



# Tecniche di indagine statistica

---

Lezione 22



# Quando si può fare di meglio del CCS ?

- Si hanno informazioni *ausiliarie* sulla popolazione *stimatori rapporto* (in genere non corretti ma più precisi)
- La popolazione è suddividibile in *gruppi omogenei* al loro interno.
- Le liste sono presenti per *gruppi di unità* e non per l'intera popolazione (*struttura gerarchica* delle liste).
- I costi per raggiungere le unità possono *variare notevolmente* e disegni diversi comportano costi molto inferiori.

# Campionamento su più stadi /1

Popolazione può essere ben definita ma **non necessariamente può essere agevole** raggiungere le sue unità (*unità di analisi*)

Es: campione di 400 famiglie residenti in una area per stimare il n.ro di assicurazioni possedute (10.000 famiglie in totale)

- a. CCS di 400 famiglie (se possibile, anche stratificato)
- b. suddivisione dell'area in **500 blocchi** di ( $\approx$ ) 20 famiglie ciascuno e indagine a tutte le famiglie residenti in 20 blocchi selezionati a caso dai 500 totali:

blocchi = unità primarie o 1<sup>o</sup> livello/clusters (*primary sampling units PSU*)

famiglie = unità 2<sup>o</sup> livello (*secondary sampling units SSU*)

Quali conseguenze di *b*) rispetto ad *a*) ?

# Campionamento su più stadi /2

Popolazione può essere ben definita ma **non necessariamente può essere agevole** raggiungere le sue unità (*unità di analisi*)

Es: campione di 1500 studenti scuole superiori FVG per stimare le conoscenze di matematica (49.516 a.s. 2020/21)

Liste studenti disponibile per scuola (*gruppi di unità*) e non per l'intera popolazione

**Struttura gerarchica** delle liste poiché **organizzazione** della pop.ne è gerarchica:

- popolazione finale di unità (di analisi) è contenuta in un aggregato di unità di livello (stadio) superiore, le quali - a loro volta - possono essere contenute in ulteriori unità sempre più ristrette in numero e ampie in dimensione



# Campionamento su più stadi - motivazione

- Necessarie solo le **liste delle sottopopolazioni** contenute nelle unità selezionate al livello superiore
- Rilevazione concentrata nelle unità primarie:
  - agevolata l'organizzazione del lavoro sul campo (formazione delle liste, selezione del campione, reclutamento del personale per la rilevazione, supervisione del lavoro sul campo)
  - facilità di esecuzione della rilevazione (minori spostamenti, i rilevatori conoscono e sono conosciuti dai rispondenti, ecc.); controllo di copertura
  - riduzione dei costi e tempi di esecuzione
  - unità reperibili presso le comunità (famiglie, scuole, reparti operativi, ecc.)

## Tuttavia

- Campione complesso, stime complesse
- Rischio di inefficienza delle stime (dovuta a **omogeneità interna delle unità primarie**)

# Campionamento a grappoli (one-stage cluster sampling)

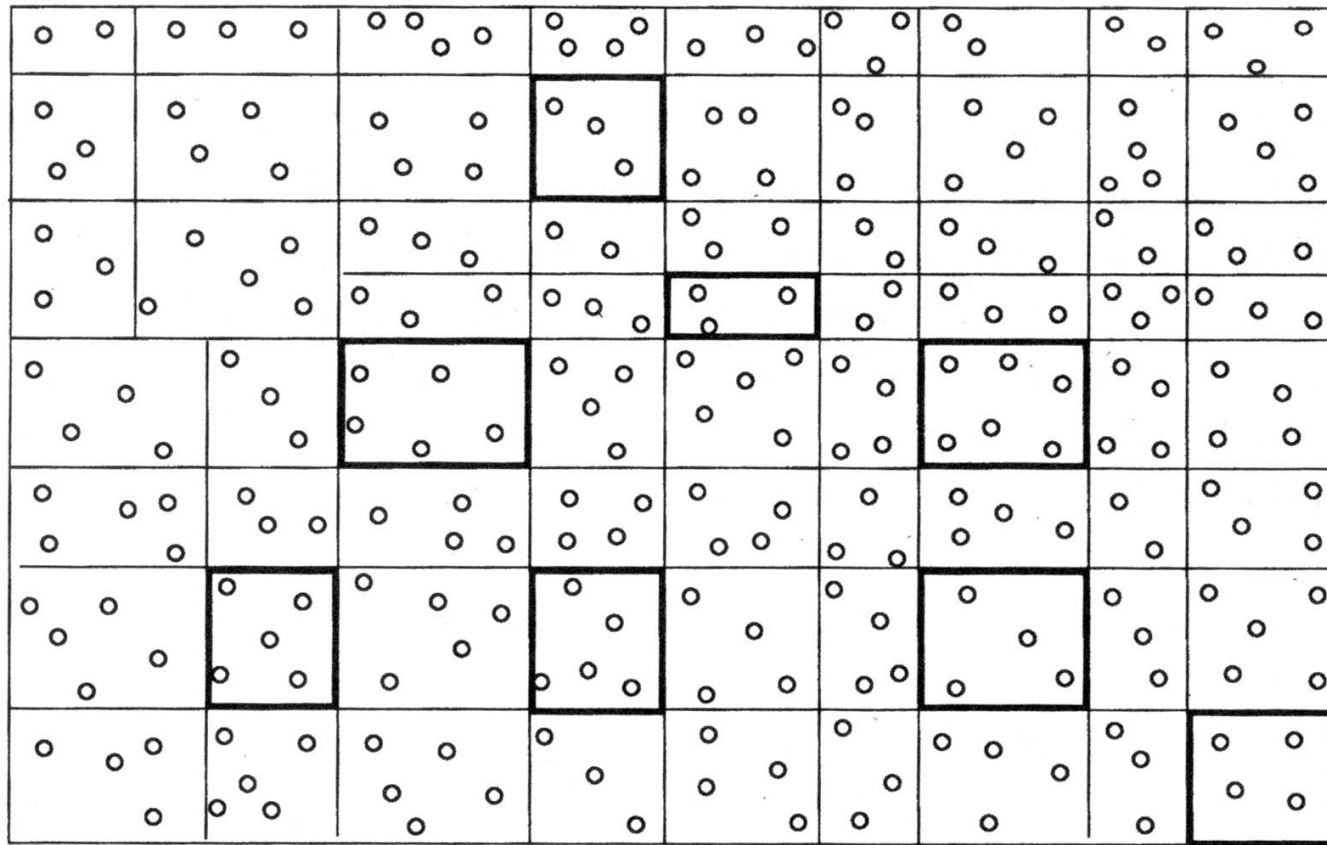


Fig. 3. Campionamento a grappoli: grappolo ( $\begin{matrix} \square \\ \circ \end{matrix}$ ), unità elementare ( $\circ$ ).

# Campionamento a due stadi (two-stage cluster sampling)

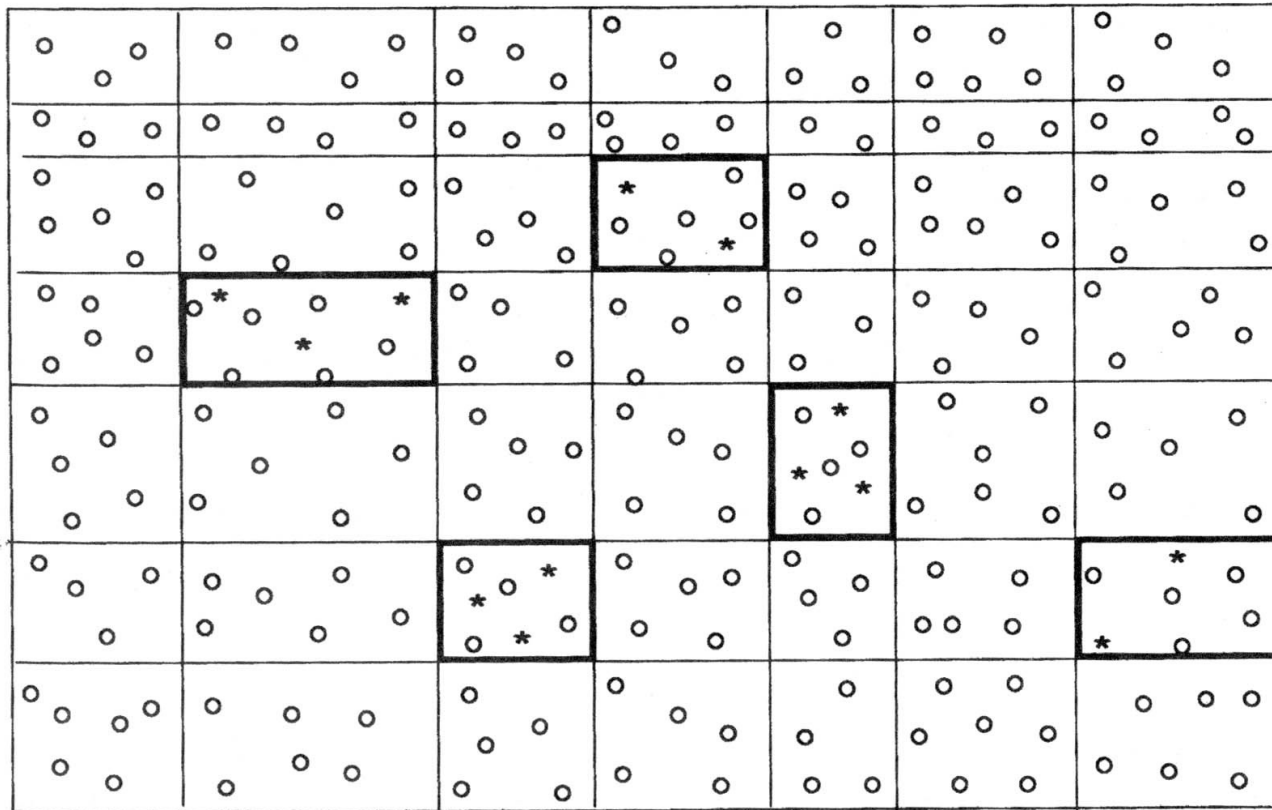


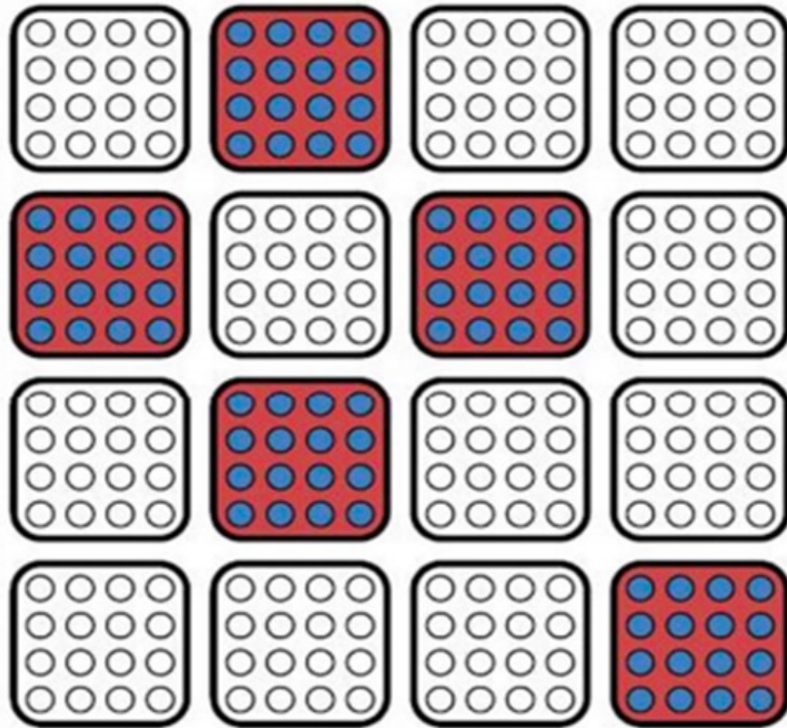
Fig. 4. Campionamento a due stadi: unità della popolazione (○), unità primaria (□), unità elementare (\*).



# Campionamento a grappoli vs a (due) stadi

## Campionamento a grappoli

**G = 5**



## Campionamento a 2 stadi

**I stadio: G = 5; II stadio:  $n_i = 3$**

