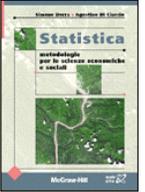


Indagine statistica

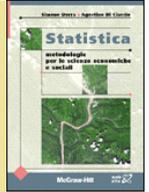


- ◆ Indagine Totale
- ◆ Indagine Campionaria

Fasi dell'indagine

- **definizione degli obiettivi**
 - definizione delle unità e delle variabili da rilevare
 - scelta del periodo di riferimento
- **individuazione della popolazione e della lista delle unità statistiche**
- **definizione del piano di campionamento**
- **raccolta dei dati**
 - scelta della tecnica di rilevazione
 - formulazione del questionario e pretest
 - rilevazione sul campo
- **registrazione dei dati**
 - registrazione su supporto magnetico
 - controllo e correzione
- **elaborazione e analisi dei dati**

La rilevazione dei dati



La raccolta delle informazioni può essere **completa** oppure **parziale**.

È **completa** quando si esaminano tutte le unità statistiche che compongono la popolazione oggetto di studio.

Pregi:

- Accuratezza delle stime anche a livelli territoriali molto spinti
- Ricchezza delle informazioni raccolte
- Esaustività

Difetti:

- Costo elevato
- Tempi di elaborazione dei dati molto lunghi
- Qualità dei dati non elevata

È **parziale** quando ci si limita a studiare un sottoinsieme, detto "**campione**" dell'insieme di riferimento.

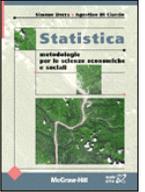
Pregi:

- Continuità della rilevazione
- Economicità
- Indagini più mirate e approfondite

Difetti:

- Riferimento territoriale non spinto
- Variabilità campionaria

Statistica descrittiva e inferenza



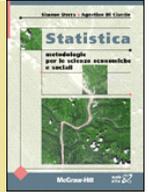
La **statistica descrittiva** fornisce gli strumenti per sintetizzare ed esplicitare in forma corretta il modo in cui il fenomeno si è manifestato nel collettivo osservato

Mediante l'**inferenza** statistica è possibile misurare e controllare l'attendibilità delle informazioni provenienti da un campione



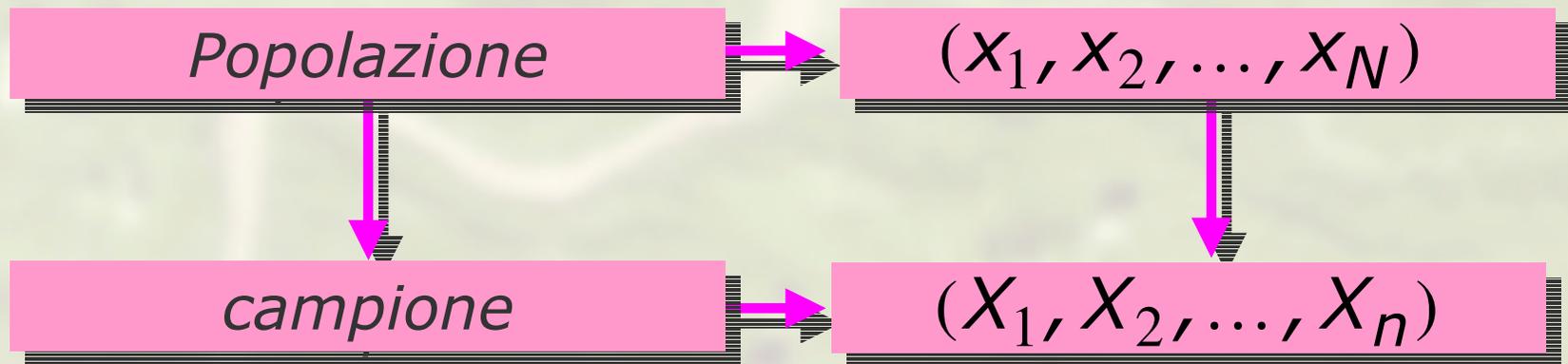
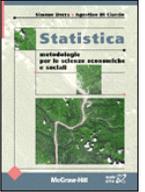
estrazione
del
campione

Estrazione del campione



- ◆ **Campionamento casuale:** insieme di tutte quelle tecniche di formazione del campione in cui la selezione delle unità è affidata a regole probabilistiche.
- ◆ **Campionamento casuale semplice:** i campioni della stessa dimensione estraibili da una popolazione hanno uguale probabilità di essere estratti.
- ◆ **Campionamento casuale stratificato:** la popolazione viene suddivisa in un certo numero di strati. Da ogni strato in maniera indipendente viene poi estratto un campione casuale semplice.

Campionamento da Popolazioni finite



n = dimensione campionaria

n/N = frazione di campionamento

Campionamento da Popolazioni finite

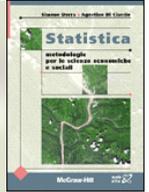


Se la **regola di selezione** del campione è di tipo probabilistico, l'estrazione del campione avviene in accordo con qualche specifica distribuzione di probabilità. Per una **selezione probabilistica** è necessario individuare:

- ✦ lo spazio campionario Ω , formato da tutti i possibili campioni estraibili con una medesima tecnica da una popolazione.
- ✦ la probabilità di ogni campione c in Ω di essere estratto

La coppia $\{\Omega, \text{probabilità dei campioni in } \Omega\}$ è detta **piano di campionamento**.

Campioni casuali



I campioni possono essere estratti casualmente dalla popolazione:

- **con ripetizione**: una volta estratta un'unità viene rimessa dentro la popolazione e quindi potrebbe essere nuovamente estratta;
- **senza ripetizione**: una volta estratta un'unità questa viene messa da parte e quindi non può essere estratta più di una volta.

Due campioni **non ordinati** di uguale numerosità sono diversi tra loro se almeno un'unità del primo campione non è contenuta nel secondo campione. Nei campioni **ordinati** conta invece anche l'ordine con cui si presentano le diverse unità.

Esempio disegno campionario

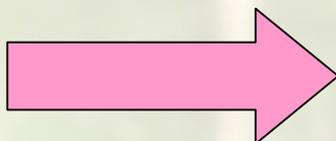


Popolazione composta da 4 grandi aziende.
($N=4$); Carattere="Fatturato annuo";

$$x_1 = 52, x_2 = 49, x_3 = 65, x_4 = 74$$

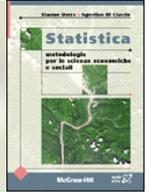
Spazio campionario Ω , costituito dai campioni *ordinati* di dimensione 3, estratti *senza ripetizione*.

$c_1 = 52 \ 49 \ 65$	$c_7 = 49 \ 65 \ 74$	$c_{13} = 65 \ 74 \ 52$	$c_{19} = 74 \ 52 \ 49$
$c_2 = 52 \ 65 \ 49$	$c_8 = 49 \ 74 \ 65$	$c_{14} = 65 \ 52 \ 74$	$c_{20} = 74 \ 49 \ 52$
$c_3 = 49 \ 52 \ 65$	$c_9 = 65 \ 49 \ 74$	$c_{15} = 74 \ 52 \ 65$	$c_{21} = 52 \ 74 \ 49$
$c_4 = 49 \ 65 \ 52$	$c_{10} = 65 \ 74 \ 49$	$c_{16} = 74 \ 65 \ 52$	$c_{22} = 52 \ 49 \ 74$
$c_5 = 65 \ 52 \ 49$	$c_{11} = 74 \ 49 \ 65$	$c_{17} = 52 \ 65 \ 74$	$c_{23} = 49 \ 52 \ 74$



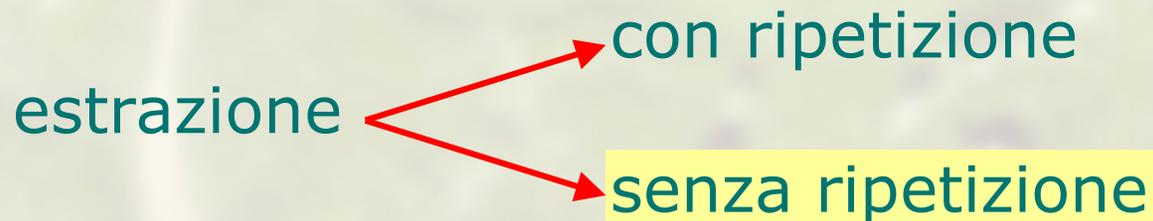
Ogni campione ha uguale probabilità di essere estratto, pari a $1/24$

Campionamento casuale semplice – Popolazione Finita



Nel campionamento casuale semplice i campioni di uguale dimensione hanno tutti stessa probabilità di essere estratti.

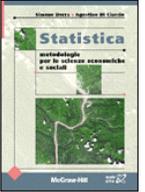
- ✦ si devono conoscere le unità della popolazione;
- ✦ tutte le unità devono essere reperibili;
- ✦ si deve procedere all'estrazione casuale delle unità.



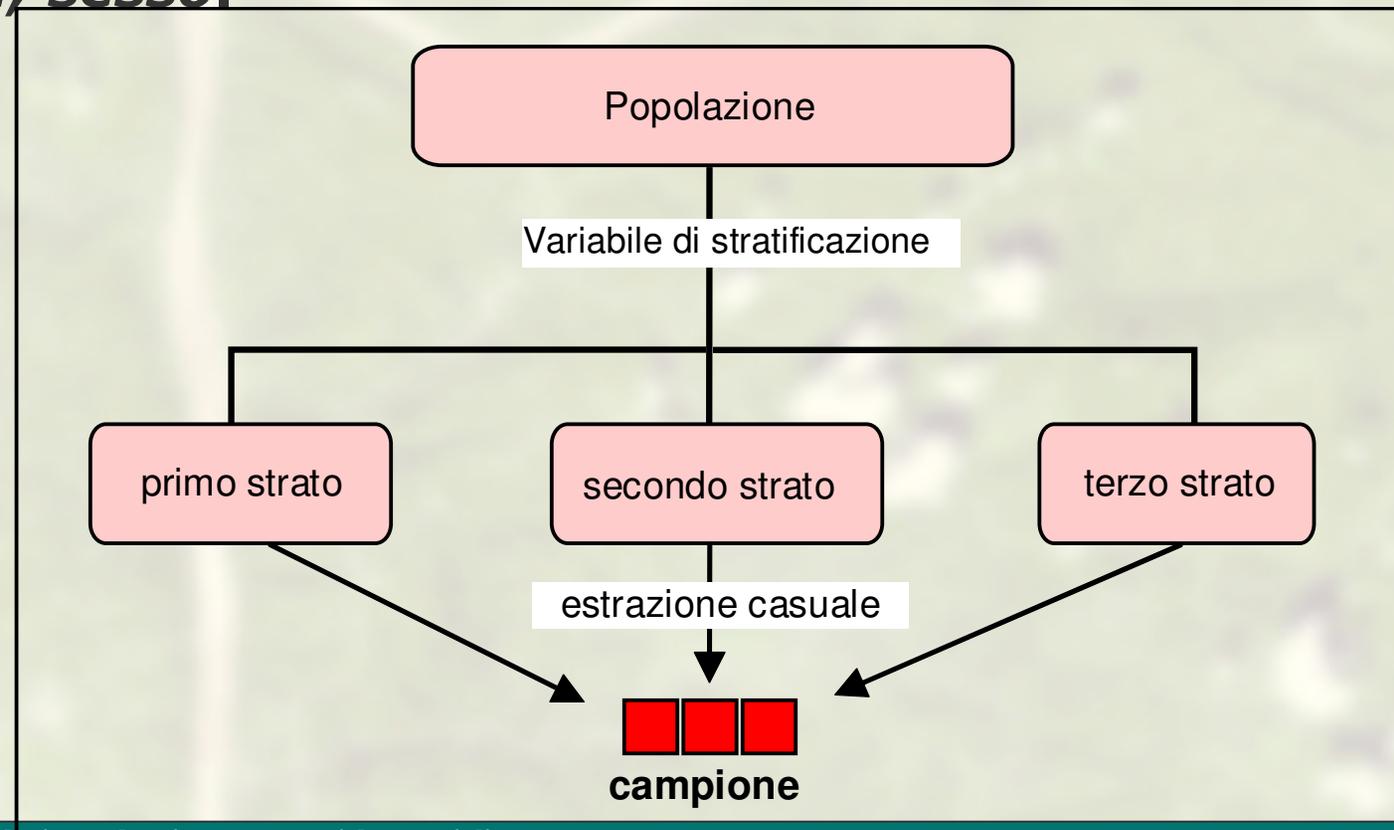
Il campione casuale ottenuto con estrazioni senza ripetizione è composto da n variabili casuali che hanno **marginalmente stessa distribuzione di probabilità ma non sono indipendenti.**

La distribuzione di probabilità della generica X_i è uguale a quella del carattere X nella popolazione.

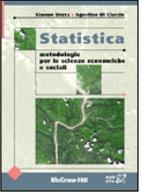
Campionamento casuale stratificato – Popolazione Finita



Nel campionamento casuale stratificato la popolazione viene suddivisa in **strati**. Da ogni strato vengono poi estratti, tramite un campionamento casuale semplice, le unità da inserire nel campione. *Esempio strati: Regioni; età; sesso.*

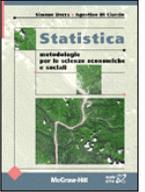


Vantaggi campionamento stratificato



- **Miglioramento della stima**, se gli strati sono stati ben scelti.
- Possibilità di ottenere anche la **stima per le singole sottopopolazioni** o strati.

Campionamento casuale a grappoli e a stadi –Popolazione Finita

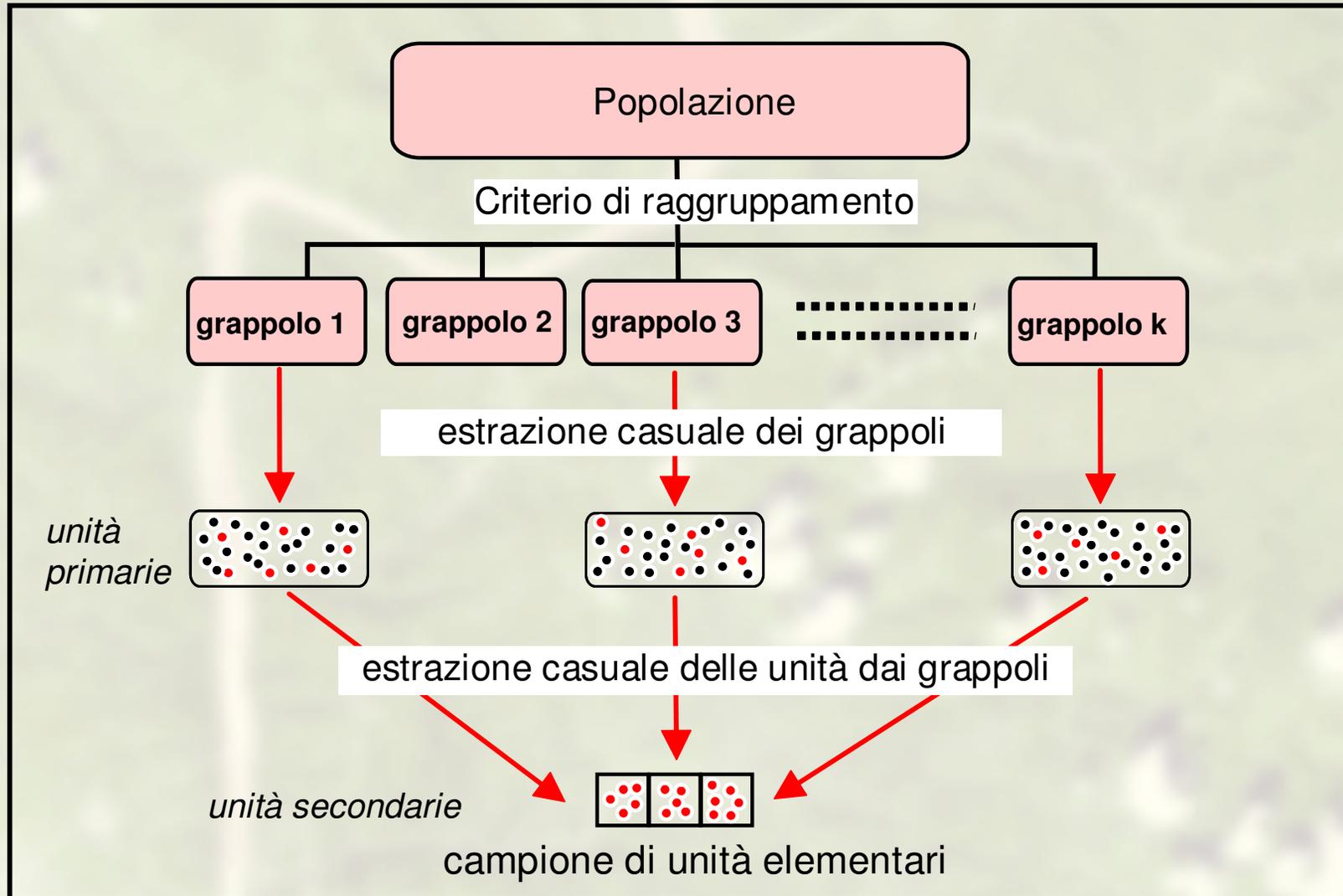
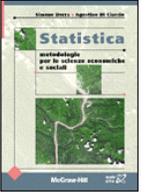


Nel campionamento **casuale a grappoli** la popolazione viene suddivisa in sottoinsiemi detti **grappoli**. Si selezionano, con un'estrazione casuale senza ripetizione, un certo numero di grappoli e si prendono come unità campionarie tutte le unità appartenenti ai grappoli estratti.

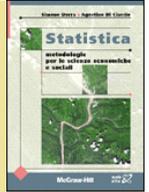
Nel campionamento **casuale a due stadi** la popolazione viene suddivisa in un certo numero di **grappoli**. Al primo stadio si estrae senza ripetizione un certo numero di grappoli. Da ciascuno di questi si estrae con ripetizione (secondo stadio) un certo numero di unità.



11 Campionamento casuale a grappoli e a stadi – Popolazione Finita

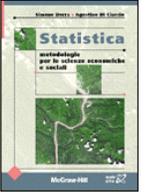


Progettazione del questionario



- ◆ **Concettualizzazione:** si individuano le entità che entrano in gioco, si descrivono le relazioni esistenti fra le entità e si individuano le possibili gerarchie fra le relazioni;
- ◆ **Individuazioni aree e sottoaree**
- ◆ **Contenuti, formulazione e successione logica delle domande**
(*ES. ad imbuto o ad imbuto capovolto*)

Tipi di Intervista



◆ Diretta →

Faccia a faccia

◆ Indiretta →

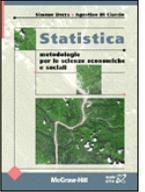
Telefono, Posta, exit poll, CATI

◆ NON Strutturata

◆ SEMI Strutturata

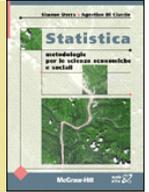
◆ Strutturata

Tipi di Domande



- ◆ *Dirette o Indirette*
- ◆ *A risposta semplice / multipla*
- ◆ *Aperta o chiusa*
- ◆ *Domande Filtro*
- ◆ *Domanda gerarchizzata*

Un esempio



Si vogliono studiare le caratteristiche demografiche e sociali (*caratteri*) della popolazione del Lazio al 1° gennaio 1998 (*collettivo*).

Un semplice questionario potrebbe raccogliere le seguenti informazioni →

Dati estratti dal questionario della terza intervista

Codice ident. intervista	3
Sesso	M
Età in anni compiuti	38
Stato civile	Coniugato
Provincia di residenza	LT
Titolo di studio	Laurea
Professione	Impiegato
Numero componenti della famiglia (oltre l'intervistato)	4
Reddito lordo annuo (in migliaia di euro)	40