

# L'Indagine Statistica

0011 0010 1010 1101 0001 0100 1011

Prof. Claudio Capiluppi



## FASI DI UNA RICERCA

Prof. Claudio Capiluppi - Università di Verona - A.A. 2006/07

Astrazione

Definizione del Problema di Ricerca

Formulazione Ipotesi di ricerca -> previsione conseguenze

Operationalizzazione: definizione variabili, unità, rif.temp.

Progettazione ricerca: scelta metodo di ricerca, disegno di rilevazione, tecniche e strumenti di osservazione e misura

Rilevazione

Formazione del campione: dimensione e tecnica selezione

Indagine pilota

Indagine principale

Indagine suppletiva di controllo

Elaborazione

Registrazione dati: codifica risposte (aperte)

Analisi qualità dati: controlli validità e coerenza

Revisione: trattamento dati anomali e mancanti

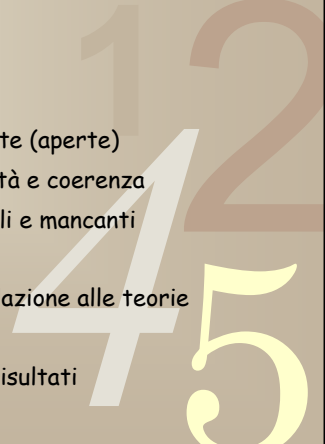
Elaborazione statistica dei dati

Interpretazione

Interpretazione dei risultati in relazione alle teorie

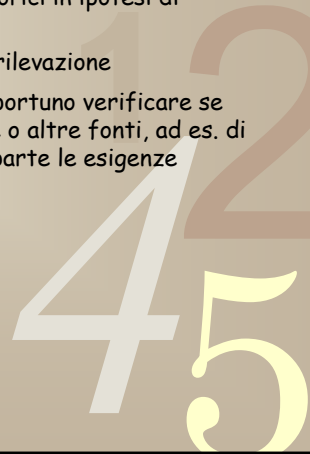
Verifica (statistica) delle ipotesi

Comunicazione, pubblicazione dei risultati



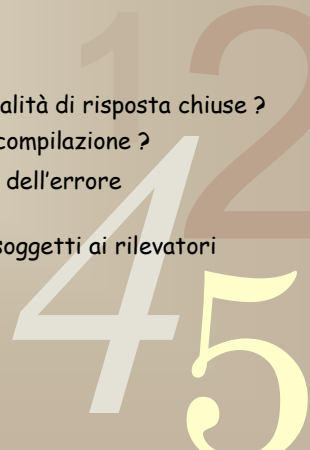
## PROGETTAZIONE DI UNA INDAGINE

- Quando ci si accinge alla realizzazione di una nuova indagine statistica, è necessario in primo luogo esplicitare:
  - definizioni e sistema di teorie ed ipotesi che costituiscono il paradigma all'interno del quale si colloca la ricerca
  - il problema di ricerca: da cui derivano gli obiettivi conoscitivi che si vogliono perseguire
  - l'operazionalizzazione delle ipotesi e dei costrutti teorici in ipotesi di ricerca e variabili misurabili
  - la definizione delle unità statistiche e delle unità di rilevazione
- Prima di avviare una nuova indagine apposita è sempre opportuno verificare se già esistono dati disponibili provenienti da altre ricerche o altre fonti, ad es. di origine amministrativa, che possono soddisfare anche in parte le esigenze informative.



## PROGETTAZIONE DI UNA INDAGINE

- Nella fase di progettazione di una nuova ricerca si dovranno decidere:
  - Il metodo di ricerca:
    - sperimentale o osservazionale ?
    - trasversale o longitudinale ?
  - Il disegno di rilevazione:
    - indagine esaustiva o campionaria ?
    - disegno di campionamento ...
  - La tecnica di indagine
    - intervista aperta o questionario strutturato con modalità di risposta chiuse ?
    - intervista diretta (faccia a faccia), telefonica , autocompilazione ?
  - Disegno per la misura della qualità dei dati, in particolare dell'errore extracampionario:
    - randomizzazione/compenetrazione assegnazioni dei soggetti ai rilevatori
    - replicazione di controllo, reintervista



## USO DI DATI AMMINISTRATIVI

- L'impiego di dati contenuti in registri di natura amministrativa presenta ovvi vantaggi ma anche diversi aspetti problematici.
- Motivi di interesse per il materiale di origine amministrativa:
  - costo ridotto: si azzerano praticamente le spese di rilevazione
  - tempi minimi di raccolta dei dati, utilizzando materiale già esistente
  - costo crescente delle indagini dirette, censuarie ma anche campionarie
  - crescente riluttanza a cooperare dei soggetti coinvolti in indagini statistiche
  - le innovazioni tecnologiche che hanno fatto il loro ingresso nel campo amministrativo: rendono trattabili grandi masse di informazioni prima inutilizzabili
  - continuità temporale e, non di rado, tempestività dell'aggiornamento
  - elevato dettaglio territoriale
  - completezza delle popolazioni osservate
  - assenza di non risposte (dipende !)
  - record individuali
  - possibilità di collegamenti temporali e di analisi longitudinali sulle singole unità (storie individuali)
  - possibilità di collegamenti con informazioni sulle stesse unità provenienti da archivi diversi

## USO DI DATI AMMINISTRATIVI

- Limiti e problemi collegati all'uso statistico di dati amministrativi
  - dati solo in parte utilizzabili a fini statistici, o comunque non adatti agli obiettivi conoscitivi, provenendo da rilevazioni svolte per altri scopi
  - nei sistemi informativi gestionali, l'attenzione è fissata su quegli elementi che sono indispensabili per la corretta gestione di una "pratica" amministrativa
  - il riferimento temporale del dato gestionale può non coincidere con i termini di interesse nella ricerca
  - l'accesso agli archivi amministrativi può essere problematico: legittimità dell'uso di dati per fini diversi da quelli per i quali sono stati ottenuti (privacy)
  - in particolare quando si collegano informazioni relative ad una unità ricavate da archivi gestionali distinti: dalla fusione scaturiscono informazioni aggiuntive sui soggetti rispetto a quelle degli archivi di partenza

## PROGETTAZIONE DI UNA INDAGINE

- Le tecniche di indagine statistica tradizionali sono essenzialmente tre:
  - intervista faccia a faccia
  - intervista telefonica
  - autocompilazione di un questionario
- Tecniche di indagine ibride: il questionario può essere
  - lasciato al rispondente da un intervistatore al termine di un'intervista diretta
  - consegnato ed in seguito ritirato da personale specializzato (es. censimento)
  - inviato per posta per introdurre e preparare una intervista telefonica.
- A partire dagli anni '70 si assiste all'introduzione nelle indagini statistiche di software realizzato per automatizzare diverse fasi dell'indagine, nascono le tecniche d'indagine computer-assisted:
  - CATI: Computer Assisted Telephone Interviewing
  - CAPI: Computer Assisted Personal Interviewing
  - CASI: Computer Assisted Self-administered Interviewing

## PROGETTAZIONE DI UNA INDAGINE

- La scelta della tecnica idonea per una particolare indagine è strettamente connessa con il tema della ricerca e i contenuti sostantivi del questionario:
  - lunghezza e complessità del questionario = durata dell'intervista
  - argomenti trattati: informazioni personali, delicate o imbarazzanti, domande retrospettive, ...
  - tipo particolare di questionario (es. diario, ...)
- Altri aspetti e fattori diversi da considerare:
  - popolazione obiettivo (copertura)
  - estensione della ricerca: numerosità e tipologia di campione
  - durata della rilevazione (tempestività)
  - necessità di identificazione o invece di anonimato
  - complessità logistiche
  - budget e risorse disponibili, non solo economiche

## L'INTERVISTA FACCIA A FACCIA

- L'intervista faccia a faccia è la tecnica di indagine più tradizionale e prevede che un intervistatore raggiunga il rispondente, ad esempio nella sua abitazione, per somministrargli le domande del questionario strutturato.
- L'intervista faccia a faccia comporta l'instaurarsi di un rapporto di interazione asimmetrica tra due individui, che presenta vari aspetti e può avere effetti rilevanti sui risultati della rilevazione e sulla qualità dei dati raccolti.
- La presenza fisica dell'intervistatore risulta vantaggiosa in quanto:
  - è possibile utilizzare tecniche per aumentare la motivazione dell'intervistato
  - è possibile fornire chiarimenti all'intervistato in caso di difficoltà di comprensione dei quesiti
  - è possibile usare materiali di supporto (cartellini con liste di risposte, ...)
  - si hanno le maggiori garanzie riguardo all'identità del rispondente
- L'intervista faccia a faccia si rivela particolarmente appropriata nelle indagini con domande che:
  - si basano sulla memoria, perché l'intervistatore può fornire un aiuto al rispondente nel ricordare e nel collocare temporalmente gli eventi di interesse (evitando difetti di memoria o l'effetto *telescoping*);
  - mirano a verificare la conoscenza di fatti, perché il soggetto deve rispondere quello che effettivamente sa e non può ottenere suggerimenti.

## L'INTERVISTA FACCIA A FACCIA

- Problematiche
  - La presenza dell'intervistatore può agire però anche in senso negativo: non tutti gli argomenti si prestano per essere affrontati in una intervista diretta
  - Effetto dell'intervistatore sulle risposte dei soggetti intervistati (distorsione stime):
    - può condizionare i rispondenti con la sua stessa presenza o commettendo errori di comportamento inintenzionale
    - l'intervistatore può essere inconsapevolmente influenzato dalle proprie aspettative, che possono rendere selettive o distorte le sue registrazioni.
  - Organizzazione complessa: addestramento e gestione del personale coinvolto, spostamenti sul territorio
  - Difficoltà di applicazione di un disegno di controllo dell'errore dell'intervistatore (es. compenetrazione)
  - Costi elevati, soprattutto per il personale impiegato

## L'INTERVISTA TELEFONICA

- L'intervista telefonica è la tecnica di indagine più utilizzata oggi nel mondo, nei paesi con una adeguata copertura telefonica.
- Vantaggi principali dell'indagine telefonica:
  - tempestività del sondaggio
  - minori costi rispetto all'intervista faccia a faccia
  - maggiore disponibilità dei soggetti a rispondere al telefono che a ricevere la visita di un estraneo, ad es. in casa propria (con differenze legate al paese)
  - la presenza dell'intervistatore telefonico ha meno impatto, influisce meno sugli intervistati
- La tecnica ha avuto e sta ancora avendo una grandissima diffusione per diversi motivi:
  - presenza del telefono sempre più capillare
  - costi telefonici calanti
  - disponibilità di sistemi software CATI di supporto all'indagine, che garantiscono:
  - ancora maggiore tempestività a costi ridotti
  - miglioramento della qualità dei dati
  - maggiore produttività degli intervistatori (*scheduling*)

## L'INTERVISTA TELEFONICA

- Problematiche
  - copertura: non tutti comunque hanno il telefono e la distribuzione non è uniforme nella popolazione
  - qualità degli elenchi: utenze affari / residenziali, utenze riservate, cellulari
  - doppie utenze (es. seconde case)
  - mancanza di contatto diretto col rispondente, che può rendere più difficile la gestione dell'intervista
  - minore durata massima dell'intervista (max 20-30 minuti)
  - impossibilità di usare materiale di accompagnamento: es. i cartellini per le liste di risposte lunghe => randomizzare l'ordine delle modalità di risposta
  - meno adatta a domande aperte
- Eccessiva pressione sul rispondente:
  - troppe indagini telefoniche
  - "interferenza" del telemarketing

## IL QUESTIONARIO AUTOCOMPILATO

- L'autocompilazione presenta caratteristiche essenzialmente diverse:
  - minore complessità organizzativa sul campo:
    - il questionario può essere inviato e restituito per posta (indagine postale)
    - costi sensibilmente inferiori a tutte le altre tecniche
    - consente di realizzare indagini di dimensioni superiori (es. censimenti)
  
- L'autocompilazione si rivela indicata ed efficace nel caso di:
  - informazioni particolarmente delicate che i soggetti sono restii a fornire ad un interlocutore presente, fino alla necessità di anonimato
  - quando si richiedono un gran numero di informazioni molto dettagliate (es: diario della giornata)
  - il tema è coinvolgente per gli interpellati.

45

## IL QUESTIONARIO AUTOCOMPILATO

- Problematiche: esistono anche svantaggi da valutare attentamente:
  - non tutti restituiscono il questionario, soprattutto se postale; necessità di prevedere un piano di solleciti e incentivi;
  - autoselezione dei rispondenti, in quanto i mancati ritorni si riferiscono a gruppi di soggetti con caratteristiche differenti dai rispondenti
  
- Non tutte le fasce di popolazione presentano lo stesso grado di partecipazione
  - le fasce meno istruite presentano resistenze a rispondere
  - risultati usualmente ottimi con insegnanti e professionisti
  
- L'impostazione del questionario, la formulazione delle domande, la chiarezza e precisione delle avvertenze sono ancora più fondamentali nel caso dell'indagine autocompilata per limitare l'autoselezione dei rispondenti e le mancate risposte parziali.

45



## IL QUESTIONARIO

- Il questionario può essere considerato:
  - uno strumento di misura
  - uno strumento di comunicazione tra intervistatore e rispondente
- Se la comunicazione non avviene in modo efficace, il questionario non potrà risultare un valido strumento di misura, cioè
  - non permetterà di misurare ciò che ci si è prefissati in modo adeguato (*validità, fedeltà, oggettività, sensibilità*)
  - e di produrre delle stime di buona qualità (*correttezza e precisione*)
- La realizzazione dello strumento è cruciale:
  - è la prima fonte potenziale di errore extra-campionario
  - se ben progettato, permette di limitare e prevenire molti degli errori che si commettono normalmente durante e dopo la rilevazione.
- Progettare un questionario è un'operazione molto più delicata e difficile di quanto non possa sembrare, nella quale si possono commettere molti errori di realizzazione, che si ripercuotono sulla qualità dei dati rilevati.

## IL QUESTIONARIO

- Indicazioni generali per la costruzione del questionario:
  - dimensionare la lunghezza del questionario tenendo presente la tecnica di somministrazione prescelta
  - chiedere solo cose cui l'intervistato può rispondere
  - evitare quesiti "inutili", che forniscono poca o nulla informazione, o che non discriminano, non sono in grado di rilevare differenze
  - strutturare il questionario in aree tematiche
  - studiare con attenzione la sequenza dei quesiti
  - disegnare con attenzione il grafo del questionario: verificare tutti i *percorsi* generati da domande e condizioni *filtro*
  - progettazione concettuale del *data model* del questionario



## IL QUESTIONARIO

- La formulazione dei quesiti
- E' necessario studiare attentamente la formulazione di ciascuna domanda, per arrivare ad individuare quella più appropriata, in relazione agli obiettivi conoscitivi
- Individuare la popolazione obiettivo: utilizzare un linguaggio che possa risultare comprensibile a tutti i rispondenti
- La difficoltà di mettere a punto una buona formulazione è sottolineata dal fatto che:
  - diversi soggetti possono intendere diversamente il testo di una stessa domanda
  - ma anche lo stesso soggetto in momenti differenti ...
- Considerare la tecnica di rilevazione: intervista vs. auto-compilazione

1  
2  
4  
5

## IL QUESTIONARIO

- Indicazioni per la formulazione dei quesiti:
  - utilizzare i termini più semplici
  - usare termini tecnici solo se comprensibili alla popolazione a cui ci si rivolge
  - sono da evitare le espressioni che riducono la comprensibilità e la neutralità della formulazione di una domanda
  - evitare l'uso di termini ambigui
  - evitare l'uso di doppie negazioni
  - evitare l'uso di termini connotati (carichi di valenze o intimidatori)
  - valutare quanto spazio dare al rispondente, cioè quanto usare le domande a risposta aperta
  - valutare il grado di dettaglio a cui si vuole arrivare nella formulazione delle alternative di risposte
  - standardizzare la formulazione (operazionalizzazione) delle variabili chiave: fare riferimento a classificazioni standard, per rendere comparabili i risultati dell'indagine con quelli di altre ricerche
  - quando occorre un livello di dettaglio diverso, usare classificazioni comunque riconducibili a classificazioni standard, mediante aggregazione di modalità
  - valutare con particolare attenzione il ricorso a quesiti retrospettivi

1  
2  
4  
5

## IL QUESTIONARIO

- Lo scopo di una classificazione (o nomenclatura) standard è quello di definire i livelli (modalità) di interesse di una variabile complessa, per:
  - operare confronti nello spazio e nel tempo
  - integrare dati provenienti da fonti diverse
- Una classificazione tipo è uno schema semplificatore di un fenomeno complesso, che deve risultare adatto a diversi utenti e molteplici usi, cosa che impone talora scelte di compromesso, in particolar modo nel caso di confronti internazionali.
- Caratteristiche di una classificazione tipo:
  - struttura gerarchica ad albero, che scende dal livello più generale di massima sintesi fino al massimo dettaglio
  - ad ogni livello, si cerca di ottenere la massima omogeneità tra le entità raggruppate in ciascuna modalità
  - ad ogni modalità è associato un codice, solitamente numerico, che la individua univocamente
  - ad ogni livello la lista deve essere esaustiva, e in generale le voci devono essere disgiunte (senza sovrapposizioni)
  - la ripartizione delle unità tra le modalità dovrebbe essere ragionevolmente bilanciata
  - la classificazione deve essere flessibile, nel senso che deve potere essere adattabile a situazioni nuove senza che sia necessario rivedere l'intera struttura
  - la dimensione deve restare maneggevole, perché sia realisticamente applicabile nelle circostanze concrete.

## IL QUESTIONARIO: LE ISTRUZIONI

- Le istruzioni per la compilazione hanno lo scopo di rendere il processo di acquisizione dai dati più affidabile (stabile) ed oggettivo.
- Nell'intervista faccia a faccia si predispone il libretto delle istruzioni:
  - rappresenta un punto di riferimento stabile per risolvere i dubbi che si possono presentare riguardo il significato dei quesiti
  - si configura come un vademecum di facile consultazione per i rilevatori
  - spiega come condurre l'intervista e gestire la comunicazione
- Nell'indagine telefonica le istruzioni devono:
  - essere impostate, più che come un manuale, come una raccolta di schede situazione-problema, con relativa soluzione
  - poter essere consultate rapidamente anche nel corso della telefonata
- Nel questionario autocompilato, le istruzioni sono rivolte al rispondente con l'obiettivo di agevolare la corretta comprensione del questionario:
  - sintetiche e semplici
  - inserite sul questionario stesso e non su un libretto a parte
  - in nota alla domanda e comunque sulla stessa pagina della domanda
  - le indicazioni di carattere generale devono essere spiegate all'inizio del questionario (brevi altrimenti le saltano)

## IL QUESTIONARIO

- La progettazione concettuale del questionario è un modo diverso di considerare il questionario statistico:
  - non solo come una sequenza di domande con una logica di somministrazione,
  - ma anche come un *data model*, cioè un insieme di *entità* e di *relazioni*.
- Il progettista del questionario ha il compito di individuare la definizione del modello concettuale più idonea per la registrazione delle informazioni previste dal questionario, evitando l'introduzione di entità ridondanti:
  - una modellizzazione corretta della struttura dati assicura non solo un modo più efficiente per la registrazione dei dati
  - ma anche una maggiore facilità d'uso e di analisi dei dati raccolti, soprattutto nel caso di indagini che si ripetono nel tempo
- Spesso accade che chi opera la stesura cartacea del questionario sottovaluti o trascuri completamente la progettazione concettuale:
  - forse a causa di uno scarso interesse o sensibilità o comprensione per gli aspetti informatici del problema, e in particolare della teoria relazionale dei database
  - conoscenze teoriche di base dei sistemi per la gestione informatica dei dati sono necessarie non solo per l'informatizzazione del questionario, ma anche per operare la stesura cartacea.

## IL QUESTIONARIO

- Il problema non può essere ignorato delegando ad un tecnico l'informatizzazione dello strumento, infatti:
  - solo avendo una chiara visione del modello concettuale del questionario, è possibile operare una adeguata stesura cartacea della sequenza di domande, che risulti anche idonea ad essere implementata informaticamente
  - se questo non avviene in fase di stesura cartacea, la successiva fase di informatizzazione richiederà una profonda riprogettazione del questionario
  - spesso il tecnico incaricato non si sente o non è in grado di fare questa riprogettazione, finendo col limitarsi ad una trasposizione pedissequa del cartaceo, che risulterà del tutto inadeguata dal punto di vista concettuale
- Il prezzo di una inadeguata progettazione si paga nel momento in cui si vogliono utilizzare i dati rilevati, cioè in fase di analisi, che risulterà oltremodo laboriosa a causa della frammentazione dell'informazione nelle entità ridondanti:
  - domande che si trovano su più percorsi del questionario
  - domande con liste di risposte diverse per categoria di soggetto
- La normalizzazione del modello concettuale del questionario fornisce delle linee guida, che aiutano a tenere sotto controllo la complessità del questionario:
  - valutare l'opportunità di ridisegnare, se necessario, percorsi troppo intricati in una sequenza di domande più semplice e lineare
  - trovare un compromesso tra controllo dell'intervista e ottimizzazione del *data model*