

Tecniche di indagine statistica

Lezione 10, 11 e 12



Confronto metodi raccolta dei dati /1

- Costi
 - CAWI costi più bassi
 - Tempi
 - apparentemente CAWI più bassi
 - Complessità argomenti trattati
 - supporto intervistatore
 - Implicazioni su disegno d'indagine:
 - **Copertura/frame**
 - CAWI: disponibilità computer/telefono e accesso Internet, contatto unità
 - CATI: disponibilità telefono (fisso, mobile)
- 

Confronto metodi di rilevazione dei dati /2

- Implicazioni su disegno d'indagine

- **Non risposta** (ammontare e informazioni disponibili su unità NR = NR totali)

- Tassi di risposta più elevati:

graduatoria ben definita (soprattutto per indagini su pop.ne adulta):

- 1) Ftf
- 2) tel (soprattutto nel passato)
- 3) posta/web

(dipende anche da a quanto/modo di sollecitare i rispondenti)

- Info su NR: strettamente collegata alle caratteristiche dell'archivio di base (se usato)

web: in generale, permette di definire vari tassi di NR rispetto alle modalità di accesso e compilazione dell'indagine web

Confronto metodi di rilevazione dei dati /3

- Implicazioni su disegno d'indagine

- **Qualità risposte**

- completezza informazioni (= *NR parziali*):

minore in modalità self

risultati non definitivi per web-survey (sembra dipendere da come vengono sollecitate le risposte - risposte obbligatorie o no)

Errori di misura

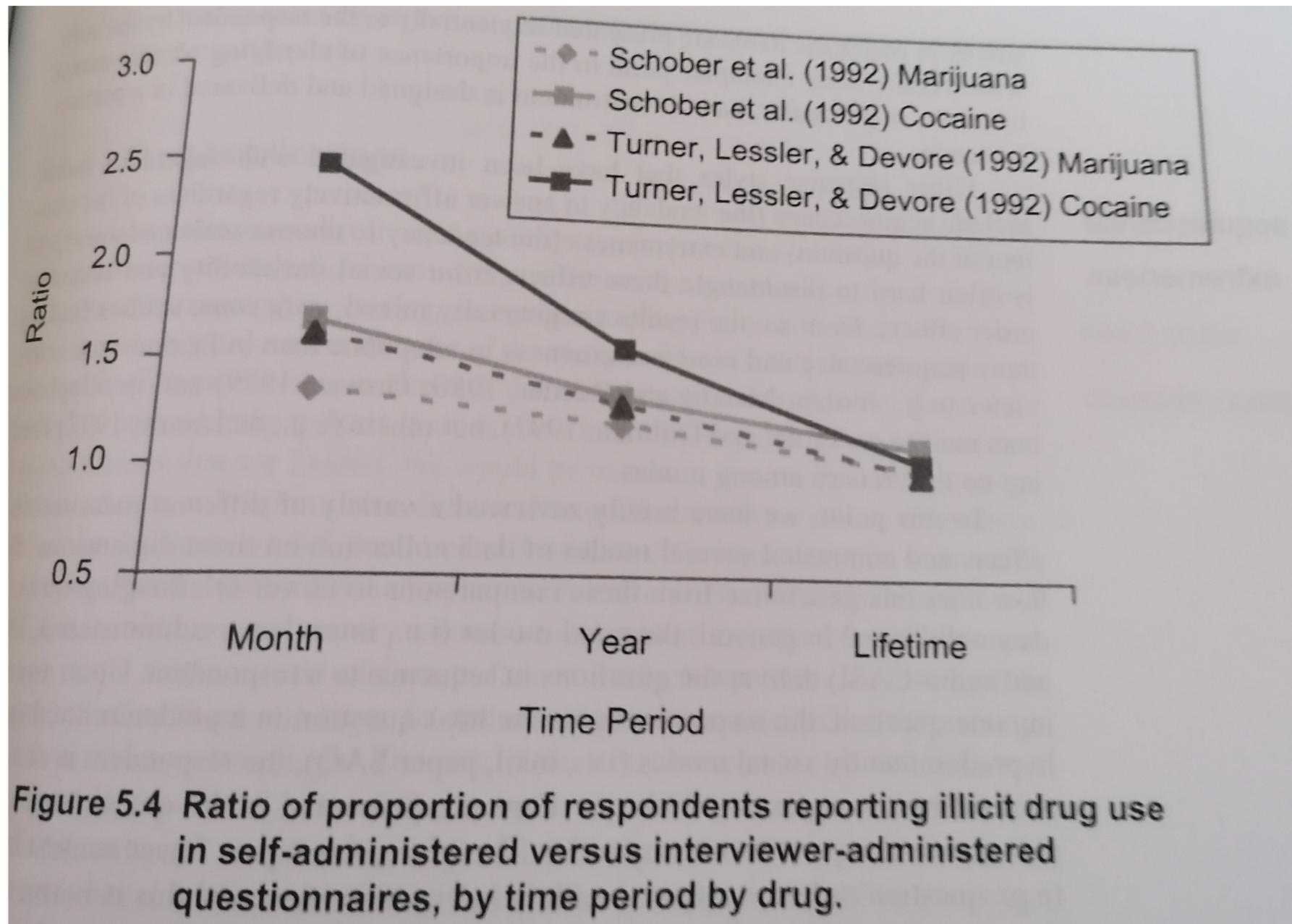
- desiderabilità sociale (= *tendenza a mostrarsi meglio di quello che si è*): maggiore se intervistatore/rice
- “*response effects*” (= risposta influenzata da metodi di raccolta):
 - *wording* (termini/espressioni usate)
 - ordine di presentazione *quesiti*
 - ordine presentazione *risposte*
 - acquiescenza (= tendenza a rispondere affermativamente o ‘abbastanza’)

minori in modalità self rispetto a tel, la direzione degli effetti è comunque legata al metodo - “a voce” o “a vista”

- utilizzo strumenti visivi

(risultati ottenuti spesso da studi sperimentali)

Desiderabilità sociale - Esempio (anche se datato!!!)



Confronto metodi di rilevazione dei dati /3

- Implicazioni su disegno d'indagine:
 - **Disegni multimodali (multiple/mixed-mode design)**

Più di una modalità di raccolta dei dati per:

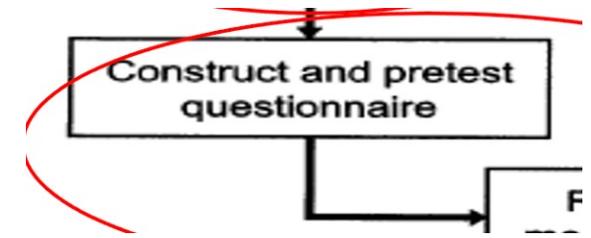
sfruttare il meglio di ciascuna modalità e ridurre gli svantaggi combinando tutti i risultati

Tre motivi principali:

1. Riduzione dei costi. **Come?**
2. Aumentare numero rispondenti. **Come?**
3. Riduzione costi in indagini longitudinali. **Come?**

Utilizzo in aumento, anche se non sempre i risultati sono quelli attesi

Costruzione strumento di misura (*questionario*)



- obiettivi indagine sono tradotti in variabili che devono essere misurate con l'indagine:

- variabili obiettivo/d'interesse (Y)
- variabili ausiliarie/supplementari/covariate (X)

- dai valori osservati (popolazione/campione) di entrambi i tipi di variabili:

- calcolo/stima di parametri relativi alla popolazione (media/mediana... di Y, % di casi con una certa caratteristica, misure di associazione, ...)

- i valori delle variabili (Y, X) sono ottenuti, usualmente, come *risposte a quesiti* (domande) *specifici*

- insieme di quesiti (più qualcos'altro): **questionario**

Questionario

- strumento di **misura**
- strumento di *comunicazione* (spesso unico momento di interazione con unità oggetto di indagine/rispondenti)
- *standardizzato*: domande e comunicazione devono essere identiche per tutte le unità (per garantire la confrontabilità dei dati raccolti) - **strutturato**

Indicazioni per redazione

- codici identificativi (per distinguere il singolo questionario/rispondente)
- quesiti su caratteristiche strutturali (*età, residenza, istruzione, ..., attività economica, sede, n.ro dipendenti, ...*)
- *quesiti specifici (scelta della forma)*
- quesiti per controllo qualità (dipende dalla modalità di raccolta dei dati)

Attenzione a: **processo di formazione delle risposte**

- comprensione quesito (valutare informazione richiesta)
- capacità di rispondere (ricordare/formarsi un'opinione)
- motivazione a fornire una risposta corretta

Tipologie di quesiti/ 1

Quesiti fattuali (*factual*): informazioni su fatti e comportamenti

esiste sempre un valore “vero” che può essere determinato anche senza chiedere al rispondente

salario orario, titolo di godimento dell'abitazione, disponibilità connessione internet, ricovero ospedale, ...

n.b.: è richiesta una precisa definizione del fatto da misurare (es. n.ro di stanze dell'abitazione, salario orario netto o lordo)

Tipologie di quesiti/ 2

Quesiti non fattuali (*nonfactual*): informazioni su attitudini e opinioni

opinione: riflette il parere su un argomento

attitudine: concetto più ampio, in generale su argomento complesso

non esiste un valore “vero”: misurazione di uno stato soggettivo del rispondente che esiste solo nella sua mente in un dato momento

Opinione/attitudine: comportamento di voto; ruolo Unione Europea e singoli stati, soddisfazione, ...

Ci sono modi diversi in cui si ‘forma’ l’opinione

Perché tenerne conto? **Per decidere come proporre i vari quesiti**

- **Online processing model**: impressione generale che le persone hanno di idee, eventi,... che vengono aggiornate e fornite spontaneamente se richieste (es. opinioni politiche)

- **Memory-based model**: l’opinione si forma nel momento in cui è richiesta utilizzando ogni tipo di informazione memorizzata, in genere la più recente (opinioni meno stabili nel tempo)

Testo della domanda /1

(esempi tratti da: Bethlehem, 2009, Applied survey methods, A statistical perspective)

Pensa che i prezzi dei prodotti alimentari aumenteranno allo stesso tasso dello scorso anno, ad un tasso più veloce o più lento?

Usare *termini familiari*

Recentemente è andato al cinema?

Precisione nei *rif.ti temporali*

Possibile alternativa:

La settimana scorsa è andato al cinema?

Testo della domanda /2

E' soddisfatto dei servizi del suo quartiere?

Evitare domande *ambigue*

Nell'ultima settimana, lavorava come dipendente o lavorava in proprio presso una azienda familiare, o faceva entrambe o non lavorava?

Evitare testi *lunghi* (valutare, caso per caso, rispetto al contenuto)

Quante volte ha sentito il suo medico di base negli ultimi 2 anni?

Evitare (se possibile) domande che richiedono il *ricordo di eventi (non particolarmente rilevanti) passati e lontani* nel tempo (più il periodo è lontano nel tempo, maggiore è la possibilità di errore)

Testo della domanda /3

E' d'accordo con la maggior parte della gente che ritiene che la qualità del sistema sanitario in Italia sia scadente?

Evitare domande che possono influenzare la risposta - *leading (loaded) questions*

Molti medici ritengono che il fumo causi il cancro ai polmoni.

E' d'accordo ?

Che cosa si dovrebbe fare con terroristi assassini che minacciano la libertà dei cittadini e la sicurezza dei nostri bambini ?

Evitare l'uso di termini - *loaded words* - che evocano situazioni estreme e che possono influenzare la risposta

Esempio: *esperimento* su opinioni della pop.ne olandese circa il potere del Parlamento Europeo (PE)

Assegnazione *casuale* dei rispondenti ai due quesiti

Q1:

L'aumento del potere del PE può avvenire solo a spese del potere dei parlamenti nazionali.

Ritiene che il potere del PE debba essere aumentato?

Risposte Q1:

Sì = 33% No = 42%
(*valenza negativa*)

Q2:

Molte problematiche superano i confini nazionali. Ad esempio, il 50% delle piogge acide in Olanda viene da altri paesi.

Ritiene che il potere del PE debba essere aumentato?

Risposte Q2:

Sì = 53% No = 23%
(*valenza positiva*)

Testo della domanda /4

Quante ore ha ascoltato le radio locali negli ultimi 6 mesi?

Evitare domande a cui il rispondente **non è in grado di rispondere**

Evitare (se possibile) domande su aspetti “**delicati**” (salute, reddito,...)

Opzioni:

- “confondere” la domanda delicata insieme ad altre meno delicate
- presentare l’argomento come qualcosa di non così inusuale o

estremo

1. Anche ai genitori più *calmi* capita di arrabbiarsi con i propri figli. Nell’ultima settimana, i vostri figli hanno fatto qualcosa che vi ha fatto arrabbiare?
2. Molti medici pensano che un uso moderato di alcol possa ridurre il rischio di infarto. Ha bevuto alcolici nell’ultimo mese?

- scelta tra range di valori per quesiti relativi al reddito (o valori numerici in generale)

Testo della domanda /5

Pensa che si dovrebbe mangiare meno e fare più sport?

Evitare **più aspetti in** una unica domanda (*double questions/double-barreled questions*)

E' contrario a proibire il fumo?

Evitare formulazioni **in negativo** e **doppia negazione**

Se fosse Presidente della Repubblica, come fermerebbe la criminalità?

Evitare domande **ipotetiche**

Testo della risposta - Modalità di risposta: *domande aperte (a risposta libera/non strutturata)*

es. “Qual è l’attività principale che lei pratica nel suo tempo libero?”

.....

“Qual è l’aspetto che ritiene più importante nel suo lavoro?”

.....

“ Quali settimanali ha letto nelle ultime 2 settimane?”

.....

Vantaggi:

- minimo condizionamento
- fase esplorativa
- per argomenti delicati o descrizioni non facilmente classificabili
- maggiori errori (sintesi/registrazione/codifica)

Svantaggi:

- maggior sforzo per il rispondente (maggiore vaghezza)
- maggior dipendenza dal livello culturale del rispondente

Testo della risposta - Modalità di risposta: *domande chiuse (a risposta fissa/strutturata)*

Alternative fisse già predisposte

es. **“A quali mezzi lei ricorre per informarsi sugli avvenimenti sportivi?”**

TV	1	<input type="checkbox"/>
Radio	2	<input type="checkbox"/>
Quotidiani d'informazione generale	3	<input type="checkbox"/>
Quotidiani sportivi	4	<input type="checkbox"/>
Stampa periodica specializzata	5	<input type="checkbox"/>
Internet, social media	6	<input type="checkbox"/>
Nessuno	7	<input type="checkbox"/>

- Vantaggi:**
- codifica immediata
 - sollecitano la memoria
 - meno sforzi per il rispondente
 - lunghezza lista
- Svantaggi:**
- ordine lista
 - risposta non ragionata

* **modalità:** “Altro (specificare)”
“Non so” / “preferisco non rispondere”

Risposte /valutazione-giudizio (rating)

1) **scala verbale** (usualmente) unico giudizio e più alternative di risposta (alternative bilanciate)
(scala di Likert, 1932)
(molto, abbastanza, poco, per niente/
completamente d'accordo, d'accordo,
indifferente, in disaccordo, in completo disaccordo)

2) **rappresentazioni grafiche** - $\begin{matrix} 0 \\ \square\square\square\square\square\square\square \end{matrix} +$

3) **scala (punteggi) numerica** 1-3, 1-5, 1-7, 1-10

In generale - non solo per rilevare una valutazione/giudizio-
da considerare:

- a) scelta scala
- b) n.ro di categorie,
- c) uso di categoria "neutra"
- d) uso di "non so"/"non pertinente"

Effetto modalità di risposta: esperimento con formulazioni alternative/1

Per rilevare il n.ro di libri posseduti, in una indagine su larga scala, mediante *assegnazione casuale*, a metà rispondenti è stata proposta la versione A1 del quesito e alla seconda metà la versione A2:

<i>A1 (n.ro libri)</i>	<i>% risposte</i>	<i>A2 (n.ro libri)</i>	<i>% risposte</i>
< 5	5,2	< 25	62,5
5 – 10	10,5	25 – 30	23,4
10 – 15	35,7	30 – 40	7,8
15 – 20	14,7	40 – 50	4,7
20 – 25	17,7	50 – 70	1,6
> 25	16,2	> 70	0,0
Totale	100		100

Quali considerazioni?

Effetto modalità di risposta: esperimento con formulazioni alternative/2

In una indagine rivolta ad un campione di laureati contattati dopo alcuni anni dal conseguimento della laurea, il quesito riportato sotto è stato proposto, mediante assegnazione casuale, a gruppi diversi di intervistati

- a) *Su una scala da 1 a 10, QUANTO sente essere ADEGUATA la preparazione professionale conseguita all'Università rispetto al tipo di lavoro che svolge?*
- b) *Quanto sente essere ADEGUATA la preparazione professionale conseguita all'Università rispetto al tipo di lavoro che svolge: per niente, poco, abbastanza o molto?*
- c) *Su una scala da 1 a 5, quanto sente essere ADEGUATA la preparazione professionale conseguita all'Università rispetto al tipo di lavoro che svolge?*

Quali risultati attesi?

Questionari e controllo della qualità

- * **sostituzioni** dovute a:
 - (se il piano di campionamento lo prevede)
 - errori lista contatti
 - intervistatore
 - irreperibilità/indisponibilità/rispondente

- motivo sostituzione
- caratteristiche unità non intervistata

- * **chi risponde alle domande** (*proxy-respondent*)

- * **intervistatore**
 - (se presente)
 - codici identificativi
 - caratteristiche strutturali

- * **informazioni su intervista**
 - data
 - ora inizio e fine
 - altre informazioni (se previsto intervistatore)

Redazione del questionario - sintesi/1

- Organizzare complessivamente il questionario in *aree omogenee per tematica trattata*, raggruppando le domande relative allo stesso tema possibilmente nella stessa area.
- Tener conto della modalità scelta per la raccolta dei dati (f-t-f, tel, etc.)
- Individuare per quali argomenti proporre *domande aperte o chiuse*.
- Adottare un *linguaggio comprensibile* a tutti, non ambiguo e preciso e chiedere cose a cui l'intervistato può rispondere.
- Essere precisi nel *riferimento temporale* delle domande e fare particolare attenzione all'uso di *quesiti retrospettivi*.
- Non fare *domande delicate* se non di specifico interesse dell'indagine.

Redazione del questionario - sintesi/2

- Collocare le domande in modo che *non influenzino* le risposte alle successive e controllare la sequenza delle risposte a una stessa domanda.
- Saper individuare la *lunghezza* ottimale del questionario per impegno di tempo, approfondimento di temi, ridondanza di informazioni.
- Impostare *graficamente* il questionario in modo adeguato.
- Adottare *criteri standard per le variabili strutturali*.
- Prevedere una parte del questionario per i *codici identificativi* e una per i *quesiti sui controlli di qualità* (se del caso).
- **Fornire informazioni su normativa privacy** (GDPR).

E dopo la redazione del questionario?

1. **test** (pre-test del questionario)
varie opzioni (da test 'interno' a test con unità che saranno intervistate)
2. indagine **pilota**
'**prova generale**' delle varie fasi di indagine: indagine condotta su piccola scala dall'inizio alla fine, inclusa l'analisi dei dati
(utile molto spesso per stime preliminari dei parametri di interesse al fine di definire/finalizzare il piano di campionamento)

INDAGINE SUL BENESSERE NELLE UNIVERSITA' ITALIANE

Progetto PRIN 2022: "Women's inclusion in Entrepreneurship and Innovation"

Gentile collega,

l'Università di Foggia in collaborazione con l'Alma Mater Studiorum - Università di Bologna, la Luiss Guido Carli, e con il supporto di CINECA, sta promuovendo un'indagine sul livello di benessere lavorativo all'interno della comunità accademica italiana.

Le domande presenti nel questionario allegato mirano a misurare le percezioni sui vari aspetti dell'attività accademica, valutare le relazioni interpersonali, l'equilibrio tra lavoro e vita privata, e il benessere nel contesto universitario. Lo scopo è quello di identificare le aree che potrebbero concorrere a migliorare l'ambiente lavorativo, eventualmente contribuendo a facilitare le attività scientifiche in senso lato.

Il questionario è anonimo e rispetta tutti gli standard etici della ricerca.

L'attività di indagine è svolta nel rispetto dei principi scientifici e di tutela della privacy. Inoltre, le risposte ai singoli questionari non sono accessibili ai ricercatori e saranno elaborate in forma aggregata.

Le chiediamo di compilarlo sperando anche che, leggendolo, possa trovarlo interessante e pertinente alla sua esperienza.

La ringraziamo per la preziosa collaborazione!

<https://forms.office.com/e/v4PqMum6eY>

fermarsi a Q.49