

# **RILEVAZIONI CAMPIONARIE E TECNICHE DI CAMPIONAMENTO**



# **FASE 1 (INDAGINE STATISTICA)**

**Programmazione e RILEVAZIONI**  
(riguardare da parr. 2.2 a 2.7 libro Girone-Pace)

**LE RILEVAZIONI CAMPIONARIE**



# I TERMINI TECNICI

Intendiamo per:

**Popolazione (universo)** l'insieme N di unità statistiche che costituiscono l'oggetto di studio

Insieme N= **ampiezza della popolazione**

**Campione** l'insieme delle n unità campionarie (*casi*) selezionate tra le N unità che compongono la popolazione

Insieme n= **ampiezza del campione**

**Campionamento** il procedimento attraverso il quale si estraggono le n unità campionarie (ovvero un numero ridotto di casi) dal complesso delle N unità della popolazione

**Rappresentativo** se si basa sulla casualità piuttosto che sulla 'comodità'



# VALORE ESATTO E STIMA

## Metodi di campionamento



# ESEMPIO CAMPIONE

Se costruiamo un campione di studenti universitari estraendo a sorte un certo numero di studenti fra quelli presenti un qualsiasi giorno in università, non otteniamo un campione probabilistico per 3 motivi:

- a) i non frequentanti hanno probabilità nulla di entrare;
- b) le matricole, che frequentano di più, hanno maggiore probabilità di essere estratte rispetto agli studenti degli anni successivi;
- c) questi ultimi hanno maggiore probabilità rispetto agli studenti fuori corso



# Formazione dei campioni

## **CAMPIONI PROBABILISTICI**

- ✓ **Campione casuale semplice**
- ✓ **Campione stratificato**
- ✓ **Campione sistematico**
- ✓ **Campione a due o più stadi**



# CAMPIONAMENTO CASUALE

- a) Campionamento casuale semplice con ripetizione
- b) Campionamento casuale senza ripetizione



# A) CAMPIONAMENTO CASUALE SEMPLICE CON RIPETIZIONE

- Per formare **il campione** basta numerare tutti gli  $N$  elementi dell'universo ed estrarre da un'urna  $n$  elementi con l'accortezza di rimettere di volta in volta il numero estratto nell'urna



## **Nei CAMPIONE CASUALE SEMPLICE (BERNOULLIANO) CON RIPETIZIONE**

- La popolazione d'origine resta invariata in quanto la scelta presuppone che ogni elemento estratto venga 'rimesso nell'urna' in modo da mantenere costante e uguale a  $1/N$  la probabilità di estrazione del  $n$  successivo



## **B) CAMPIONAMENTO CASUALE SENZA RIPETIZIONE**

L'unità che entra a far parte del campione viene esclusa dalle successive estrazioni quindi la probabilità di estrazione dei numeri successivi cambia perché cambia la composizione dell'urna



**Nel CAMPIONE SENZA RIPETIZIONE**

La popolazione d'origine si riduce di un'unità per cui la probabilità di scelta del 1° elemento, 2° elemento ecc. saranno:  $1/N$ ;  $1/N-1$ ;  $1/N-2$



# VANTAGGI E LIMITI CAMPIONAMENTO CASUALE

## **PREGI**

- ✓ Il campione è statisticamente rappresentativo
- ✓ nessuna unità o gruppo di unità è favorito
- ✓ presenta una notevole semplicità nel metodo
- ✓ ha un basso costo operativo

## **DIFETTI**

- è necessaria una lista completa delle unità della popolazione (lista di campionamento)
- la selezione è completamente affidata al caso



**CAMPIONAMENTO STRATIFICATO** quando la numerosità della popolazione è molto elevata, per cui è difficile numerare tutti gli elementi e si corre il rischio di non inserire nel campione elementi **D'INTERESSE, SI RICORRE**

Il metodo si articola in **3 fasi** :

1. si procede alla suddivisione delle unità della popolazione di riferimento in gruppi (denominati **STRATI**);
  - il raggruppamento deve essere fatto in modo che ciascun elemento compaia in un solo strato;
  - gli elementi appartenenti allo stesso strato devono essere simili tra loro (*strati/classi omogenee*);
  - per raggruppare la popolazione in strati omogenei si utilizza come criterio di stratificazione una variabile correlata al fenomeno studiato (*vedi secondo esempio*)



## ESEMPIO CAMPIONAMENTO STRATIFICATO

Se si intende stimare il **reddito** di una popolazione tramite un campione di individui, i soggetti della popolazione di riferimento possono essere suddivisi in strati sulla base della variabile *occupazione*.

Se gli strati sono 4: operai, impiegati, ceti medi autonomi, liberi professionisti,  
-da ognuno di questi strati estraiamo un campione  
-quindi uniamo i 4 campioni parziali in un campione complessivo



## ESEMPI CAMPIONI STRATIFICATI

Il Rettore conduce un'indagine relativa all'opinione sui servizi dell'Ateneo. La popolazione studentesca e gli amministrativi è composta da 21.909 persone, così suddivisi:

Dati:  $N_1 = 6.204$  studenti residenti;  $N_2 = 13.304$  studenti non residenti;  $N_3 = 2.401$  amministrativi

La grandezza del campione deve essere pari a  $n=100$  in cui il numero di individui selezionato da ogni strato è proporzionale al peso che lo strato ha sulla popolazione

$$n_1 = f N_1 = 6204/21909 = 0,28 * 100 = 28 \text{ studenti residenti}$$

$$n_2 = f N_2 = 13304/21909 = 0,61 * 100 = 61 \text{ studenti residenti}$$

$$n_3 = f N_3 = 2.401/21909 = 0,11 * 100 = 11 \text{ amministrativi}$$

Dove  $f=n/N$  è la frazione di campionamento



# CAMPIONAMENTO SISTEMATICO

- Questa tipologia di campionamento è equivalente a quella del campionamento casuale semplice da cui differisce solo per la tecnica di estrazione
- Le unità campionarie non vengono più estratte con sorteggio ma selezionate sistematicamente una ogni intervallo

## CAMPIONAMENTO SISTEMATICO

Selezioniamo un numero casuale compreso tra 1 e  $k$  (oppure lo scegliamo rapportando la 'popolazione' alla numerosità del campione da noi stabilito) e includiamo nell'indagine l'individuo corrispondente e i  $k$ -esimi individui successivi.

Per esempio facciamo un'indagine ogni  $k=8$  individui. Selezioniamo casualmente un numero tra 1 e 8, ad esempio 5. Questo significa che includeremo nell'indagine il quinto, il tredicesimo ( $5+8$ ), il 21-esimo ( $13+8$ ), il 29-esimo ( $21+8$ ) individuo e così via sino a raggiungere l'ampiezza campionaria desiderata.

**la prima unità è scelta casualmente, le altre secondo il criterio prefissato che è definito passo di campionamento**

TAVOLETTA



## ESEMPIO CAMPIONAMENTO SISTEMATICO

- Da una popolazione di  $N=8235$  unità si desidera estrarre un campione di  $n=500$ , si sceglierà una unità ogni 16 arrotondando per difetto ( $K=8235/500=16,7$ ) cominciando con un numero estratto a caso tra 1 e 16. Se questo è 12, allora i soggetti scelti saranno il 12°, il 28° ( $12+16$ ), il 44° ( $28+16$ ), etc.



# VANTAGGI E LIMITI CAMPIONAMENTO SISTEMATICO

## PREGI

- ✓ Semplifica l'operazione di estrazione
- ✓ Permette di ottenere campioni casuali anche nella situazione in cui manchi una lista della popolazione e l'ampiezza di questa (cioè  $N$ ) sia sconosciuta

**Esempi: exit poll, ricerche di mercato**

**N.B.: una stima dell'ampiezza della popolazione aiuterebbe a determinare un valore appropriato di  $k$**

## DIFETTI

- Risente molto di più rispetto al campionamento casuale del grado di accuratezza della lista di campionamento



# CAMPIONAMENTO A STADI

Nel caso in cui la popolazione dovesse risultare molto elevata, per semplificare la procedura di estrazione e per ridurre i costi di rilevazione, si usano i **campioni a più stadi.**



- La popolazione viene suddivisa su più livelli gerarchicamente ordinati
- Si ricorre ad uno schema che comporta l'individuazione di *unità primarie, secondarie e di ordine superiore di rilevazione*



## ESEMPIO CAMPIONE A STADI

**Esempio 1:** dovendo costruire un campione nazionale di maestri elementari, il campionamento si può effettuare in 2 stadi:

- Al *1° stadio* si estrae un campione di unità primarie rappresentate dai circoli didattici
- Al *2° stadio* si estrae un campione di unità secondarie, ovvero i maestri, da ognuna delle unità primarie selezionate alla prima estrazione

**Esempio 2:** Se volessimo indagare su un campione di famiglie italiane sarebbe conveniente estrarre:

- prima un campione di province (*1° stadio- unità primarie*),
- poi nell'ambito di ciascuna provincia selezionata campioneremo un numero di comuni (*2° stadio- unità secondarie*)
- infine nell'ambito di ciascun comune selezionato sceglieremo un numero di famiglie (*3° stadio- unità finali*)



# VANTAGGI E LIMITI CAMPIONAMENTO A STADI

## **PREGI**

- risparmio di tempo e di costo operativo
- è necessaria la sola lista delle unità appartenenti ai stadi di interesse

## **DIFETTI**

- non assicura la stessa rappresentatività del campionamento stratificato
- richiede un considerevole lavoro affinché si individuino le unità primarie e le unità di stadio successivo



# **CAMPIONI NON PROBABILISTICI**

Un campione non probabilistico (o di convenienza) gli individui inclusi nell'indagine sono inseriti in maniera **NON CASUALE** (scelte condizionate dal ricercatore)

- ✓ **Campione a scelta ragionata**
- ✓ **Campione accidentale**
- ✓ **Campione a valanga**



# **CAMPIONAMENTO A SCELTA RAGIONATA**

Le unità campionarie si ottengono non in base a criteri di casualità, ma su una selezione soggettivamente impostata.

La scelta del campione viene effettuata

- sulla base di considerazioni razionali
- si fonda sul giudizio del ricercatore il quale conosce il fenomeno e con criteri, più o meno personali, cerca di individuare un campione della popolazione



# CAMPIONAMENTO ACCIDENTALE (.....)

Il ricercatore sceglie le prime persone che capitano come soggetti su cui indagare.

In questo caso è sufficiente l'adesione volontaria dei rispondenti

- Es. campioni formati dai lettori di un giornale che rispondono spontaneamente a determinate indagini

(**risposta volontaria**)

- **Vantaggi:** risparmio di tempo e denaro
- **Limiti:** si perde l'accuratezza del campionamento



# CAMPIONAMENTO A VALANCA

**La procedura consiste nell'individuare i soggetti da inserire nel campione a partire dagli stessi soggetti intervistati**

- **Fase 1:** si identificano e si intervistano le persone in possesso di determinate caratteristiche. Tali soggetti sono utilizzati come informatori per identificare altri individui con peculiarità simili da inserire nel campione.
- **Fase 2:** consiste nell'intervistare i soggetti su cui indagare e far in modo che questi ultimi suggeriscano a loro volta altri nominativi da inserire nel campione
- Metodo utilizzato nello studio *di popolazioni clandestine* (gruppi sociali che tendono ad occultare la loro identità)



# QUANTO DEVE ESSERE GRANDE UN CAMPIONE?

La scelta della dimensione del campione dipende ed è direttamente proporzionale :

1. ALLA VARIABILITÀ TRA GLI ELEMENTI DELLA POPOLAZIONE  
(natura della popolazione)
2. AL LIVELLO DI PRECISIONE RICHIESTO PER UNA STIMA  
(finalità dell'indagine)



# QUANTO DEVE ESSERE GRANDE UN CAMPIONE?

ATTENZIONE:

**La dimensione del campione non necessariamente deve essere proporzionale alla dimensione della popolazione.**

- Se la grandezza della popolazione aumenta, quella del campione richiesto per una stima accurata non deve necessariamente diventare più grande
- Un campione di 1000 casi potrebbe produrre stime affette da errori molti simili sia che sia stato estratto da una popolazione di 10 mila che di 100 mila soggetti



# LA SCELTA DELLA DIMENSIONE DEL CAMPIONE

## 1. VARIABILITÀ TRA GLI ELEMENTI DELLA POPOLAZIONE

- L'ampiezza del campione necessaria perché si rispecchi il valore che una variabile assume nella popolazione dipende dal grado di eterogeneità della variabile all'interno della popolazione
- Le popolazioni con una variabilità (o eterogeneità) maggiore (un  $\sigma$  più grande) necessitano di un campione più grande
- Se la popolazione è omogenea rispetto ad una variabile sarà sufficiente un campione piccolo per ottenere una stima accurata (la grandezza del campione è ininfluenta)



# LA SCELTA DELLA DIMENSIONE DEL CAMPIONE

## 2. LIVELLO DI PRECISIONE RICHIESTA PER UNA STIMA

- Più è grande la precisione richiesta, maggiore dovrà essere la dimensione del campione in esame
- L'ampiezza de campione si costruisce a partire dal livello di fiducia, ovvero *dall'intervallo di confidenza*, stabilito per la stima e *dal grado di errore* che siamo disposti ad accettare



# AMPIEZZA DEL CAMPIONE

## **Altro aspetto importante:**

Il ricercatore deve tener conto del tipo di analisi che intende effettuare e quindi dei frazionamenti ai quali sottoporrà il campione (analisi multivariata).

**ESEMPIO:** in una ricerca sul comportamento elettorale se si vogliono studiare le caratteristiche degli elettorati di piccoli partiti, un campione di 50 casi può:

-essere sufficiente per una stima sul numero dei voti che vanno ad un partito piccolo;

-non può essere utilizzato per mettere in relazione il voto alle caratteristiche (genere, età, titolo di studio) degli elettori



# BONTÀ DI UN CAMPIONE

Per **bontà di un campione** intendiamo:

- la sua validità scientifica
- la sua rappresentatività

**Un campione è rappresentativo quando fornisce un'immagine in piccolo ma senza distorsioni della popolazione**

