

# Lezione 2B

Le fonti statistiche primarie su imprese e famiglie: censimenti, indagini campionarie, registri di imprese, fonti amministrative.

Metodi statistici per l'analisi socio-economica

Docente: Giovanni Giuseppe Ortolani

Corso di Laurea Magistrale in Economia dei settori produttivi e dei mercati internazionali

a.a. 2021/2022



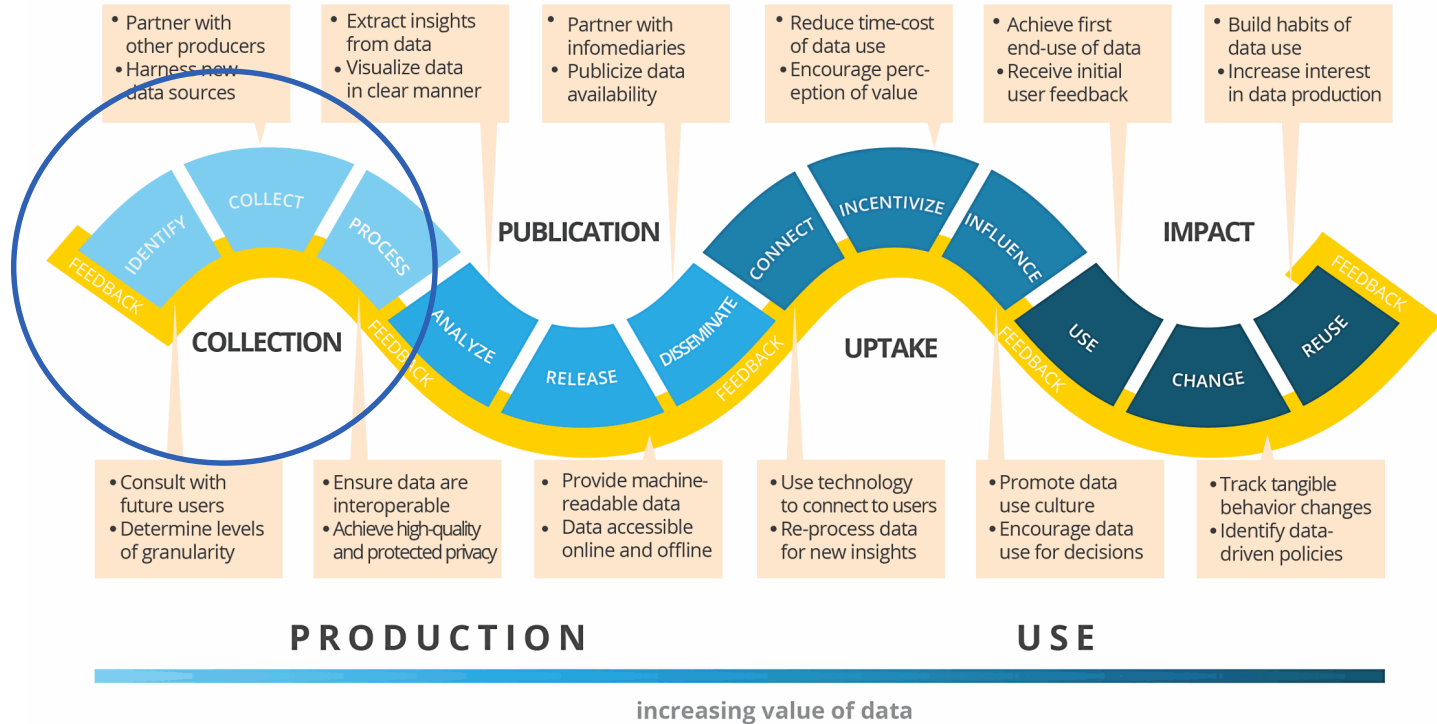
UNIVERSITÀ  
DEGLI STUDI DI TRIESTE

## Le fonti primarie delle statistiche

- Il soddisfacimento della **crescente domanda** di statistiche ufficiali ha indotto un notevole aumento dello sforzo produttivo delle autorità statistiche.
- Queste ultime forniscono al pubblico informazioni con vari **livelli di aggregazione** (dal micro al macro) e su un'ampia **varietà** di fenomeni.
- A tal fine, i produttori di statistiche gestiscono **sistemi informativi statistici** sempre più **complessi** e **integrati**, con numerose **rilevazioni** e **indagini** condotte con dispiego di avanzate tecniche e metodologie.
- Viene perseguita la minimizzazione dei **costi**, anche riguardo agli **oneri dei rispondenti** (*reporting burden*), e la massimizzazione della **qualità**, nel rispetto degli standard internazionali.
- Nel seguito tratteremo le caratteristiche generali delle indagini e presenteremo sinteticamente le **principali rilevazioni statistiche** condotte in Italia.

# **Le indagini statistiche: aspetti generali**

# La catena del valore dei dati: progettazione, produzione e uso delle statistiche



## L'indagine statistica

- Un'indagine statistica è una **raccolta** di dati effettuata seguendo un **processo** produttivo complesso che ha come **obiettivo** la produzione di informazione statistica, cioè la produzione di descrizioni riassuntive di carattere **quantitativo** riguardanti un fenomeno collettivo.
- L'**oggetto** dell'osservazione di ogni fenomeno individuale che costituisce il fenomeno collettivo è detto **unità statistica**.
- Il fenomeno collettivo viene studiato attraverso l'osservazione o la misurazione di una o più **caratteristiche** delle unità statistiche.
- I **mezzi** utilizzati per raccogliere le informazioni possono essere molteplici: **questionari**, strumenti di **misurazione** (ad es. una bilancia che fornisce il peso dei prodotti di un'impresa), **conteggi** e **osservazioni** (ad es. numero di individui entrati alla frontiera di un paese), ma possono anche essere dedotti da documentazione **amministrativa** (ad es. dati dei bilanci delle imprese).

# Fasi di un'indagine statistica

L'indagine statistica è un processo che si articola in diverse fasi, ognuna orientata alla qualità delle statistiche prodotte e diffuse.

## **Le principali fasi sono:**

1. definizione degli obiettivi dell'indagine
2. definizione del disegno di indagine
3. acquisizione dei dati
4. registrazione
5. revisione e validazione
6. elaborazione metodologica
7. presentazione e utilizzazione dei risultati
8. diffusione

# Popolazione

**Popolazione** o *universo di riferimento*: è l'**insieme di unità elementari**, ben identificabili, delle quali interessa studiare una o più **caratteristiche** ad esse **comuni**.

Ciò che fa di un aggregato di elementi una popolazione, infatti, è proprio l'esistenza di una caratteristica presente in tutte le unità, che può diventare oggetto di analisi.

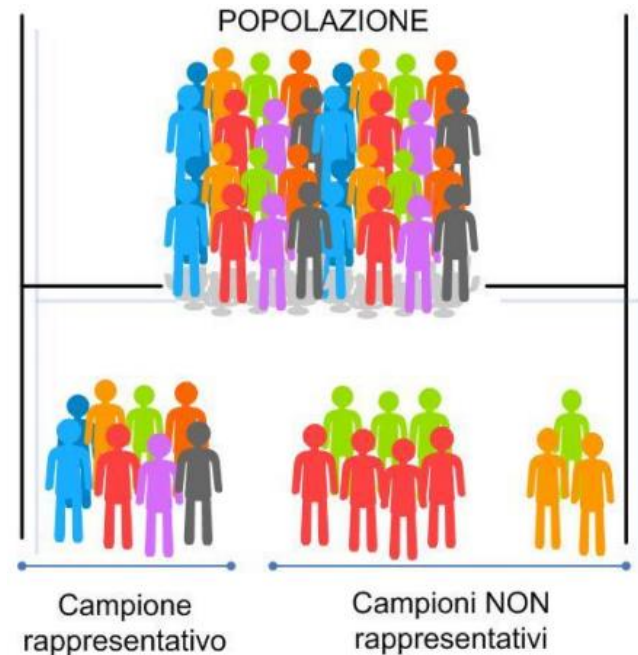
*L'insieme delle **persone** residenti in un territorio è una popolazione se si esamina una qualche loro proprietà, come il **reddito disponibile**.*

*L'insieme delle **aziende** residenti in un paese è una popolazione se si esamina il loro **fatturato**.*



# Campione

- **Campione:** è il **sottoinsieme** di una popolazione oggetto di indagine che viene selezionato per **rappresentarla**.
- I criteri usati per la formazione dei campioni possono essere diversi ma, affinché le informazioni ottenute dal campione possano essere estese alla popolazione di provenienza, è necessario che il campione sia **rappresentativo** ovvero riproduca le caratteristiche più rilevanti ai fini dell'indagine della popolazione di origine.





## Rilevazioni dirette

### **Definizione:**

Una rilevazione si definisce **diretta** quando le osservazioni sono desunte dalle singole unità che costituiscono la popolazione e che vengono contattate nella fase di rilevazione.

### **Esempi:** indagini su:

- condizioni di salute
- tempo libero
- sicurezza dei cittadini
- spese delle famiglie
- forze lavoro
- ...



## Rilevazioni dirette

### ***Tecniche di indagine:***

- interviste dirette o faccia a faccia
- interviste postali
- interviste telefoniche
- autocompilazione
- osservazione diretta

### ***Tipologie di questionario e di intervista:***

- PAPI (*Pencil and paper*)
- CAPI (*Computer Assisted Personal Interview*)
- CATI (*Computer Assisted Telephone Interview*)
- CAWI (*Computer Assisted Web Interview*)

## Rilevazioni indirette

### Definizione:

Una **rilevazione indiretta** (o *indagine statistica basata su dati di fonte amministrativa*) è un'indagine che utilizza **dati amministrativi** con finalità di tipo statistico.

Una parte rilevante delle rilevazioni della statistica ufficiale ha base amministrativa.

### Caratteristiche del dato amministrativo:

- è generato dalla normale attività amministrativa;
- nasce dalla registrazione di uno stato di fatto (ricovero ospedaliero, cambio di residenza, etc.);
- è tendenzialmente esaustivo.



## Rilevazioni dirette e indirette: pro e contro

	<b>Dirette</b>	<b>Indirette</b>
<b>Caratteristica</b>	L'informazione è espressamente raccolta al fine di conoscere un determinato fenomeno	Si utilizzano i dati amministrativi con finalità di tipo statistico
<b>Pro</b>	Rispondenza tra gli obiettivi dell'indagine e l'informazione raccolta	<ul style="list-style-type: none"><li>• Minori costi di rilevazione</li><li>• Minor "disturbo statistico"</li></ul>
<b>Contro</b>	Maggiori costi di rilevazione Maggior "disturbo statistico"	<ul style="list-style-type: none"><li>• La "popolazione amministrativa" può non coincidere con la popolazione statistica</li><li>• I concetti, le definizioni e le classificazioni adottate possono non coincidere con quelli della ricerca statistica</li><li>• Minore qualità dei dati amministrativi rispetto agli obiettivi della ricerca</li><li>• Alcuni fenomeni sociali sfuggono agli atti amministrativi</li></ul>

## INDAGINE TOTALE O CAMPIONARIA?

- Si ricorre **certamente** all'indagine per campione quando la rilevazione completa è impossibile o quando la determinazione delle modalità possedute dalle unità in esame ne comporta la distruzione (Es.: lo studio della durata di un nuovo tipo di contenitore sigillante);
- Si **può** scegliere l'indagine campionaria quando:
  - la popolazione che si studia è numerosa;
  - è insediata su un territorio vasto;
  - deve essere osservata ripetutamente nel tempo.
- Se si considera l'aspetto economico, la rilevazione parziale quasi sempre **ha costi minori** rispetto a quella totale
- Le indagini **sono da preferirsi** quando:
  - si voglia andare in profondità nella ricerca della informazione;
  - serva rapidità nel raccogliere, trattare e pubblicare i risultati.



QUINDI SPESSO SI PONE L'ESIGENZA DI EFFETTUARE UN'INDAGINE CAMPIONARIA

## Indagini totale e campionarie: pro e contro

	<b>Totali</b>	<b>Campionarie</b>
<b>Caratteristica</b>	<ul style="list-style-type: none"><li>• Viene osservata l'intera popolazione oggetto di studio</li></ul>	<ul style="list-style-type: none"><li>• Viene osservato solo un sottoinsieme della popolazione, definito "campione"</li></ul>
<b>Pro</b>	<ul style="list-style-type: none"><li>• Elevato dettaglio di analisi (ad es. territoriale)</li></ul>	<ul style="list-style-type: none"><li>• Riduzione dei costi</li><li>• Riduzione dei tempi</li><li>• Riduzione del carico organizzativo</li></ul>
<b>Contro</b>	<ul style="list-style-type: none"><li>• Costi elevati</li><li>• Possibile minore accuratezza dei dati (più complesso il controllo dei dati raccolti)</li><li>• Ottenimento delle informazioni in tempi più lunghi</li></ul>	<ul style="list-style-type: none"><li>• Comportano imprecisioni dovute all'errore campionario</li><li>• Minore dettaglio nelle disaggregazioni</li></ul>

## I vari tipi di errori nelle indagini

- **Errore campionario**
  - differenza inevitabile tra valori campionari e della popolazione, intrinseco alla «parzialità» delle indagini campionarie. Non presente nelle indagini censuarie.
- **Errori di rappresentatività**
  - liste di estrazione del campione inadeguate; mancate risposte. Possono produrre effetti molto negativi sulla possibilità di estendere i risultati alla popolazione. Le mancate risposte possono avere effetti nefasti anche nelle indagini censuarie.
- **Errori nella raccolta dei dati**
  - implementazione degli strumenti di ricerca. Esempio, scarsa attenzione alla formulazione delle domande. Possibile anche nelle indagini censuarie.
- **Errori nella elaborazione dei dati**
  - difetti nel trattamento informatico-statistico dei dati raccolti, ad esempio errori nella codifica. Possibile anche nelle indagini censuarie.

## Il Censimento

Il Censimento è tipicamente una rilevazione:

### **Totale o universale**

Viene osservata l'intera popolazione di riferimento

### **Generale**

Riguarda i principali aspetti strutturali che caratterizzano le unità di rilevazione

### **Simultanea**

Le notizie devono essere riferite allo stesso istante (la situazione rilevata deve essere quella del momento stabilito come data di riferimento del censimento)

### **Periodica**

Viene ripetuta a intervalli regolari





## Dal Censimento generale al Censimento permanente

Fino al **2011**, l'Istat ha realizzato i **Censimenti generali** attraverso rilevazioni dirette a cadenza **decennale**:

- Censimento Generale della **Popolazione** e delle **Abitazioni**
- Censimento Generale dell'**Agricoltura**
- Censimento Generale dell'**Industria**, dei **Servizi** e delle **Istituzioni Non Profit**
- Censimento delle **Istituzioni Pubbliche**

## Dal Censimento generale al Censimento permanente

Dal 2016 il Censimento è diventato *permanente*.

A differenza del passato, i censimenti permanenti **non** coinvolgono **tutte** le unità della popolazione, ma di volta in volta una **parte di esse**, ovvero dei campioni rappresentativi.

Tuttavia, la restituzione al Paese dei dati ottenuti è di tipo censuario, quindi riferibile all'intero campo d'osservazione.

Questo è possibile grazie all'**integrazione** di fonti amministrative con rilevazioni campionarie.

Il 7° Censimento generale dell'Agricoltura, in corso a inizio 2021, l'ultimo a cadenza decennale, chiude definitivamente la lunga storia dei censimenti generali dell'Istat.

CENSIMENTI PERMANENTI



POPOLAZIONE E ABITAZIONI

IMPRESE

ISTITUZIONI PUBBLICHE

ISTITUZIONI NON PROFIT

AGRICOLTURA

# **Cenni sulle tecniche campionarie**

## IL CAMPIONAMENTO E IL "CASO"

- Si ricorre preferenzialmente a tecniche di campionamento che hanno in comune la **casualità** nella determinazione delle unità che entrano a far parte dell'insieme da osservare;
- La casualità si attribuisce:
  - assegnando a ogni unità della popolazione una **probabilità positiva** di essere selezionata,
  - utilizzando in modo appropriato le **tecniche per la selezione casuale** del campione.
- La **casualità** della selezione si propone essenzialmente come metodo finalizzato all'ottenimento di un campione **rappresentativo** della popolazione con riguardo a ogni variabile di interesse per lo studio in oggetto.
- I campioni **non probabilistici** riflettono, invece, nel bene e nel male, i particolari criteri di selezione di chi li forma.
- Da **rigettare** sono i "campioni" formati **"a casaccio"**, o con **adesione volontaria** dei rispondenti.

## ATTENZIONE ALL'ULTIMA FRASE...

- Da **rigettare** sono i "campioni" formati "**a casaccio**", o con **adesione volontaria** dei rispondenti.

Riflettiamo invece su quanto spesso si formano **generalizzazioni** sulla base:

- affermazioni, «informazioni», «dati», ecc. pubblicati su siti social, gruppi di utenti, blog, o anche sui media tradizionali, di cui non conosciamo la «metodologia di produzione»
- delle proprie, limitate e, spesso, poco obiettive, percezioni ed esperienze dirette personali
- delle limitate e, spesso, poco obiettive, percezioni ed esperienze di persone di nostra diretta conoscenza
  
- «Negli ultimi tempi i prezzi degli affitti a ... sono saliti parecchio. Sarà meglio cercare a ...»
- «I furti nelle case sono molto cresciuti, converrà installare un antifurto»
- «Laurearsi non conviene più. E' meglio cercare subito un lavoro dopo il diploma»

# Non assicurare la casualità di selezione può essere fatale anche con una rilevazione strutturata...

## Il caso delle elezioni presidenziali USA del 1936



- ↳ La rivista Literary Digest condusse un sondaggio cui risposero 2,3 milioni di persone.
- ↳ Utilizzarono questionari spediti a 10 milioni di persone scelte dall'**elenco telefonico**, e dalle liste dei **club**
- ↳ Previsione: Landon wins
  - Landen 57% Roosevelt 43%

# Meglio un piccolo campione ben costruito di uno grande ma distorto...

## Il caso delle elezioni presidenziali USA del 1936

### Risultati



- ↯ Bias nelle risposte:
- ↯ Il campione era formato da elettori che avevano un telefono o erano iscritti a un club
- ↯ Erano ricchi
- ↯ Il voto era correlato allo status socio-economico
- ↯ Lo statistico George Gallup predisse invece la vittoria di Roosevelt.
- ↯ Utilizzò un campione molto più ristretto (50.000 persone).
- ↯ La lezione di Gallup fu fondamentale: dimostrò la minore importanza della **dimensione** del campione statistico rispetto alla sua **composizione**, che dev'essere necessariamente casuale e probabilistica, e quindi al **disegno di campionamento**.

# IL DISEGNO DI CAMPIONAMENTO

E' l'insieme delle **decisioni** prese nel formare il campione.

Si tratta di rispondere alle seguenti domande:

Quale **struttura** deve avere il campione? Quali **probabilità di selezione**?

Quanto **numerioso** deve essere il campione?

## LA STRUTTURA DEL CAMPIONE

- **Semplice**, se la lista della popolazione è unica
- **Complessa**, se sono necessarie più liste

## PROBABILITA' DI SELEZIONE

- **Costanti**, se sono uguali per tutte le unità di una lista;
- **Variabili**, quando ad ogni unità si assegna una probabilità di selezione sua propria e ammonta a 1 per il complesso delle unità listate

## LA DIMENSIONE OTTIMA DEL CAMPIONE

E' determinabile in funzione del disegno di campionamento.



# IL CAMPIONAMENTO CASUALE SEMPLICE

Si consideri una popolazione di  $N$  unità. Si dice **casuale semplice** il campione di  $n$  unità estratte dalla popolazione con uguale probabilità ( $1/N$ ) a ogni passo dell'estrazione.

La probabilità che un'unità compaia nel campione è la somma della probabilità che vi compaia in una qualsiasi delle  $n$  prove, e cioè  $n/N$ .

Un campione casuale si può ottenere con:

- Selezione **casuale con reinserimento**;
- Selezione **casuale senza reinserimento**;
- Selezione **sistematica** (anche detta “pseudo-casuale”) ma solo se le posizioni della lista sono mescolate fino all'ottenimento di un ordinamento assolutamente casuale.

# IL CAMPIONAMENTO CASUALE SEMPLICE: LA STIMA

- Siano  $Y_1, \dots, Y_i, \dots, Y_N$  i valori delle  $N$  unità della popolazione e  $y_1, \dots, y_i, \dots, y_N$  i valori osservati presso le  $n$  unità campionarie.

- MEDIA CAMPIONARIA: 
$$\bar{y} = \frac{\sum_j^n y_j}{n} = \frac{\hat{Y}}{N}$$

- La VARIANZA della media campionaria con un campione con reimmissione è

$$\text{Var}(\bar{y}) = \frac{S_y^2}{n}$$

- $S_y^2$  è la VARIANZA “elementare”, o “della popolazione”, della variabile osservata:

$$S_y^2 = \frac{1}{N-1} \sum_i^N (Y_i - \mu)^2$$

- da stimare con

$$s_y^2 = \frac{1}{n-1} \sum_i^N (y_i - \bar{y})^2$$

# Campionamento casuale semplice: errore e numerosità campionaria

L'**errore campionario** che si è disposti a commettere, indicato in valore assoluto o in percentuale, per un'estrazione di un campione senza ripetizione, come è il caso più comune nella ricerca sociale, ha la forma

$$\partial = \frac{z_a \sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

Da cui la **numerosità campionaria** per la stima di una media può essere scritta come

$$n = \frac{z_a^2 N \sigma^2}{\partial^2 (N-1) + z_a^2 \sigma^2}$$

**Esempio:** per una popolazione di 2000 unità,  $\sigma=36$ , una percentuale di attendibilità del 95% (quindi  $z=1,96$ ), errore massimo del 2% si ha:  $n=(1,96^2 * 2000 * 36^2) / (2^2 * 1999 + 1,96^2 * 36^2) = 767$

# IL CAMPIONAMENTO STRATIFICATO

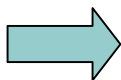
“Stratificare” una popolazione significa ripartirla in sottopopolazioni, dette “STRATI”.

Motivi per stratificare:

- Introdurre sulla selezione il massimo **controllo**, pur mantenendola **casuale**;
- Individuare sottopopolazioni al massimo **omogenee** rispetto alla variabile o alle variabili da rilevare e ricavare così stime più **efficienti** di quelle ottenibili con un campione casuale semplice.

Per stratificare, si devono specificare gli **attributi** in base ai quali va **suddivisa** la popolazione e, correlatamente, il **numero** di **strati**.

Le caratteristiche per la stratificazione devono essere note prima della selezione per ogni unità statistica. Ogni unità statistica appartiene ad un solo strato.



Stratificato è un campione tratto da una popolazione stratificata. Poiché ogni strato è una popolazione in senso proprio, il campione che si seleziona da uno strato è idoneo a rappresentarla. Il campione totale, ottenuto per unione di quelli dei singoli strati, mira a rappresentare l'intera popolazione.

## Esempio di Campionamento Stratificato Proporzionale

Per un sondaggio di opinione tra gli studenti di una università, è necessario estrarre un campione tale da garantire a tutte le Facoltà di essere rappresentate in relazione al numero di iscritti.

Supponiamo che in una Università gli studenti iscritti siano 10.000

Lettere ha 2.000 studenti, Ingegneria 1.500, Matematica 1.000, Medicina 3.000, Economia 2.500

La numerosità campionaria calcolata  $n=1200$ . Ogni Facoltà rappresenta uno strato, occorre dunque estrarre da ciascuno strato un numero di studenti proporzionale al numero degli studenti iscritti in ciascuna Facoltà

<b>es. Lettere</b>	$x : 2000 = 1200 : 10000 \rightarrow x = 2000 * 1200/10000 =$	<b>240</b>
<b>Ingegneria</b>	$x : 1500 = 1200 : 10000 \rightarrow x = 1500 * 1200/10000 =$	<b>180</b>
<b>Matematica</b>	$x : 1000 = 1200 : 10000 \rightarrow x = 1000 * 1200/10000 =$	<b>120</b>
<b>Medicina</b>	$x : 3000 = 1200 : 10000 \rightarrow x = 3000 * 1200/10000 =$	<b>360</b>
<b>Economia</b>	$x : 2500 = 1200 : 10000 \rightarrow x = 2500 * 1200/10000 =$	<b>300</b>
	<b>Totale</b>	<b>1.200</b>

# IL CAMPIONAMENTO STRATIFICATO: LA STIMA

- La MEDIA CAMPIONARIA del singolo strato h

$$\bar{y}_h = \frac{\sum_i^{n_h} y_{hi}}{n_h} \quad [h = 1, \dots, H]$$

- La MEDIA dell'intero campione  $\bar{y} = \sum_h^H W_h \bar{y}_h = \sum_h^H \frac{N_h}{N} \bar{y}_h \quad [h = 1, \dots, H]$

- La VARIANZA della MEDIA dell'intero campione

$$\text{Var}(\bar{y}) = \sum_h^H W_h^2 \text{Var}(\bar{y}_h)$$

... che si stima con

$$\hat{s}_h^2 = \frac{1}{n_h - 1} \sum_j^{n_h} (y_{hj} - \bar{y}_h)^2 \quad [h = 1, \dots, H]$$

Se la somma delle varianze interne dei singoli strati è inferiore alla varianza complessiva del campione ottengo stime più efficienti rispetto al campionamento casuale semplice

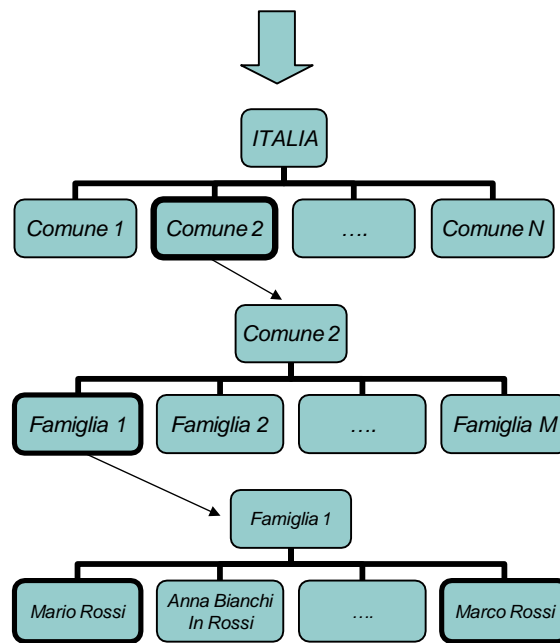
# IL CAMPIONAMENTO SU *PIU' STADI*

Un esempio di campionamento su più stadi è quello condotto dall'ISTAT per svolgere l'indagine sulle forze del lavoro in Italia.

Il campione di unità da intervistare per questa indagine si forma selezionando alcuni Comuni ...

...poi un insieme di famiglie dalle anagrafi dei Comuni...

...infine si seleziona il “grappolo” dei componenti delle famiglie individuate.



## L'EFFETTO DEL DISEGNO DI CAMPIONAMENTO

In un campione casuale **complesso** (diverso da quello **semplice**), l'errore di campionamento può essere espresso in una forma che evidenzia il **guadagno** o la **perdita di precisione** delle stime rispetto all'analogica stima ottenibile con un campione casuale semplice di pari numerosità.

Poiché un campione **stratificato** è generalmente **più efficiente** di un campione casuale semplice, la numerosità campionaria necessaria per ottenere stime di pari efficienza è **inferiore** con un campionamento stratificato.

Il campionamento su **più stadi** è spesso **utilizzato su vasta scala** perché la rilevazione dei dati è "concentrata" sui punti selezionati al 1° stadio e quindi più economica (e quindi, ad es., i rilevatori non devono spostarsi tra troppi comuni, magari per raggiungere solo pochi intervistati).

**Risultano facilitati l'organizzazione del lavoro in loco e l'esecuzione della rilevazione.**

**Risultano minori i costi**



**E' più complessa la stima**

**Vi è rischio di inefficienza**



## Disegni campionari misti

I campionamenti a **più stadi** possono essere **anche stratificati**. Nell'esempio precedente, ad es., le unità di primo stadio potevano essere stratificate per area geografica.

In alcune indagini dell'ISTAT le unità di primo stadio, i comuni, sono invece stratificate per **ampiezza demografica**, nel senso che sono raggruppate secondo le dimensioni in termini di popolazione residente. Così si individuano due tipi di comune: quelli AR (autorappresentativi) e quelli NAR (non autorappresentativi).

I comuni AR sono quelli **grandi** e costituiscono **da soli uno strato**. Essi **devono** essere selezionati perché sono gli unici di quello strato. I comuni NAR **sono riuniti insieme** in alcuni strati. Da questi strati si devono **estrarre** i comuni che possono rappresentare quello strato con un campionamento casuale semplice.

# **Alcune delle principali rilevazioni statistiche italiane**

**Le fonti della contabilità nazionale**

## Le principali rilevazioni statistiche italiane

- Esamineremo fra poco **alcune** delle più importanti **rilevazioni statistiche** condotte dall'Istat e dalla Banca d'Italia.
- Le rilevazioni considerate sono, in primo luogo, funzionali alla produzione di **statistiche settoriali**, cioè riferite a specifici segmenti o aspetti del sistema economico o sociale (ad es. la condizione economica delle imprese, le spese delle famiglie, ecc.). Di seguito trovate l'ampia gamma di rilevazioni Istat, con le relative informazioni metodologiche: [\*\*ELENCO DELLE RILEVAZIONI ISTAT\*\*](#).
- Inoltre, molte di loro concorrono alla definizione di aggregati statistici a un più elevato livello di generalità, in primo luogo i dati «macro» della contabilità nazionale (ad es. i consumi, gli investimenti, il prodotto interno lordo).
- Pertanto, le rilevazioni possono essere considerati come i «**building blocks**» che vanno a comporre i moderni **sistemi informativi statistici**. E' pertanto fondamentale che sia perseguita, per quanto possibile l'**interoperabilità** dei risultati forniti dalle singole indagini, ai fini della loro successiva **integrazione**, e, in ultima analisi, alla composizione della rappresentazione del sistema economico complessivo.

## La contabilità nazionale attinge da una molteplicità di fonti...

- Per costruire le stime degli aggregati che compongono la **contabilità nazionale** si utilizzano dati provenienti da un ampio insieme di **fonti statistiche di base** (o primarie).
- Quelle gestite dall'Istat includono sia le **rilevazioni** presso le **imprese**, le **famiglie** e gli altri agenti economici, sia i **registri statistici**, costruiti attraverso l'integrazione di dati di indagine e archivi amministrativi.
- A queste si affiancano specifici **archivi amministrativi** (quali, ad esempio, quelli dell'INPS, dell'Agenzia delle Entrate, ecc.), nonché informazioni e indagini di altra fonte, pubblica e privata.
- Una parte dei dati provenienti dalle rilevazioni o da altre fonti sono già definiti in modo coerente con le definizioni dei conti nazionali, mentre per quelli definiti in base a principi diversi è necessario operare **correzioni** e **riclassificazioni** che permettono di rispettare i criteri dettati dal SEC2010.

## Le fonti della CN per la stima dell'offerta

Dal lato delle **stime dell'offerta**, le fonti più rilevanti sono quelle che forniscono informazioni **censuarie** sulle **imprese**:

1. l'**archivio** statistico delle imprese attive (**ASIA-Imprese**), e
2. il sistema informativo sui risultati economici delle imprese (di seguito indicato come **Frame-SBS**).

**Frame-SBS** è una base di **microdati** che include i **conti economici** di ciascuna delle **imprese italiane attive** che operano per il mercato (a esclusione di agricoltura e intermediazione finanziaria).

E' costruita attraverso una complessa procedura di **integrazione** di dati di fonte amministrativa, trattati statisticamente e combinati con i dati della **rilevazione dei conti economici delle imprese e del Censimento permanente delle imprese**.

Fonti amministrative: bilanci depositati, archivio Inps, studi di settore e archivi fiscali.

## Il Registro delle imprese ASIA

- Il registro Asia è un fondamentale archivio sulle **imprese industriali, commerciali e dei servizi alle imprese e alle famiglie**.
- Oltre a costituire la base informativa per le analisi sull'evoluzione della struttura delle imprese italiane e sulla loro demografia, il registro Asia rappresenta l'**universo di riferimento delle indagini sulle imprese** condotte dall'Istat e viene utilizzato, ad es., per le stime di contabilità nazionale
- Il registro fornisce informazioni sulle imprese integrando quelle desumibili da **fonti amministrative**, gestite da enti pubblici o da società private (v. slide seguenti), e quelle da **fonti statistiche**.

## Il Registro delle imprese ASIA

Le **variabili** comprese nel registro sono classificate secondo tre tipologie:

- variabili **identificative** (ragione sociale, indirizzo e altri caratteri per l'esatta individuazione dell'unità sul territorio)
- variabili di **stratificazione** (attività economica dell'impresa classificata secondo la **classificazione Ateco**, forma giuridica, dimensione dell'impresa, in termini di addetti indipendenti e dipendenti medi annui e di fatturato, appartenenza a gruppo)
- variabili **demografiche** (data di nascita e cessazione dell'impresa, data di eventi quali scorpori, fusioni o procedure concorsuali, fallimenti, liquidazioni, ecc.).

## Il Registro delle imprese ASIA

Le principali **fonti amministrative** di ASIA sono:

- Agenzia delle entrate
- il Registro delle imprese delle Camere di commercio, industria, artigianato e agricoltura
- INPS
- INAIL
- l'archivio delle utenze telefoniche
- l'archivio dei Bilanci consolidati e di esercizio
- l'archivio degli intermediari finanziari gestito dalla Banca d'Italia
- l'archivio delle società di assicurazione gestito dall'IVASS



## La rilevazione sui conti economici delle imprese

La rilevazione dei **conti economici delle imprese** è rivolta a tutte le imprese italiane che operano nei settori dell'industria, delle costruzioni, del commercio e dei servizi, con l'esclusione di alcune divisioni dell'intermediazione monetaria e finanziaria, delle assicurazioni e dei servizi domestici.

Il questionario, raccolto via web, rileva i **dati economico-finanziari e patrimoniali** delle imprese, classificate secondo l'attività economica prevalente. Rileva, inoltre, i dati sull'**occupazione**, sugli **investimenti** e sui **costi del personale**.

La rilevazione adotta una strategia **campionaria** per le imprese con **meno di 250 addetti** (PMI: 75.715 imprese coinvolte) mentre è **totale** per le imprese **con 250 addetti ed oltre** (SCI: 3.667 imprese coinvolte).

## Il Censimento permanente delle imprese

Il primo **Censimento permanente delle imprese** si è svolto nel 2019. L'obiettivo della rilevazione è quello di aggiornare il quadro sulla **struttura** e sulla **competitività** delle **imprese** per cogliere i cambiamenti del sistema produttivo nazionale.

Le informazioni sono desunte dall'integrazione fra i registri statistici e le indagini economiche correnti, cui si aggiungono nuove rilevazioni dirette mirate all'analisi di aspetti emergenti di grande rilevanza per la competitività, la sostenibilità sociale e ambientale delle imprese, la crescita economica del Paese.

A differenza dei censimenti tradizionali, il Censimento permanente delle imprese è di tipo **campionario**, mentre la restituzione dei dati ottenuti è di tipo censuario. La rilevazione, svolta per la prima volta con **cadenza triennale**, non più decennale, consente un rilascio di informazioni continue e tempestive. Il campione coinvolto, che cambia ogni tre anni, è di circa **280.000 imprese** che impiegano **3 o più addetti**.

I principali vantaggi introdotti dal **nuovo processo censuario** sono l'aumento della quantità e della qualità dell'offerta informativa, il contenimento dell'onere statistico sugli operatori economici, la riduzione dei costi complessivi della produzione statistica ufficiale.

La raccolta dei dati è *paperless*, condotta esclusivamente attraverso un questionario online, riservato alle sole imprese campione.

[Censimento permanente delle imprese 2019: i primi risultati](#)



## Le fonti della CN per l'intermediazione finanziaria e il settore primario

- Le informazioni sui conti economici delle imprese che operano nell'intermediazione finanziaria provengono dalle istituzioni preposte alla loro vigilanza, ossia la **Banca d'Italia**, l'Istituto per la vigilanza sulle assicurazioni (**IVASS**) e la Commissione di vigilanza sui fondi pensione (**COVIP**).
- Le stime relative al **settore primario** (agricoltura, silvicoltura e pesca) sono basate su molteplici fonti specifiche, oltre che su quelle derivanti dal Censimento dell'agricoltura, l'indagine sui risultati economici delle aziende agricole (REA) e la rilevazione sulla struttura e produzioni delle aziende agricole (SPA).

## Le fonti della CN per il non-profit e per la PA

- Per il **settore non-profit** si dispone di un data base costruito a partire dal **Registro statistico delle istituzioni non-profit** private e su quello delle istituzioni pubbliche, integrato con informazioni che derivano dalla rilevazione campionaria sul settore e da una serie di archivi amministrativi (UNIEMENS INPS, IRAP, UNICO, IVA).
- Per le **Amministrazioni pubbliche** si utilizza un approccio articolato, basato sulla raccolta e il trattamento dei dati di bilancio dell'insieme delle unità istituzionali incluse in tale settore, a livello centrale e locale. Per alcune tipologie di amministrazioni (ad esempio lo Stato, i Comuni, le Regioni, le aziende socio sanitarie territoriali, le Università, gli Enti di previdenza) la raccolta avviene tramite canali centralizzati.

# **Alcune delle principali rilevazioni statistiche italiane**

**Indagini sulle famiglie, le forze lavoro, i prezzi al consumo e  
l'intermediazione finanziaria**

## Il Censimento permanente Istat su popolazione e abitazioni

- Dal 2018 l'Istat ha avviato il **Censimento permanente della popolazione e delle abitazioni**, la rilevazione che consente, con cadenza annuale e non più decennale, di rilasciare informazioni continue e tempestive sulle principali **caratteristiche socio-economiche della popolazione** dimorante abitualmente in Italia.
- Il Censimento permanente non coinvolge più tutte le famiglie nello stesso momento, ma solo un campione selezionato di esse, e grazie all'integrazione dei dati raccolti con le rilevazioni campionarie con quelli provenienti dalle fonti amministrative consente di restituire informazioni rappresentative dell'intera popolazione.
- Le prime due rilevazioni del 2018 e 2019 hanno visto il coinvolgimento per anno di un campione di circa **1.400.000 famiglie** in oltre **2.800 comuni**.
- Nel 2020, a fronte dell'emergenza Covid-19, l'Istat ha modificato parzialmente il disegno censuario. Nel 2021 è prevista la ripresa delle rilevazioni sul campo.

[Prima diffusione dei dati definitivi 2018 e 2019](#)

# L'indagine Istat sulle spese delle famiglie

La **rilevazione Istat sulle spese delle famiglie** rileva i comportamenti di spesa e i movimenti turistici delle famiglie residenti in Italia. Rappresenta inoltre la base informativa per le stime ufficiali di **povertà** relativa e assoluta in Italia.

L'indagine è di tipo **campionario** e coinvolge, ogni anno, circa **32.000 famiglie residenti** in circa **540 comuni** italiani. Le famiglie vengono estratte casualmente utilizzando la Lista Anagrafica Comunale (LAC) oppure attraverso l'Anagrafe Nazionale della Popolazione Residente (ANPR).

La raccolta dei dati avviene attraverso un **questionario CAPI**, somministrato in due momenti distinti da un rilevatore, e un **diario cartaceo autocompilato** dalla famiglia.

Sul diario cartaceo le famiglie devono invece annotare quotidianamente, **per 14 giorni**: le spese sostenute da tutti i componenti per l'acquisto di alimentari e bevande, pasti e consumazioni fuori casa, cibi pronti, prodotti presso distributori automatici e beni e servizi non alimentari di largo consumo; le quantità consumate o regalate (da tutti i componenti) di beni autoprodotti; i luoghi di acquisto più frequente di alcuni beni, alimentari e non.

[LE SPESE PER I CONSUMI DELLE FAMIGLIE - 2019](#)

# L'indagine Banca d'Italia sui bilanci delle famiglie italiane

L'indagine nasce negli anni '60 con l'obiettivo di raccogliere informazioni sui **redditi** e i **risparmi** delle famiglie italiane. Nel corso degli anni l'oggetto della rilevazione si è andato estendendo per includere anche la **ricchezza** e altri aspetti inerenti i comportamenti economici e finanziari delle famiglie, come ad esempio l'**uso dei mezzi di pagamento**.

Nelle ultime indagini il campione è formato da circa **8.000 famiglie (20.000 individui)**, distribuite in circa **300 comuni italiani**.

I dati dell'indagine sui bilanci delle famiglie italiane sono distribuiti sul sito della Banca e all'interno di basi dati internazionali, opportunamente armonizzate. In particolare, dal 2010 l'indagine fornisce i dati per l'Italia per l'indagine dei bilanci delle famiglie dell'Eurosistema coordinata dalla Banca Centrale Europea.

[Indagine sui bilanci delle famiglie italiane nell'anno 2016](#)





## La rilevazione Istat sulle forze di lavoro

La **rilevazione sulle forze di lavoro** rappresenta la principale fonte di informazione statistica sul mercato del lavoro italiano. Le informazioni rilevate costituiscono la base sulla quale vengono derivate le stime ufficiali degli **occupati** e dei **disoccupati**, nonché le informazioni sui principali aggregati dell'offerta di lavoro – professione, settore di attività economica, ore lavorate, tipologia e durata dei contratti, formazione.

Si tratta di un'indagine **campionaria** e ogni anno vengono intervistate oltre 250.000 famiglie residenti in Italia (per un totale di **600.000 individui**) distribuite in circa **1.400 comuni italiani**. Le famiglie da intervistare vengono casualmente estratte dalle liste anagrafiche comunali secondo una strategia di campionamento volta a costruire un campione statisticamente rappresentativo della popolazione residente in Italia relativamente alle variabili oggetto d'indagine.

L'indagine prevede che la prima intervista venga effettuata di norma presso il domicilio della famiglia con tecnica CAPI. Le interviste successive alla prima vengono effettuate di norma telefonicamente.

## Le indagini sui prezzi al consumo

L'**indagine mensile sui prezzi al consumo** registra le variazioni nel tempo dei prezzi di un insieme di prodotti (paniere) rappresentativo di tutti i beni e servizi destinati al **consumo finale delle famiglie**, acquistabili attraverso transazioni monetarie.

I dati che concorrono alla costruzione degli indici mensili dei prezzi al consumo sono raccolti attraverso una pluralità di fonti:

- la *Rilevazione territoriale dei prezzi al consumo*, condotta da rilevatori degli Uffici Comunali di Statistica, che registrano i dati su tablet connessi online. Unità di rilevazione: punti vendita, imprese e istituzioni; famiglie per i canoni d'affitto.
- la *Rilevazione centralizzata dei prezzi al consumo*, condotta dall'Istat direttamente o attraverso la collaborazione con grandi fornitori di dati
- gli *Scanner data* provenienti dalla Grande Distribuzione organizzata
- la *Fonte amministrativa*.

Nel complesso sono rilevate circa 500.000 quotazioni al mese.

La rilevazione contribuisce a definire i vari tipi di **indici dei prezzi al consumo** (per l'intera collettività - **NIC**, per le famiglie di operai e impiegati – **FOI**, e armonizzato a livello comunitario - **IPCA**), di cui parleremo in successive lezioni.

# Le rilevazioni della Banca d'Italia sugli intermediari creditizi e finanziari

L'evoluzione del sistema di riferimento dell'azione della Banca d'Italia ha reso più stringente la necessità di effettuare con sempre maggiore tempestività l'analisi dei **fenomeni monetari, finanziari ed economici**, mediante l'acquisizione e l'elaborazione dei dati necessari.

Al patrimonio informativo utilizzato dalla Banca d'Italia concorrono in misura rilevante i dati segnalati dagli **intermediari creditizi e finanziari**. Le norme vigenti pongono infatti a carico degli intermediari creditizi e finanziari obblighi segnaletici nei confronti della Banca.

I flussi segnaletici trasmessi, su **base censuaria**, forniscono una rappresentazione dell'impresa bancaria o finanziaria che consente di apprezzare la **situazione patrimoniale ed economica** ed i **rischi** che ne caratterizzano la gestione,

La **matrice dei conti** è lo strumento fondamentale attraverso il quale la Banca raccoglie dati statistici e di vigilanza sull'operatività degli intermediari. E' uno schema segnaletico che contiene le informazioni necessarie a soddisfare flessibilmente le esigenze conoscitive delle diverse funzioni dell'Istituto.