

# L'indagine Statistica

Massimo Aria

Dipartimento di Matematica e Statistica  
Università di Napoli Federico II

## Contenuti

- Le attività di un'indagine sul campo
- L'indagine statistica
- Le tecniche di rilevazione dei dati
- I concetti di base
  - Unità statistica, popolazione e campione
  - Carattere o variabile
  - Modalità o attributi
  - Le scale di misura dei caratteri
- L'organizzazione dei dati

## Il piano delle attività

- Coordinamento scientifico
- Coordinamento operativo
- Analisi dello scenario
- Analisi desk (*protocollo di intervista, interviste a testimoni privilegiati*)
- Analisi del target (*selezione banche dati, progettazione del questionario per l'estensiva*)
- Campionamento (*liste, lotti di lancio, errore*)
- Indagine estensiva (*rilevazione, elaborazione statistica*)

L'indagine sul campo

3

## L'indagine statistica

Insieme delle operazioni effettuate per:

- Descrivere una o più caratteristiche del fenomeno oggetto di studio
- Verificare ipotesi sul fenomeno
- Esplorare relazioni tra i diversi aspetti del fenomeno

L'indagine sul campo

4

## Fasi di un'indagine statistica

1. Definizione degli obiettivi della ricerca
2. Rilevazione dei dati
3. Elaborazione metodologica
4. Presentazione ed interpretazione dei risultati
5. Utilizzazione dei risultati della ricerca

### Definizione degli obiettivi della ricerca

- *inquadrare il problema definendo sia l'obiettivo globale che gli obiettivi parziali in maniera compatibile con i vincoli di tempo e di costo*
- *individuare le informazioni da ricercare*
- *circoscrivere tempo e spazio dell'indagine*

*attraverso* → **studi preliminari volti a delineare lo scenario di riferimento**

## Rilevazione dei Dati

*consiste nella raccolta delle informazioni (precedentemente individuate) necessarie al raggiungimento degli obiettivi preposti*

Tipi di rilevazione:

- **completa:** quando si esaminano tutti gli elementi oggetto di studio (*la rilevazione riguarda l'intera popolazione*);
- **parziale:** quando lo studio considera un sottoinsieme, detto *campione*, dell'insieme di riferimento (*la rilevazione riguarda una parte della popolazione*).

L'indagine sul campo

7

## ... rilevazione

parziale vs totale

- *Campione o popolazione?*
- *Analisi di tutto il fenomeno o parte di esso?*

dichiarazioni vs misurazioni

- *Sono i dati frutto di misurazioni?*
- *Provengono da un questionario?*

**Analisi della qualità dell'informazione**

Roberta Siciliano

L'indagine sul campo

8

## ... piano di raccolta dei dati

- Predisposizione del materiale di rilevazione
- Strumento e tecnica di rilevazione (questionario, indagine telefonica, rilevazione assistita da computer, ecc.)
- Addestramento del personale
- Indagine pilota (pre-testing: accertamento della qualità del materiale di rilevazione)
- Eventuale ridefinizione degli obiettivi
- Preparazione del materiale di rilevazione definitivo

## Metodi di rilevazione

Tra i metodi di rilevazione dei dati molto impiegato è quello **dell'indagine mediante questionario** da sottoporre ad una data collettività o ad un campione di essa.

Il **questionario** è un modello di rilevazione costituito da un certo numero di domande espresse in modo *standardizzato* (uguale per tutti gli interpellati) e *strutturato* (le cui domande riflettono le ipotesi e l'oggetto della ricerca).

## Il questionario può essere compilato:

- ❖ Direttamente dalla persona da interpellare (*autocompilazione*) alla quale viene consegnato (o spedito nel caso di indagine postale); il questionario sarà ritirato o rispedito una volta compilato
- ❖ Tramite la *tecnica dell'intervista*: in tal caso è l'intervistatore a porre le domande del questionario all'intervistato;

## Il questionario

Sezioni

Domande di controllo

*Briefing* degli intervistatori

Primo controllo su una parte dei questionari rientrati

Memorizzazione delle informazioni

- *Al fine di migliorare la leggibilità e la gestione delle informazioni!*

- *Vanno inserite in più punti per verificare l'attendibilità delle risposte!*

- *E' una fase delicatissima e cruciale! Può pregiudicare l'analisi!*

- *Di solito viene scelto un 5% dei questionari.*

- *Orientarsi ad un Data Base!*

## ***La tecnica dell'intervista***

### ***può aver luogo con l'ausilio dell'elaboratore elettronico***

- Sistema *C.A.T.I.* (*Computer aided telephone Interview*): l'intervista viene realizzata telefonicamente e contemporaneamente avviene l'inserimento delle risposte sull'elaboratore;
- Sistema *C.A.W.I.* (*Computer aided web Interview*): si realizzano ricerche su internet somministrando questionari e raccogliendo dati on – line. Il sistema permette di controllare il campione o la popolazione di riferimento e di gestire le quote per ottenere dati rappresentativi e corretti della popolazione di internet.

## **I vantaggi del C.A.T.I. e del C.A.W.I**

- Controllo dell'errore
- Controllo di coerenza (attraverso controlli logici)
- Rapidità
- Controllo del campione da intervistare
- Codifica automatica dei dati
- Disponibilità *real – time* dei dati

# L'informazione statistica

*consiste in ogni risultato ottenuto da una indagine sulle unità statistiche in rapporto ai diversi fenomeni (caratteri) analizzati*

## Tipologia delle informazioni statistiche:

- **Variabile:** è un carattere *quantitativo* che assume come modalità dei numeri reali;
- **Mutabile:** è un carattere *qualitativo* che assume attributi non numerici di diversa natura.

L'indagine sul campo

15

## Esempi di caratteri statistici

- **Variabili:** →

| età | peso | altezza |
|-----|------|---------|
| 19  | 86   | 180     |
| 20  | 78   | 174     |
| 21  | 84   | 192     |
| 22  | 63   | 158     |
| 34  | 71   | 169     |
| ... | ...  | ...     |
| ... | ...  | ...     |
| ... | ...  | ...     |

Modalità

- **Mutabili:** →

| Colore Capelli | Professione | Grado istruzione |
|----------------|-------------|------------------|
| Rosso          | Operaio     | elementare       |
| Bruno          | Impiegato   | laurea           |
| Biondo         | Dirigente   | diploma          |
| Bruno          | Disoccupato | media            |
| Rosso          | Operaio     | laurea           |
| ...            | ...         | ...              |
| ...            | ...         | ...              |
| ...            | ...         | ...              |

L'indagine sul campo

16

## Tipi di caratteri

*Sulla base della scala di misurazione, si distinguono i caratteri:*

- **con scala nominale:** sono **mutabili** le cui modalità non assumono alcun ordine precostituito;  
(esempio: genere, residenza, occupazione, ecc.)
- **con scala ordinale:** sono **mutabili** che, pur utilizzando una scala qualitativa, assumono modalità logicamente ordinabili;  
(esempio: titolo di studio, grado di preferenza, ecc.)
- **con scala ad intervallo:** sono **variabili**, la cui scala di misurazione fa riferimento ad un'origine arbitraria;  
(esempio: punteggio, temperatura, ecc.)
- **con scala di rapporto:** sono **variabili** per le quali è intrinseca la definizione di uno zero assoluto.  
(esempio: durata, arrivi, presenze, ecc.)

## Esempio di scala nominale

- Stato civile
- Settore di attività economica di un'impresa
- Titolo di studio
- Canale televisivo

## Esempio di scala ordinale

- **Giudizio** (scarso, mediocre, sufficiente, buono, ottimo)
- **Reddito in fasce**
- **Titolo di studio** (ordinato per anni di studio)

## Esempio di scala di rapporto

- **Reddito pro-capite**
- **Numero di figli per famiglia**
- **Numero di addetti per impresa**
- **Età**

## Proprietà dei caratteri

Sulla base della scala di misurazione, si distinguono le seguenti proprietà:

- **mutabili con scala nominale:**
  - *simmetria* ossia se  $x_i = x_j$ , anche  $x_j = x_i$
  - *transitività* ossia se  $x_i = x_j$ ;  $x_j = x_k$ ; allora  $x_i = x_k$
- **mutabili con scala ordinale:**
  - *simmetria, transitività*
  - *ordine*  $x_i < x_j$
- **variabili con scala ad intervallo e di rapporto:**
  - *simmetria, transitività, ordine*
  - *differenza*  $x_i - x_j$

L'indagine sul campo

21

## ...ancora sulle variabili

- Le operazioni possibili

- Conteggio
- Organizzazione
- Classificazione
- Differenza
- Differenza relativa

$X =$  variabile  
 $x_1 =$  modalità 1 di  $X$   
 $x_2 =$  modalità 2 di  $X$   
 $x_1 - x_2 = \text{diff}$ .

$\frac{x_2 - x_1}{x_1} =$  tasso di variaz.

L'indagine sul campo

22

## Tipi di variabili

sulla base dei valori assunti dalle modalità, le variabili sono:

- **continue:**  
possono assumere un qualunque valore contenuto in un intervallo reale predefinito;  
(esempio: età, peso, altezza, temperatura, la temperatura massima giornaliera in una località balneare, ecc.)
- **discrete:**  
possono invece assumere un numero discreto di modalità che possono essere sempre poste in corrispondenza con l'insieme dei numeri naturali.  
(esempio: numero figli, voto di laurea, numero di visitatori al museo, ecc.)

## ... organizzazione dei dati

- Rilevazione dei dati e loro ricodifica
- Controllo della qualità dei dati
- Categorizzazioni, dicotomizzazioni, etc.
- Trattamento dei dati mancanti
  - esclusione
  - imputazione

## Variabile dicotomica

È una variabile con due sole  
modalità

- **E' l'espressione del possesso o meno di un determinato attributo oppure l'appartenenza a una categoria di unità**

L'indagine sul campo

25

## Es. di variabile dicotomica

- Genere
- Avere superato un esame
- Appartenere ad un corso
- Reddito (basso/alto)

L'indagine sul campo

26

## Organizzazione dei dati

- Sintesi e sua importanza
- Rappresentazione dei dati statistici
  - Enumerativa
  - Tabellare
  - Grafica
- Distribuzioni di frequenza
  - Semplici
  - Congiunte
- Serie storiche
- Serie territoriali
- Matrici di dati

L'indagine sul campo

27

## Elaborazione metodologica

*consiste nella scelta e nell'applicazione della  
metodologia statistica più adeguata  
(in coerenza con la tipologia dei dati rilevati  
e gli obiettivi preposti)*

nella scelta della metodologia giocano un ruolo importante:

- la distinzione tra fenomeni *qualitativi* e *quantitativi*
- la natura della rilevazione: *completa* o *parziale*
- l'obiettivo di fondo dell'analisi: *descrittiva* o *inferenziale*

L'indagine sul campo

28

## ... elaborazione

Identificazione metodi e tecniche

- *Semplice descrizione dei dati?*
- *Modellistica?*
- *Previsione?*

Scelta del software

- *Per memorizzare i dati*
- *Per elaborare i dati*
- *Per trasferire i risultati*

Validazione risultati

Feedback

L'indagine sul campo

29

## Presentazione ed interpretazione dei risultati

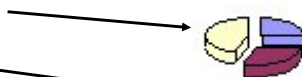
*attraverso una nota metodologica e l'uso di tecniche visuali, i risultati dell'indagine diventano usufruibili per i soggetti destinatari della stessa*

(come manager, decisori economici, esperti finanziari, ecc.)

alcuni esempi di strumenti visuali sono:

- tabelle e rapporti sintetici
- grafici a torta
- istogrammi
- serie temporali
- ecc.

| Car Name          | No. of Models | Total Sales (Millions) |      |      |      | Average Price |      |
|-------------------|---------------|------------------------|------|------|------|---------------|------|
|                   |               | 1996                   | 1997 | 1998 | 1999 |               |      |
| Chrysler PT       | 8             | 207                    | 430  | 1044 | 14   | 70            | 18   |
| Chrysler Stratus  | 4             | 136                    | 51   | 1013 | 15.5 | 76            | 22   |
| Ford LTD          | 4             | 50                     | 72   | 2106 | 11.5 | 74            | 18   |
| Chrysler Voyager  | 4             | 200                    | 110  | 2000 | 10   | 77            | 17   |
| Vauxhall Viva     | 4             | 65                     | 62   | 1042 | 12.9 | 81            | 20.7 |
| Mercedes Sprinter | 4             | 107                    | 78   | 1207 | 14.9 | 88            | 20   |
| Ford Escort       | 4             | 120                    | 70   | 1027 | 10.6 | 82            | 20   |



L'indagine sul campo

30

## ... presentazione

Selezione dei risultati da presentare

**NON TUTTI**  
*i risultati derivanti dalle*  
*elaborazioni*  
**SONO**  
**INTERESSANTI!**

Scelta del mezzo di diffusione

*A volte un foglio di carta*  
*può essere più efficace*  
*di una relazione di 300*  
*pagine!!!!*

trasportabilità

*usare piattaforme*  
*standard*

L'indagine sul campo

31

## Utilizzazione dei risultati della ricerca

*lo statistico contribuisce ad una corretta*  
*utilizzazione dei risultati dell'indagine scientifica,*  
*circoscrivendo l'ambito interpretativo e delineando*  
*i vincoli entro cui essa assume validità*

In tal senso è importante considerare:

- Le ipotesi su cui la metodologia impiegata si fonda;
- L'ambito spazio-temporale in cui la ricerca è condotta;
- Nell'ottica previsiva: la validità, nel futuro, delle variabili esplicative considerate.

L'indagine sul campo

32

## ... utilizzo dei risultati

### **Definire chiaramente le caratteristiche della ricerca**

*Specificare esattamente il problema analizzato e le informazioni di partenza*

### **Citare ricerche pregresse**

*Analizzare e riportare i risultati di ricerche fatte da altri ricercatori sullo stesso argomento*

### **Indicare chiaramente le condizioni di applicabilità**

*Chi utilizza i risultati deve sapere bene i pro e contro!!!!*