



Tecniche di indagine statistica

Lezione 8 e 9



Fonti di errore e *qualità* (dati) indagine

Qualità dati: concetto complesso che coinvolge *diverse dimensioni* (alcune collegate a *discrepanze* tra fasi/passi successivi di un'indagine - differenza tra valore osservato e vero) e che porta l'attenzione su:

possibili fonti di *errore* e impatto su (stime) risultati (**errore totale**)

Fonti di errore aumentano **incertezza** rispetto alla **accuratezza delle stime** (*distribuzione dello stimatore*):

1. deviazione **sistematica** dal vero valore di popolazione

(**distorsione - bias**) $B(z) = E(z) - Z$ con z stimatore del parametro di pop. $ne Z$

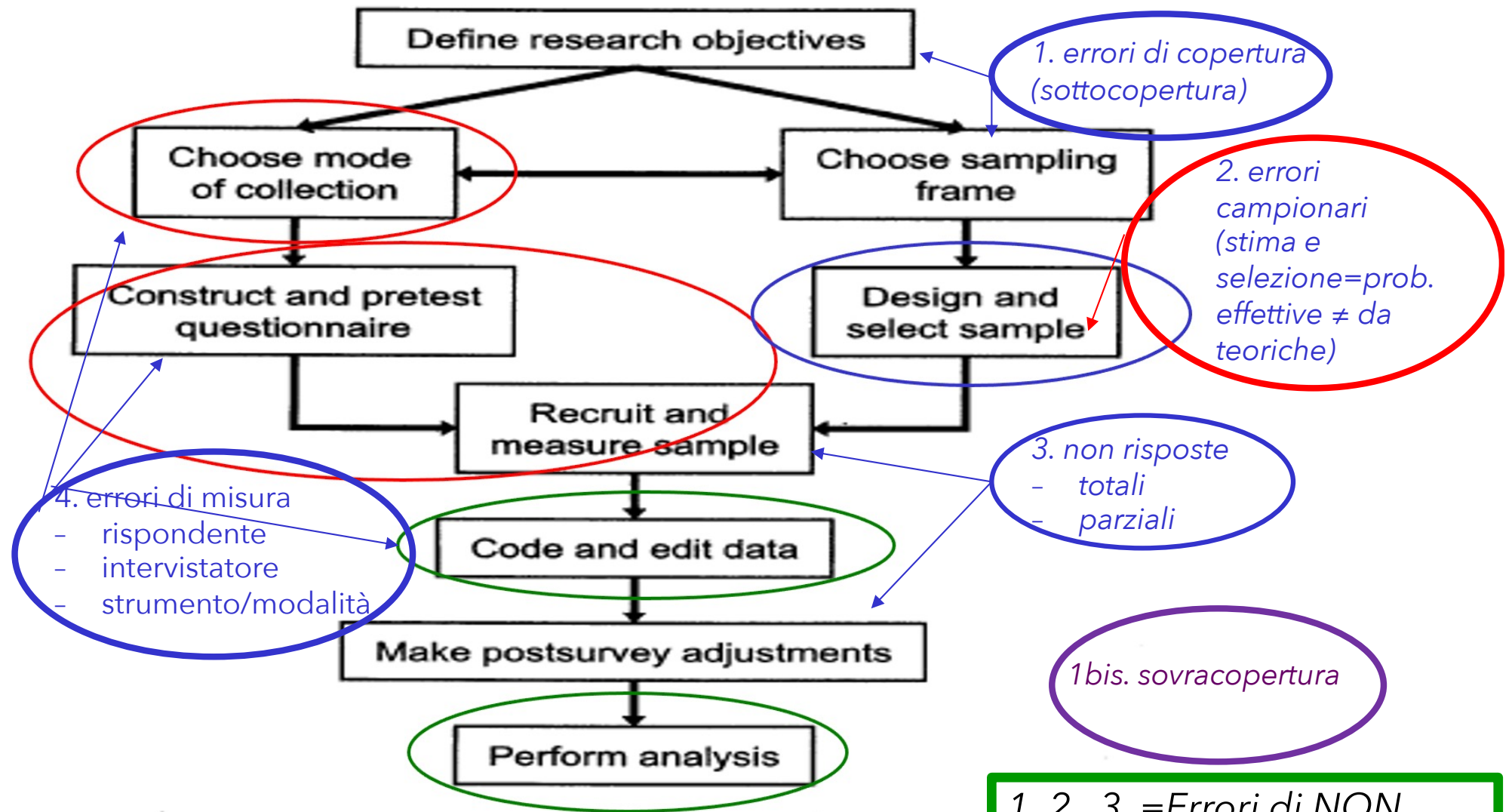
2. aumento della **variabilità** intorno al vero valore di popolazione

$$V(z) = E(z - E(z))^2$$

Poiché uno stimatore può essere **preciso** ma distorto, indicatore più adatto è:

$$MSE(z) = E(z - Z)^2 = V(z) + B^2(z)$$

Ciclo di vita indagine come processo (produttivo) – Fasi indagine e errori

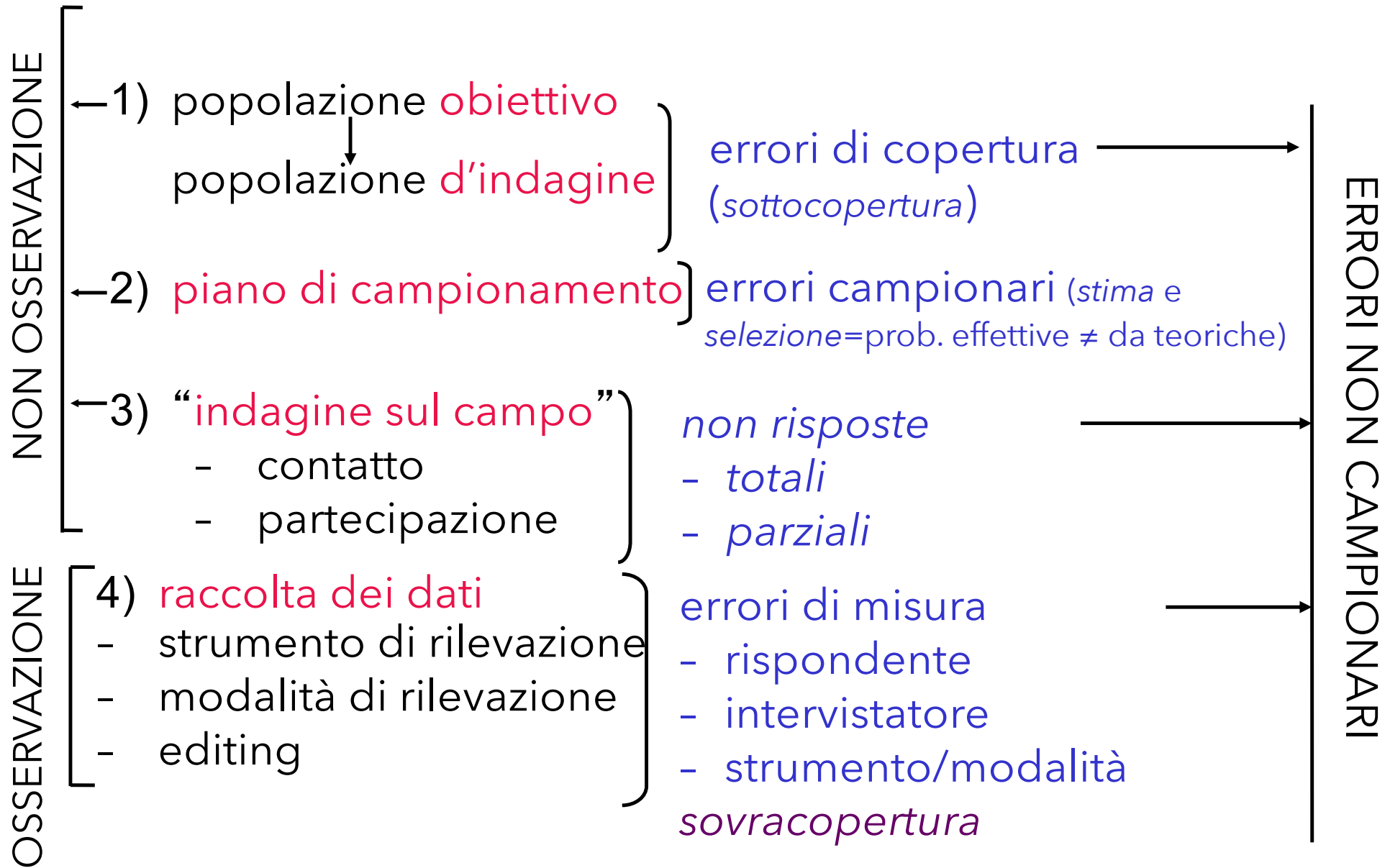


■ Errori **NON** campionari ■ Errore campionario

1., 2., 3. = Errori di NON osservazione
4. = Errore di osservazione

Fasi indagine campionaria e errori

(slide precedente in versione diversa)



Total Survey Error (TSE)

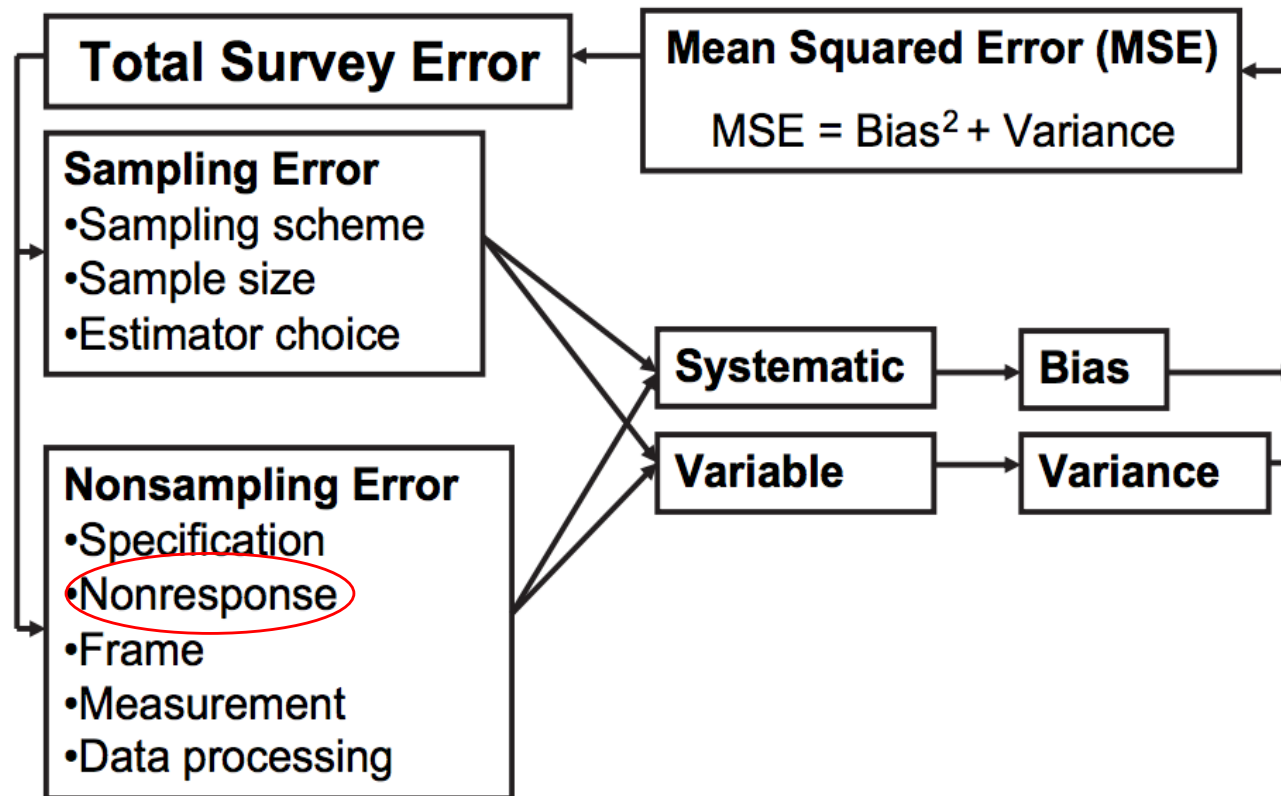


Figure 1. Total Survey Error, Its Components, and the Mean Squared Error.

(Biemer P.P. (2010) Total Survey Error: Design, Implementation, and Evaluation
Public Opinion Quarterly, Volume 74 (5), 817-848

(Processo di) Raccolta dei dati



1. “cattura”

- Metodo/modalità di contatto/recupero informazioni

2. data entry

- registrazione dati (in relazione con p.1)

3. editing

- controlli e correzione errori

4. *aggiustamenti* (eventuali)

- imputazione
- pesi (campionamento, non risposta, post-stratificazione)

(Processo di) Raccolta dei dati e 'tecnologia' /PAI

Metodo:

- intervista diretta (*face to face, f-t-f, usualmente con intervistatore/ricce*)
- intervista *via posta*
- intervista telefonica

PAI - Paper Aided Interviewing (anche P&P)

- modulo (questionario) cartaceo con domande e trascrizione risposte
 - PAPI, Paper and Pencil Interviewing (f-t-f, intervistatore/ricce)
 - postale (cartacea)
 - TEL (intervistatore/ricce)

(Processo di) Raccolta dei dati e 'tecnologia'/CAI

Metodo:

- intervista diretta (*face to face, f-t-f, usualmente con intervistatore/ricce*)
- intervista *via posta*
- intervista telefonica

Data entry e editing
unificate con metodi
CAI

CAI - Computer *Aided* Interviewing

- modulo (questionario) su schermo con domande e trascrizione risposte

- **CAPI**, Computer Assisted PERSONAL Interviewing
- **CATI**, Computer Assisted TELEPHONE Interviewing (**CAMI**)
- **CASI**, Computer Assisted SELF Interviewing
 - **CASQ**, *Computer Self Administrated Questionnaire* (*no intervistatore/ricce e computer rispondente*)

CAWI - uso di Internet (e-mail/web anche su telefono cellulare)

Differenze tra metodi /1

- **Presenza intervistatore/ricce** (effetto)
 - essenziale in *ftf*
 - non prevista in *self*
 - situazioni intermedie (presenza più discreta in TEL, parte *self* in *ftf*)
- **Interazione/coinvolgimento del rispondente**
 - (in modo diverso) massima in *ftf* e *self*
 - moderata in TEL
 - minima se dati ricavati da misurazioni o record esistenti
- **Riservatezza e 'contesto'**
 - chi può sentire/vedere risposte (intervistatore/ricce e/o altri)
 - argomenti "delicati"/sensibili

Differenze tra metodi /2

- **Stimoli comunicativi**
 - verbali (ascolto)
 - visivi (lettura e altro)
 - **Tecnologia**
 - competenze organizzative
 - abilità dei rispondenti
-

- Quali effetti ha $\left\{ \begin{array}{l} \text{metodo} \\ \text{tecnologia} \end{array} \right.$ sulle risposte?
- Quali effetti in generale sulla qualità dell'indagine?

*Numerosi studi condotti per rispondere a tali domande con alcuni risultati abbastanza consolidati su **PAI vs CAI** (spesso studi sperimentali e/o meta-analisi)*

Metodi CAI vs PAI

Pro

- tempi e costi minori
 - controlli maggiori e migliori
(editing può essere immediato, collegamento con altre informazioni, verifica e correzione on line)
- maggiore qualità dei dati

Contro

- copertura (in relazione allo strumento usato)
- in alcune modalità, tassi di risposta molto bassi