

Indagini Statistiche

Alcune definizioni

Popolazione di riferimento: la popolazione target dell'indagine a cui si fa riferimento e a cui poi saranno riportati i risultati

- ▶ Es: Responsabile acquisti alimentari
- ▶ Es: Famiglie Italiane

Unità Statistica: Singola unità appartenente alla Popolazione di Riferimento

- ▶ Es: Individuo Over 18
- ▶ Es: Nucleo Familiare

Sampling Frame: Elenco delle Unità Statistiche appartenenti alla Popolazione di Riferimento

- ▶ Es: Numeri mobili generati casualmente
- ▶ Es: Elenco telefonia fissa

Campione: Sottinsieme del Frame estratto casualmente

Indagini Statistiche ...

in teoria

Le indagini statistiche si dividono in 2 tipi:

- ▶ **Indagini Censuarie:** viene contattato tutto il Frame
- ▶ **Indagini Campionarie:** viene estratto un Campione Casuale dal Frame

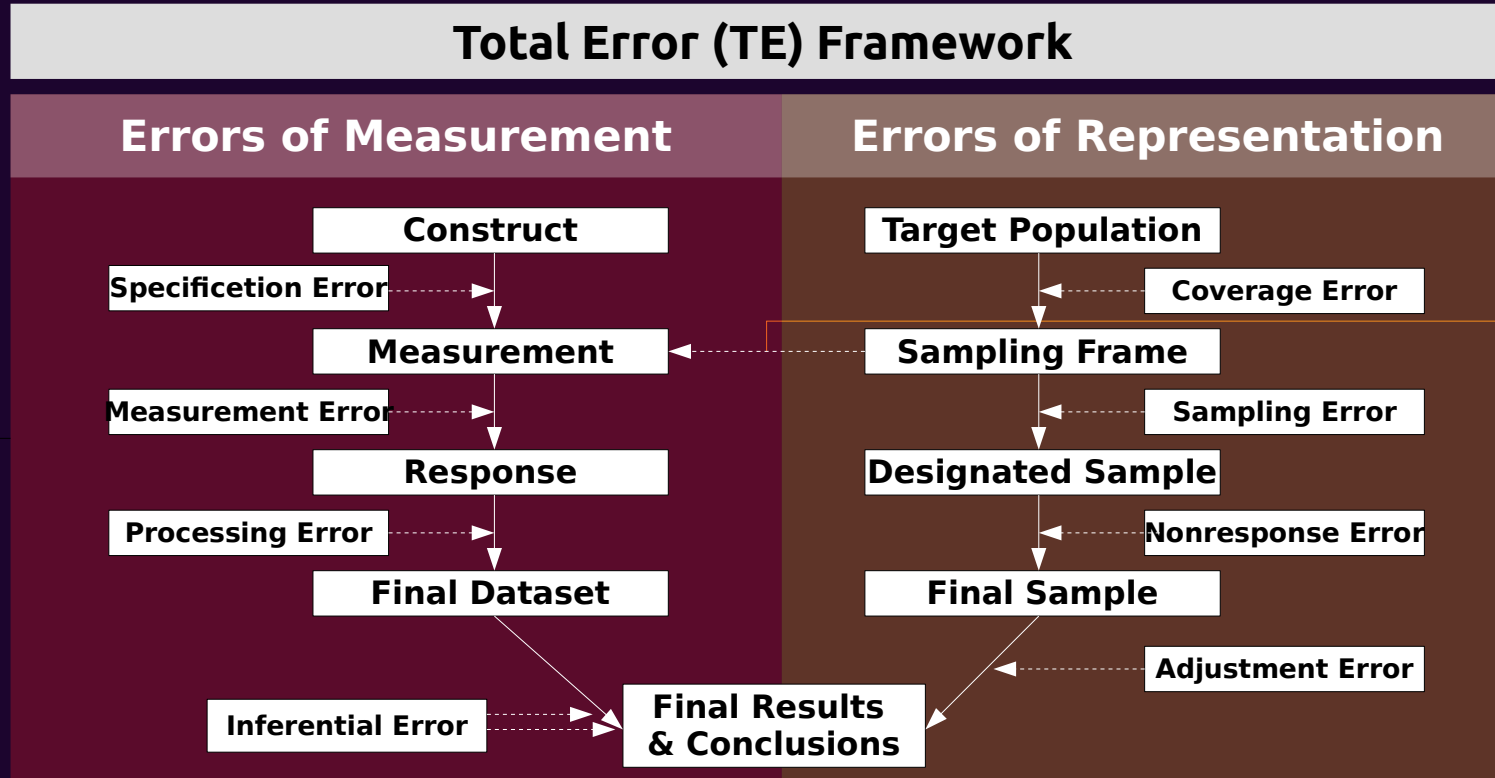
Se il campione è estratto casualmente, i risultati ottenuti dal campione, possono essere estesi a tutta la popolazione.

Questo processo si chiama **INFERENZA**.

In questo caso si dice che il campione è **RAPPRESENTATIVO** della popolazione.

Indagini Statistiche ...

in pratica



Il **Sampling Frame** influenza il modo attraverso cui vengono raccolti i dati

L'errore

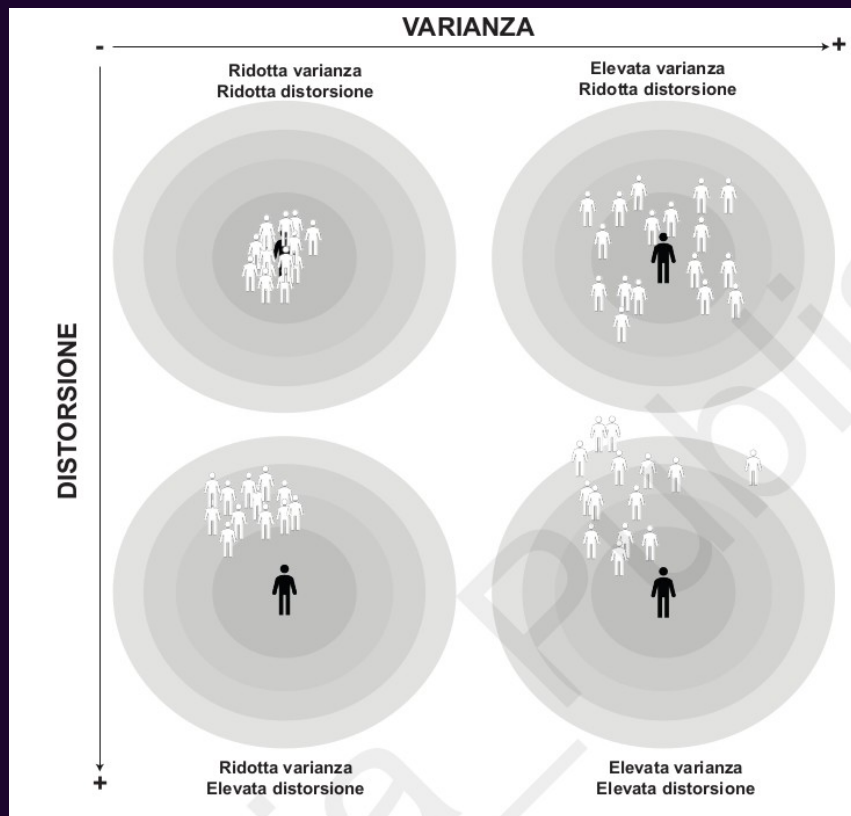
Errore Totale dell'Indagine

=

Varianza

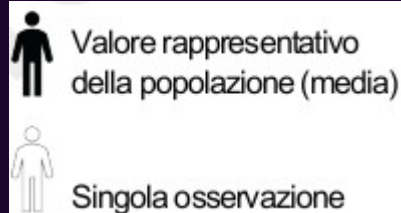
+

Distorsione



Errori casuali

Errori sistematici



Coverage Error

Le liste a disposizione

Elenco telefonico:

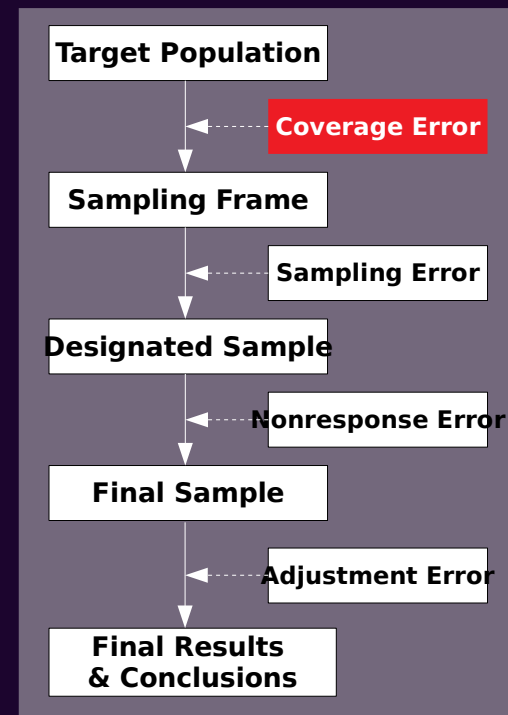
- ▶ 51,3% delle famiglie ha la linea fissa, di cui il 66,1% ha il numero nell'elenco telefonico
- ▶ → 34% della popolazione è rintracciabile nell'elenco

Numeri mobili generati casualmente:

- ▶ 95,9% delle famiglie ha il cellulare

Panel online commerciale:

- ▶ 76,1% ha accesso a internet
- ▶ Opinione.net: 22,500 panelisti



Sampling Error

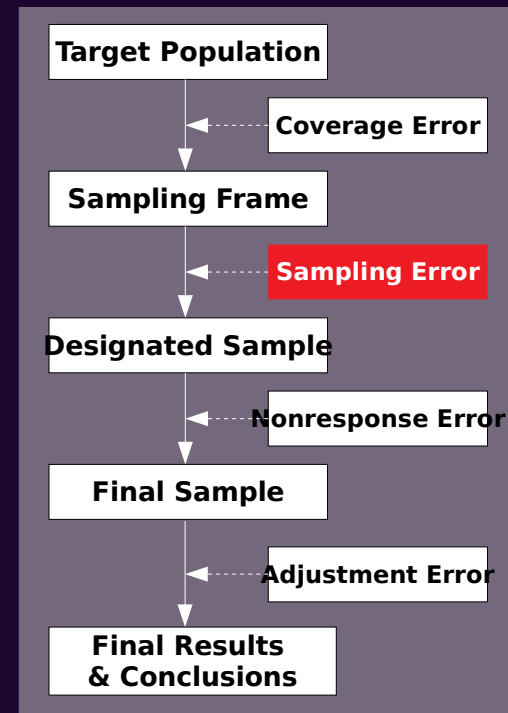
I tipi di campionamento

Casuale Semplice: Le Unità Statistiche vengono estratte casualmente dal Sampling Frame

Stratificato: Le Unità Statistiche vengono estratte in base all'appartenenza a degli strati definiti a priori (**es:** zona di residenza)

- ▶ **Proporzionale:** la numerosità di ogni strato è calcolata in maniera proporzionale rispetto all'universo
- ▶ **Ragionato:** la numerosità degli strati è definita in modo 'ragionato' per ottenere numerosità minime in ogni strato

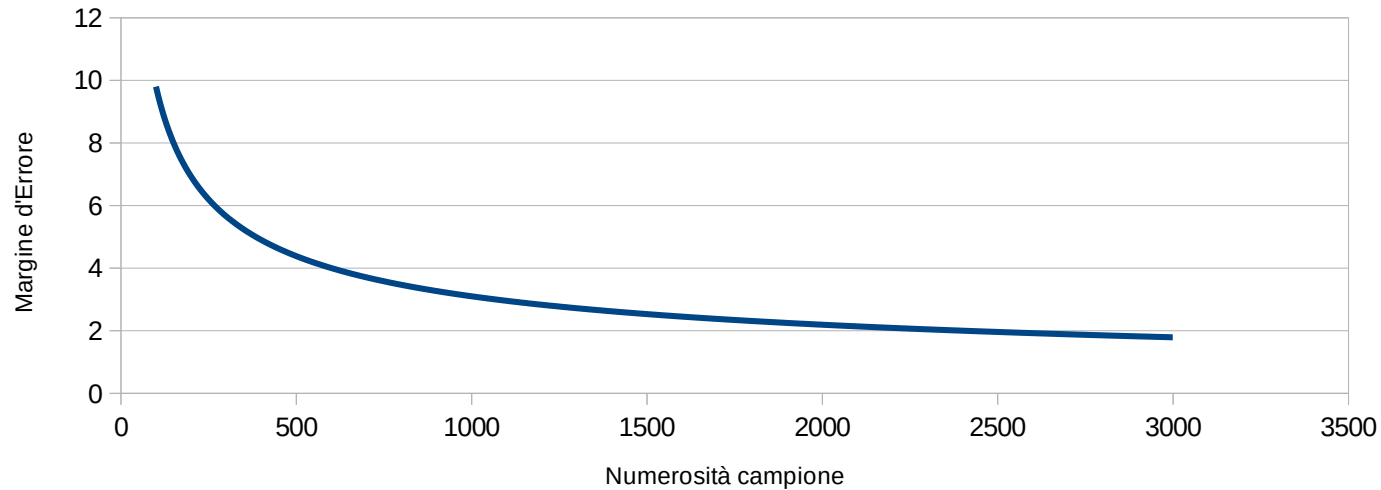
Per Quote: Si definiscono delle quote (**es:** genere-età) che vanno a chiudersi mano a mano durante l'indagine



Sampling Error

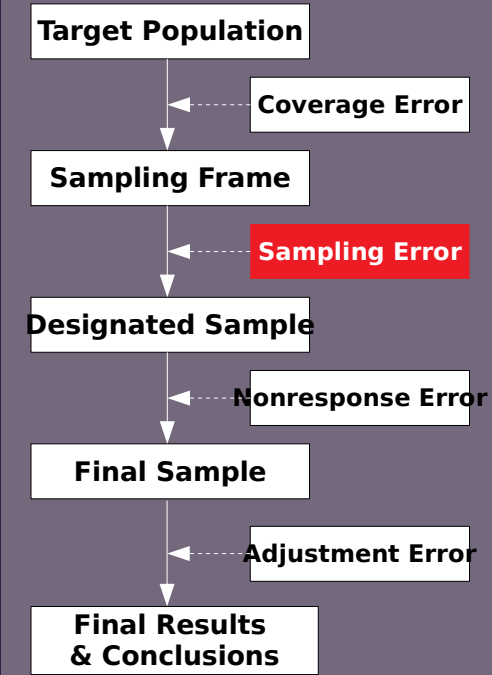
I tipi di campionamento

Andamento del Margine d'Errore al crescere della Numerosità Campionaria



Media in popolazione = media campione $\pm$ margine di errore

▶ **Es:** media campione = 15
Margine di errore = 2
 $13 \leq \text{Media in popolazione} \leq 17$



Sampling Error

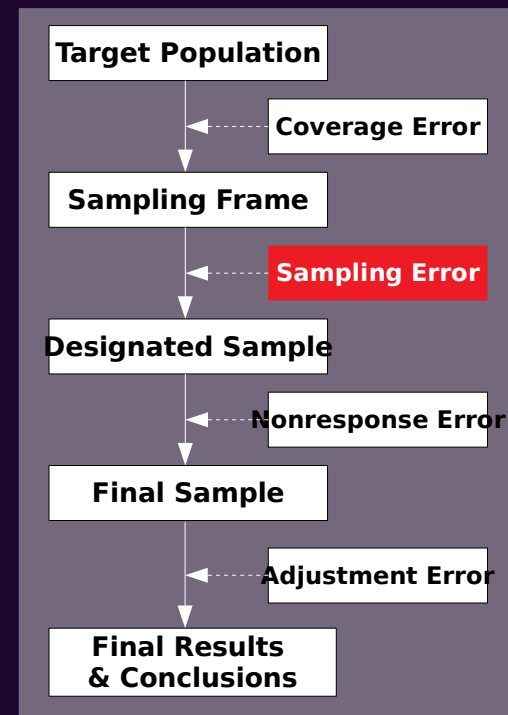
Calcolo Numerosità e Margine d'Errore

Calcolo numerosità campionaria: su **demetra.com** potete trovare lo strumento per il calcolo della numerosità campionaria e del margine d'errore.

Calcolo numerosità campione	
Livello di confidenza:	☐ 95% ☐ 99%
Intervallo di confidenza:	
Numerosità Universo:	
Numerosità Campionaria:	
Calcola	

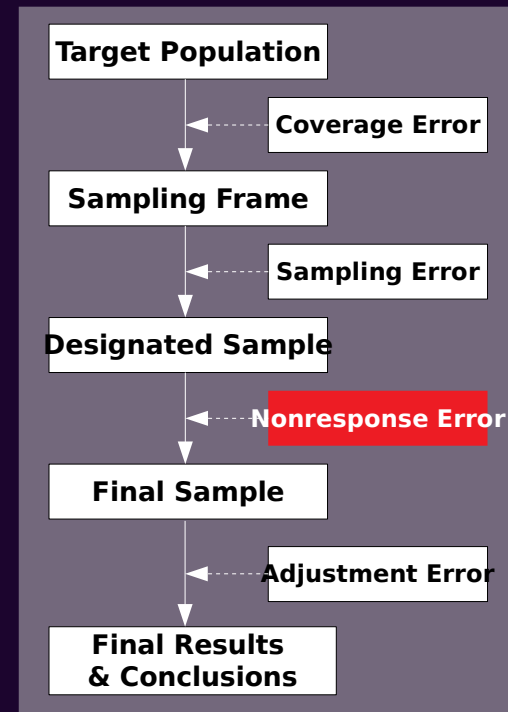
Calcolo intervallo di confidenza	
Livello di confidenza:	☐ 95% ☐ 99%
Dimensione campione:	
Numerosità Universo:	
Percentuale della risposta:	
Intervallo di confidenza:	
Calcola	

Percorso per lo strumento: **demetra.com** → **Campioni Statistici** → **Calcolo numerosità campione**



Non Response Error

- ▶ Il **Campione selezionato** non è sempre quello poi effettivamente ottenuto a causa della **Non Risposta**
- ▶ Se la **Non Risposta** è in qualche modo collegata alla **Variabile d'Interesse**, allora si ha l' **ERRORE di NON RISPOSTA**
- ▶ L'**Errore di Non Risposta** è tanto più grande quanto più la **Non Risposta** è collegata alla **Variabile d'Interesse** e tanto più è alto il tasso di **Non Risposta**



Adjustment Error

Pesatura

Sample weights:

Basati sulle probabilità di selezione, si utilizzano quando non tutte le unità hanno la stessa probabilità di essere estratte
(**es:** campioni stratificati non proporzionali)

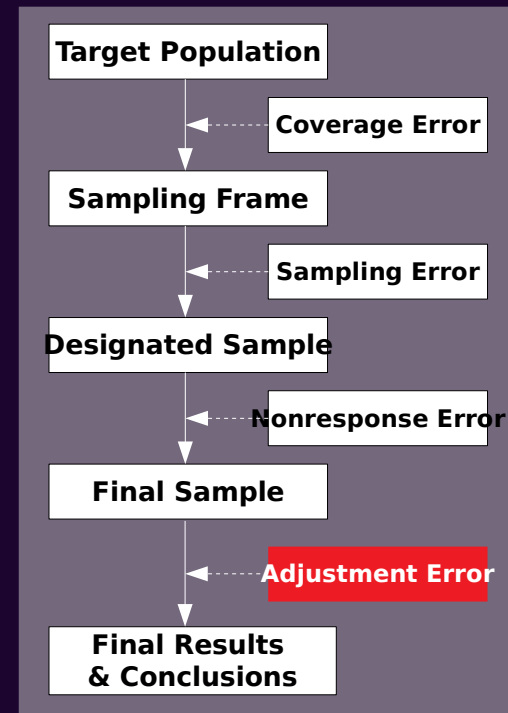
Nonresponse weights:

Servono per correggere i tassi di risposta differenti in diversi sottogruppi, si basano sulle variabili utilizzate per la definizione delle quote

Post-stratification weights:

Utilizzati per riportare ai totali di popolazione la distribuzione di variabili che non erano state usate nella stratificazione iniziale ma che si pensa siano collegate con la variabile d'interesse

(**es:** titolo di studio)



Conclusioni

Errori & Soluzioni (o almeno ci si prova)

Coverage error

- Utilizzo più frame
- Domande di screening a inizio questionario

Sampling Error

- Utilizzo campioni casuali
- Calcolo numerosità ottimale

NonResponse Error

- Procedure per ridurre le non risposte
- Pesatura dei dati

Adjustment Error

- Calcolo di diversi tipi di pesi
- Valutazione delle stime ottenute

