

Corso di Statistica Facoltà di Economia

francesco mola

a.a. 2010-2011

Lezione n° 1

Sommario

- Che cosa è la *statistica*?
Lo scopriremo a fine corso ... altrimenti che gusto c'è?
- L'indagine
- I concetti di base
 - Unità statistica, popolazione e campione
 - Carattere o variabile
 - Modalità e attributi
 - Frequenze, serie e seriazioni
- Le variabili
- Le scale di misura
- Le fonti statistiche
- Strutturazione generale del corso

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

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

Fasi di un'indagine statistica

1. Definizione obiettivi

- Obiettivo globale
- Obiettivi parziali
- Vincoli
 - Tempo
 - Costi

Fasi di un'indagine statistica (cont.)

Rilevazione dati

parziale vs totale

dichiarazioni vs misurazioni

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

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

Analisi della qualità dell'informazione

Popolazione di 25 marche di sigarette

MARCA	condensato	peso	monossido	prezzo
Alpine	14.1	0.86	13.6	medio
Benson&Hedges	16	1.06	16.6	alto
BullDurham	29.8	2.03	23.5	alto
CamellLights	8	0.67	10.2	medio
Carlton	4.1	0.4	5.4	basso
Chesterfield	15	1.04	15	alto
GoldenLights	8.8	0.76	9	medio
Kent	12.4	0.95	12.3	medio
Kool	16.6	1.12	16.3	alto
L&M	14.9	1.02	15.4	medio
LarkLights	13.7	1.01	13	medio
Marlboro	15.1	0.9	14.4	alto
Merit	7.8	0.57	10	medio
MultiFilter	11.4	0.78	10.2	medio
NewportLights	9	0.74	9.5	medio
Now	1	0.13	1.5	basso
OldGold	17	1.26	18.5	alto
PallMallLight	12.8	1.08	12.6	medio
Raleigh	15.8	0.96	17.5	alto
SalemUltra	4.5	0.42	4.9	basso
Tareyton	14.5	1.01	15.9	medio
True	7.3	0.61	8.5	medio
ViceroyRichLight	8.6	0.69	10.6	medio
VirginiaSlims	15.2	1.02	13.9	alto
WinstonLights	12	0.82	14.9	medio

Campione di 10 marche di sigarette (primi 10)

MARCA	condensato	peso	monossido	prezzo
Alpine	14.1	0.86	13.6	medio
Benson&Hedges	16	1.06	16.6	alto
BullDurham	29.8	2.03	23.5	alto
CamellLights	8	0.67	10.2	medio
Carlton	4.1	0.4	5.4	basso
Chesterfield	15	1.04	15	alto
GoldenLights	8.8	0.76	9	medio
Kent	12.4	0.95	12.3	medio
Kool	16.6	1.12	16.3	alto
L&M	14.9	1.02	15.4	medio

Campione di 10 marche di sigarette (ultimi 10)

MARCA	condensato	peso	monossido	prezzo
Now	1	0.13	1.5	basso
OldGold	17	1.26	18.5	alto
PallMallLight	12.8	1.08	12.6	medio
Raleigh	15.8	0.96	17.5	alto
SalemUltra	4.5	0.42	4.9	basso
Tareyton	14.5	1.01	15.9	medio
True	7.3	0.61	8.5	medio
ViceroyRichLight	8.6	0.69	10.6	medio
VirginiaSlims	15.2	1.02	13.9	alto
WinstonLights	12	0.82	14.9	medio

Campione di 10 marche di sigarette (a caso)

marche	condensato	peso	monossido	prezzo
True	7,3	0,61	8,5	medio
Alpine	14,1	0,86	13,6	medio
Chesterfield	15	1,04	15	alto
GoldenLights	8,8	0,76	9	medio
Kent	12,4	0,95	12,3	medio
MultiFilter	11,4	0,78	10,2	medio
Kool	16,6	1,12	16,3	alto
Tareyton	14,5	1,01	15,9	medio
L&M	14,9	1,02	15,4	medio
ViceroyRichLight	8,6	0,69	10,6	medio

Dilemma: quale dei tre campioni appena considerati scegliamo?

... piano di raccolta dei dati

- Predisposizione del materiale di rilevazione
- Strumento 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

Le rilevazioni statistiche (cont.)

Il Questionario

Sezioni

Domande di controllo

Briefing degli intervistatori

Primo controllo su una parte dei questionari rientrati (*pre-testing*)

Memorizzazione delle informazioni

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*

Fasi di un'indagine statistica (cont.)

Elaborazione

Identificazione metodi e tecniche

Scelta del software

Validazione risultati

Feedback

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

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

Presentazione ed interpretazione dei risultati

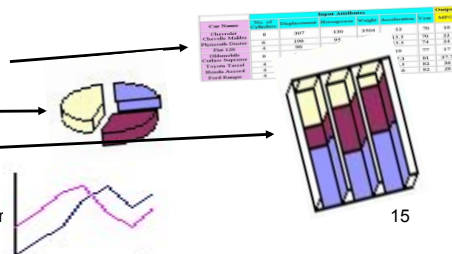
attraverso una nota metodologica e l'uso di tecniche visuali.

i risultati dell'indagine diventano usufruibili per i soggetti destinatari della stessa

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

alcuni esempi di strumenti visuali sono:

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



Fasi di un'indagine statistica (cont.)

Presentazione risultati

Selezione dei risultati da presentare

Scelta del mezzo di diffusione

trasportabilità

usare piattaforme standard

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

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

Fasi di un'indagine statistica (cont.)

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!!!!

Le rilevazioni statistiche

Semplici

Altezza, peso, reddito di una persona. ecc..

Complesse

Monitorare un titolo in borsa per giorni. seguire un sciame sismico. ecc.

...alcune termini con cui familiarizzare...

- “popolazione” o “universo” o “collettivo statistico”
- Campione
- “unità statistica” o “osservazione” o “caso” o “individuo”
- “carattere” o “variabile”

...alcuni termini con cui familiarizzare... (cont.)

- **Modalità** (modo in cui si presenta una variabile)
 - **Numero** (numero figli. esami sostenuti. ecc.)
 - **Attributo** (colori. sapori. qualità. ecc.)
- **Frequenza** (numero di volte che si presenta un modalità)
 - **Absoluta** (numero osservazioni di una modalità)
 - **Relativa** (numero osservazioni di una modalità/totale osservazioni)
- **Serie** → caratteri qualitativi
- **Seriazione** → caratteri quantitativi

...ancora sulle variabili...

Quantitative

- *Quando si considerano caratteri quantitativi queste si definiscono "variabili"*

Qualitative

- *Dette anche più propriamente "mutabili" o "variabili categoriche"*

...ancora sulle variabili...(cont.)

Continue

- *es. reddito, numero di particelle in un liquido, ritardi alla lezione, velocità di trasmissione di un'informazione in un computer*

Discrete

- *Numero figli, numero esami sostenuti, numero errori di stampa in una pagina di un libro*

Le diverse scale di misura

1. Nominale
2. Ordinale
3. Intervalli
4. Rapporti
5. Dicotomica

Variabile su "Scala Nominale" o "Sconnesso"

È una variabile le cui modalità sono nomi

- **Vengono anche chiamate mutabili**
- **Le modalità vengono anche chiamate categorie**

Esempio di scala nominale

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

Variabile su “Scala Ordinale”

È una variabile qualitativa le cui modalità possono essere messe in sequenza

- **Bisogna definire la scala sottostante la variabile**
- **È possibili rendere ordinale una variabile nominale**

Esempio di scala ordinale

- Giudizio (scarso. mediocre. sufficiente. buono. ottimo)
- Reddito in classi
- Titolo di studio (ordinato per anni di studio)

Variabile su “Scala a intervalli”

È una variabile quantitativa le cui modalità sono valori che partono da un'origine arbitraria

- **I valori sulla scala riflettono l'ordine di grandezza delle osservazioni**
- **e.s.: temperature**

Variabile su “Scala a rapporti”

È una variabile le cui modalità sono valori valutati da un'origine che rappresenta la nullità del fenomeno osservato

- **Lo zero è dunque assoluto**

Esempio di scala a rapporti

- Reddito
- Numero di figli per famiglia
- Numero di addetti per imprese
- Età

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**

Es. di variabile dicotomica

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

...ancora sulle variabili... (cont.)

- Le operazioni possibili

- Conteggio
- Organizzazione
- Classificazione
- Differenza
- Differenza relativa

X = variabile

x_1 = modalità 1 di X

x_2 = modalità 2 di X

$x_1 - x_2 = \text{diff.}$

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

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

Le fonti statistiche

- Chi produce statistiche?
- Sotto quale forma?
- Cosa sono le statistiche “non ufficiali”?
- Esistono statistiche sul WEB?

- **ISTAT, EUROSTAT.**
Ministeri, centri di ricerca
riconosciuti

- **Cartacea (annuari, bollettini)**
- **Su supporti informatici (nastri, floppy disk, cd rom)**

- **Sì, è la nuova frontiera!**

- **E' da molto che non senti parlare di “sondaggi”?**