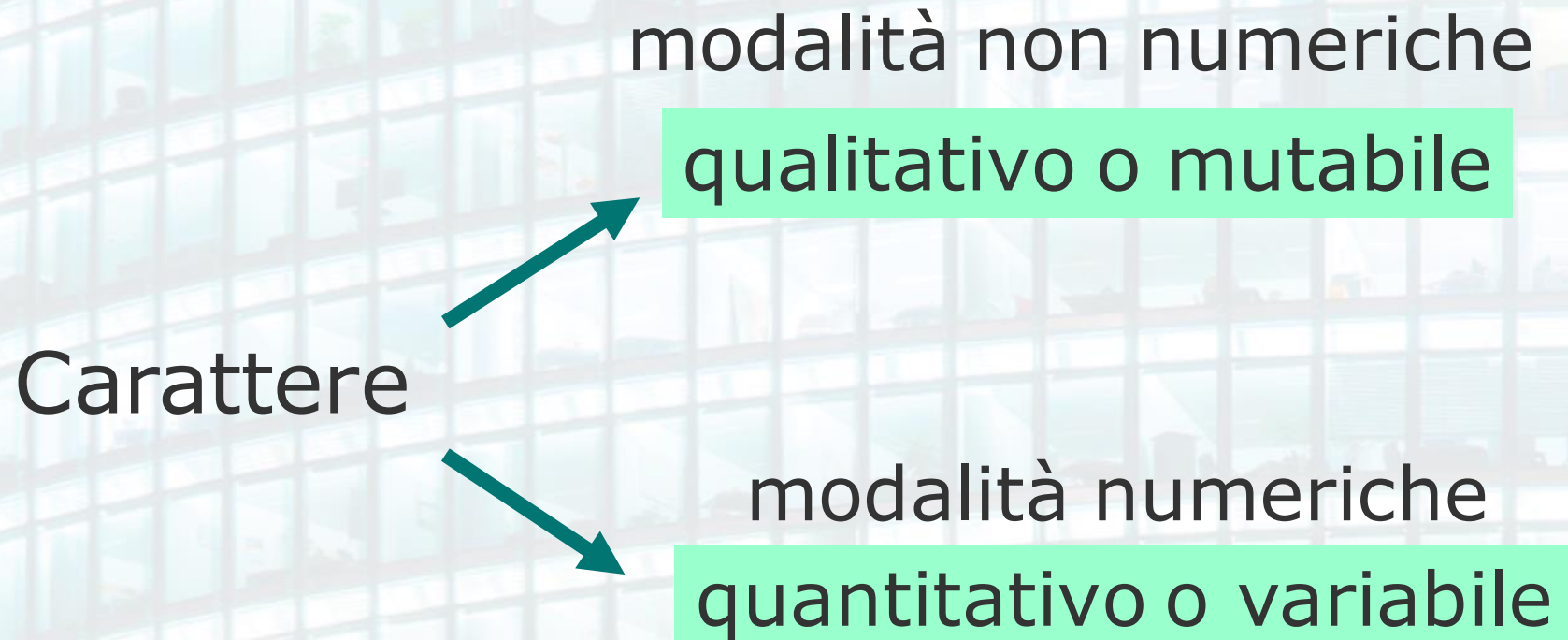
Vari tipi di collettivo:

- di stato o di movimento
- empirico o teorico
- finito o infinito

# 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

# Classificazione dei caratteri statistici



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

# Caratteri qualitativi (mutabili)

## ◆ sconnesso o nominale

(religione, luogo di nascita,...)

## ◆ ordinato

(grado di soddisfazione, posizione in graduatoria...)

### ◆ rettilineo

◆ ciclico (direzione del vento, mesi dell'anno...)

# Caratteri quantitativi (variabili)

Si distinguono in:

- **su scala di intervalli** (es. temperatura...)
- **su scala di rapporti** (es. reddito, altezza...)
  
- **continui** (es. peso, altezza...)
- **discreti** (es. n. di figli, età in anni...)
  
- **trasferibili** (es. reddito, N. di dipendenti...)
- **non trasferibili** (es. peso, età...)



# Suddivisione in classi

- n. di classi adeguato al problema
- classi disgiunte
- includere tutte le modalità del carattere
- quale ampiezza delle classi?
- intervalli aperti a destra o a sinistra

# Rilevazioni: Sperimentali/osservazionali

## Sperimentali

(Medicina, Fisica, Chimica)

1. Ipotesi di lavoro
2. Possibilità di controllo



Controllo

- **Diretto** dei Fattori Sperimentali e dei Fattori di Stratificazione
- **Indiretto:** Randomizzazione

## Osservazionali

(Indagini di mercato, sondaggi)

Non si ha la possibilità di controllo



Indagine Statistica

# Indagine statistica

- ◆ Indagine Totale
- ◆ Indagine Campionaria

## Fasi dell'indagine

- ◆ **definizione degli obiettivi**
  - ◆ definizione delle unità e delle variabili da rilevare
  - ◆ scelta del periodo di riferimento
- ◆ **individuazione della popolazione e della lista delle unità statistiche**
- ◆ **definizione del piano di campionamento**
- ◆ **raccolta dei dati**
  - ◆ scelta della tecnica di rilevazione
  - ◆ formulazione del questionario e pretest
  - ◆ rilevazione sul campo
- ◆ **registrazione dei dati**
  - ◆ registrazione su supporto magnetico
  - ◆ controllo e correzione
- ◆ **elaborazione e analisi dei dati**

# La rilevazione dei dati

La raccolta delle informazioni può essere **completa** oppure **parziale**.

È **completa** quando si esaminano tutte le unità statistiche che compongono la popolazione oggetto di studio.

## **Pregi:**

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

## **Difetti:**

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

È **parziale** quando ci si limita a studiare un sottoinsieme, detto "**campione**" dell'insieme di riferimento.

## **Pregi:**

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

## **Difetti:**

- Riferimento territoriale non spinto
- Variabilità campionaria

# Statistica descrittiva e inferenza

La **statistica descrittiva** fornisce gli strumenti per sintetizzare ed esplicitare in forma corretta il modo in cui il fenomeno si è manifestato nel collettivo osservato

Mediante l'**inferenza** statistica è possibile misurare e controllare l'attendibilità delle informazioni provenienti da un campione



estrazione  
del  
campione

# Estrazione del campione

- **Campionamento casuale:** insieme di tutte quelle tecniche di formazione del campione in cui la selezione delle unità è affidata a regole probabilistiche.
- **Campionamento casuale semplice:** i campioni della stessa dimensione estraibili da una popolazione hanno uguale probabilità di essere estratti.
- **Campionamento casuale stratificato:** la popolazione viene suddivisa in un certo numero di strati. Da ogni strato in maniera indipendente viene poi estratto un campione casuale semplice.

# Vantaggi campionamento stratificato

- **Miglioramento della stima**, se gli strati sono stati ben scelti.
- Possibilità di ottenere anche la **stima per le singole sottopopolazioni** o strati.

# Progettazione del questionario

- **Concettualizzazione:** si individuano le entità che entrano in gioco, si descrivono le relazioni esistenti fra le entità e si individuano le possibili gerarchie fra le relazioni;
- **Individuazioni aree e sottoaree**
- **Contenuti, formulazione e successione logica delle domande**  
(*ES. ad imbutto o ad imbutto capovolto*)



# Tipi di Intervista

◆ Diretta → Faccia a faccia

◆ Indiretta → Telefono, Posta, exit poll, CATI

- ◆ NON Strutturata
- ◆ SEMI Strutturata
- ◆ Strutturata

# Tipi di Domande

- ◆ *Dirette o Indirette*
- ◆ *A risposta semplice / multipla*
- ◆ *Aperta o chiusa*
- ◆ *Domande Filtro*
- ◆ *Domanda gerarchizzata*

# 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 seguenti informazioni →

## Dati estratti dal questionario della terza intervista

Codice ident. intervista	<b>3</b>
Sesso	<b>M</b>
Età in anni compiuti	<b>38</b>
Stato civile	<b>Coniugato</b>
Provincia di residenza	<b>LT</b>
Titolo di studio	<b>Laurea</b>
Professione	<b>Impiegato</b>
Numero componenti della famiglia (oltre l'intervistato)	<b>4</b>
Reddito lordo annuo (in migliaia di euro)	<b>40</b>

# Un esempio

Di seguito vediamo come si presenta la base dati relativa ad un insieme di 20 interviste effettuate con lo stesso questionario.

Mentre la prima colonna contiene un codice che permette di identificare le singole interviste, le colonne successive si riferiscono alle diverse informazioni rilevate col questionario (*variabili o caratteri*)

# Tabella dei dati

Cod. Int.	Sesso	Età in anni compiuti	Stato civile	Provincia di resid.	Titolo di studio	Professione	n. comp. Fam.	Reddito lordo annuo (migl. Euro)
1	F	29	Nubile	RM	Laurea	Impiegato	1	35000
2	F	40	Nubile	FR	Diploma	Disoccupato	3	10000
3	M	38	Coniugato	LT	Laurea	Impiegato	4	40000
4	F	42	Coniugata	FR	Laurea	Libero professionista	3	90000
5	M	48	Celibe	RI	Laurea	Libero professionista	1	90000
6	M	59	Coniugata	RM	Diploma	Casalinga	4	15000
7	M	70	Celibe	RM	Lic. Media	Pensionato	5	30000
8	F	25	Nubile	FR	Laurea	Disoccupato	3	20000
9	F	35	Coniugata	FR	Laurea	Libero professionista	3	50000
10	M	46	Celibe	RM	Laurea	Pensionato	2	70000

# Qualche osservazione

- La stessa modalità può presentarsi ripetuta – anche molte volte – nella stessa colonna
- Ogni colonna corrisponde a quella che più avanti definiremo come *distribuzione unitaria* di un carattere
- Scelto un carattere tra quelli di interesse, dalla tabella è possibile individuare subito la modalità assunta da *ogni* unità statistica del collettivo.
- Viceversa, scelta un'unità, è possibile leggere le modalità che questa unità presenta in corrispondenza di *ogni* carattere considerato.