

Le indagini statistiche

Domenico De Stefano

a.a. 2018/2019

Indice

- 1 Le indagini statistiche: introduzione
- 2 Esempi di indagine statistica
- 3 I censimenti
- 4 Le poverty surveys
- 5 Le indagini multiscopo
- 6 Rilevazione sulle forze lavoro

Dal libro di Jelke Bethlehem, “Applied survey methodology”:

We live in an information society. There is an ever-growing demand for statistical information about the economic, social, political, and cultural shape of countries. Such information will enable policy makers and others to make informed decisions for a better future. Sometimes, such statistical information can be retrieved from existing sources, for example, administrative records. More often, there is a lack of such sources. Then, a survey is a powerful instrument to collect new statistical information.

Premesse dal corso di statistica di base

La statistica è la scienza che si occupa di raccogliere dati e trarre da essi informazioni/conoscenza.

I dati possono aiutare a capire i fenomeni, occorre però raccogliere i dati che servono, e farlo bene; i dati vanno poi esaminati in modo da isolare e evidenziare le informazioni che si cercano.

La statistica si occupa, a partire da una domanda su un fenomeno, di stabilire quali dati possano essere usati per rispondere a quella domanda, e, se i dati non sono già disponibili, come debbano essere raccolti. Segue poi la fase in cui si analizzano i dati per estrarre le informazioni cercate.

Premesse dal corso di statistica di base (2)

Sommariamente le fasi di un'analisi statistica sono

- *formulare una domanda*, tradurre un'esigenza conoscitiva in modo che sia suscettibile di una risposta in termini statistici;
- *individuare o raccogliere i dati*, è un ambito vasto, che va sotto il nome di disegno dell'indagine e campionamento; i fondamenti poggiano sul calcolo delle probabilità;
- *organizzare e guardare i dati*, dalla massa di dati così com'è non è immediato estrarre le informazioni che servono, si possono però sintetizzare opportunamente e/o rappresentarli graficamente, in funzione delle informazioni cercate;

Premesse dal corso di statistica di base (3)

- *analizzare i dati e modellare* la fase di analisi dei dati di solito prevede che si formuli (specifichi) un modello statistico che possa spiegare i dati osservati in base alle ipotesi fatte sul fenomeno, si stima il modello usando i dati. Il modello potrà essere usato per confermare o smentire delle (eventuali) ipotesi fatte sul fenomeno e/o per previsioni su future istanze.

Indice

- 1 Le indagini statistiche: introduzione
- 2 Esempi di indagine statistica**
- 3 I censimenti
- 4 Le poverty surveys
- 5 Le indagini multiscopo
- 6 Rilevazione sulle forze lavoro

Alcuni esempi di indagine statistica

Il loro vasto ambito di applicazione rende quasi impossibile la definizione precisa di cos'è un'indagine statistica.

A grandi linee una **indagine statistica** è lo strumento mediante il quale si **acquisiscono informazioni su uno o più fenomeni attinenti ad una popolazione**.

Alcuni esempi di indagine statistica (in campo sociale) sono:

- il Censimento della popolazione italiana con cadenza decennale a partire dal 1861 ad oggi, con l'eccezione del censimento del 1936 (ultima rilevazione: 9 ottobre 2011)
- il Sistema di indagini multiscopo e la Rilevazione Forze Lavoro condotte dall'ISTAT
- la General Social Survey (GSS) condotta dal 1972 ad oggi negli Stati Uniti dal NORC (National Opinion Research Center) at the University of Chicago
- le ricerche di mercato condotte tipicamente da istituti di ricerca privati
- i vari sondaggi di opinione e politici (diffusi ormai in tutti i paesi industrializzati e nati negli Stati Uniti del primo dopoguerra)

Alcuni esempi di indagine statistica (2)

- innumerevoli indagini condotte da ricercatori, Università, istituti di ricerca pubblici e privati, Istituzioni pubbliche e governi

Indice

- 1 Le indagini statistiche: introduzione
- 2 Esempi di indagine statistica
- 3 I censimenti**
- 4 Le poverty surveys
- 5 Le indagini multiscopo
- 6 Rilevazione sulle forze lavoro



Domenico De Stefano

- ## Analisi dei dati - Prima parte

I censimenti

- Il primo censimento ufficiale in Italia si è svolto nel 1861, anno dell'unità d'Italia
- L'ISTAT si occupa della gestione dei censimenti della popolazione da quello del 1931
- Il censimento della popolazione del 2011 è il 15°
- Obiettivo principale del Censimento della popolazione è rilevare, *per ciascun comune*, le persone *dimoranti abitualmente* (anche se temporaneamente assenti). La popolazione censita come dimorante abitualmente, ossia residente nel territorio del Comune, costituisce la **popolazione legale**
- La **popolazione presente** è invece quella effettivamente rilevata, cioè fisicamente sul luogo al momento della rilevazione
- Il censimento della popolazione non è l'unico. L'ISTAT realizza anche:
 - ▶ il Censimento dell'Agricoltura (cadenza decennale, quello del 2010 è il 6°)
 - ▶ il Censimento dell'Industria e dei Servizi (cadenza decennale, quello del 2011 è il 9°)

I censimenti

- Il censimento nell'*accezione moderna* è una rilevazione totale della popolazione, delle abitazioni, delle imprese di un paese in un determinato momento nel tempo.
- Il livello di dettaglio delle informazioni desunte dal censimento non è deducibile da altri tipi di indagine e costituisce un supporto alle decisioni di ogni livello di governo (Stato, Regione, Provincia, Comune)
- I contenuti informativi del censimento vengono attualmente definiti sulla base dei Regolamenti dell'Unione Europea che forniscono l'elenco dei *core topics* (variabili che gli Stati membri sono obbligati a inserire nella rilevazione)
- Lo strumento di rilevazione principale è il **questionario** (nel 2011 per la prima volta è possibile compilare il questionario online)
- I dati del censimento sono disponibili sul **datawarehouse** dell' ISTAT
<http://dati.istat.it/>

I censimenti

Censimento			Popolazione residenti	Var %	Note
num.	anno	data rilevamento			
1°	1861	31 dicembre	22.176.477	-	Il primo censimento della popolazione viene effettuato nell'anno dell'unità d'Italia.
2°	1871	31 dicembre	27.299.883	+23,1%	Come nel precedente censimento, l'unità di rilevazione basata sul concetto di "famiglia" non prevede la distinzione tra famiglie e convivenze.
3°	1881	31 dicembre	28.951.546	+6,1%	Viene adottato il metodo di rilevazione della popolazione residente, ne fanno parte i presenti con dimora abituale e gli assenti temporanei.
4°	1901	10 febbraio	32.963.316	+13,9%	La data di riferimento del censimento viene spostata a febbraio. Vengono introdotte schede individuali per ogni componente della famiglia.
5°	1911	10 giugno	35.841.563	+8,7%	Per la prima volta viene previsto il limite di età di 10 anni per rispondere alle domande sul lavoro.
6°	1921	1 dicembre	39.396.757	+9,9%	L'ultimo censimento gestito dai comuni gravati anche delle spese di rilevazione. In seguito le indagini statistiche verranno affidate all'Istat.
7°	1931	21 aprile	41.043.489	+4,2%	Per la prima volta i dati raccolti vengono elaborati con macchine perforatrici utilizzando due tabulatori Hollerith a schede.
8°	1936	21 aprile	42.398.489	+3,3%	Il primo ed unico censimento effettuato con periodicità quinquennale.

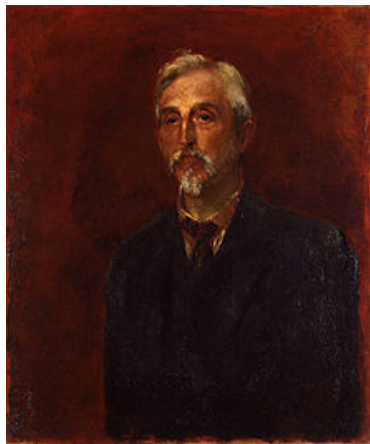
I censimenti

9°	1951	4 novembre	47.515.537	+12,1%	Il primo censimento della popolazione a cui è stato abbinato anche quello delle abitazioni.
10°	1961	15 ottobre	50.623.569	+6,5%	Il questionario viene diviso in sezioni. Per la raccolta dei dati si utilizzano elaboratori di seconda generazione con l'applicazione del transistor e l'introduzione dei nastri magnetici.
11°	1971	24 ottobre	54.136.547	+6,9%	Il primo censimento di rilevazione dei gruppi linguistici di Trieste e Bolzano con questionario tradotto anche in lingua tedesca.
12°	1981	25 ottobre	56.556.911	+4,5%	Viene migliorata l'informazione statistica attraverso indagini pilota che testano l'affidabilità del questionario e l'attendibilità dei risultati.
13°	1991	20 ottobre	56.778.031	+0,4%	Il questionario viene tradotto in sei lingue oltre all'italiano ed è corredato di un "foglio individuale per straniero non residente in Italia".
14°	2001	21 ottobre	56.995.744	+0,4%	Lo sviluppo della telematica consente l'attivazione del primo sito web dedicato al Censimento e la diffusione dei risultati online.
15°	2011	9 ottobre	59.433.744	+4,3%	Il Censimento 2011 è il primo censimento online con i questionari compilati anche via web.

Indice

- 1 Le indagini statistiche: introduzione
- 2 Esempi di indagine statistica
- 3 I censimenti
- 4 Le poverty surveys**
- 5 Le indagini multiscopo
- 6 Rilevazione sulle forze lavoro

Le poverty surveys dell'Età Vittoriana



Charles Booth, by George F. Watts, National Portrait Gallery, London

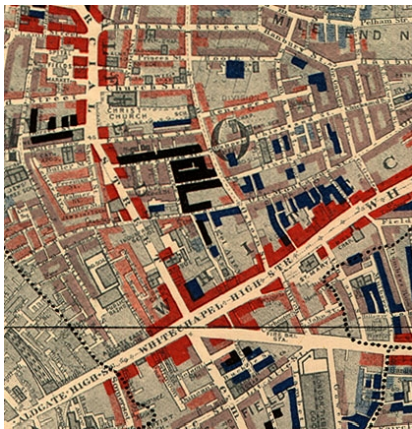
- Le prime indagini statistiche (non censuarie) in ambito sociale vennero realizzate in Inghilterra durante l'Età Vittoriana (1832-1902) e riguardavano principalmente studi sulle condizioni di povertà
- "Life and Labour of the People in London" (1886-1903) di **Charles Booth** fu sicuramente tra le prime e più influenti surveys dell'epoca (anche se non poteva ancora essere considerata un'indagine statistica di tipo moderno)
- Booth era un ricco armatore fortemente turbato dalle condizioni di povertà della classe lavoratrice (working class):

Le poverty surveys dell'Età Vittoriana

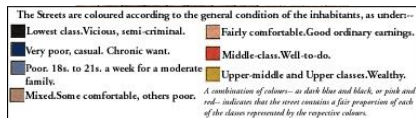


- Booth voleva descrivere le condizioni di vita nell'East End di Londra
- era però complicato collezionare informazioni su questa vasta "popolazione"
- la sua idea fu quella del c.d. metodo del "**wholesale interviewing**", che consisteva nell'intervistare coloro che (all'epoca) avevano informazioni sulle condizioni socio-economiche delle famiglie, cioè i *funzionari scolastici* (gli *attendance officer*)
- i funzionari scolastici avrebbero fornito dati sulle condizioni dei genitori dei vari alunni
- (Il suo "campione" era distorto)

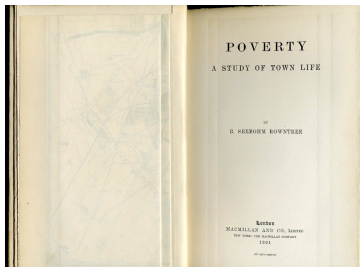
Le poverty surveys dell'Età Vittoriana



- Sulla base di tali dati Booth identificò 8 classi sociali (4 delle quali rappresentavano diversi gradi di povertà)
- Uno dei risultati fu la realizzazione di una mappa della composizione sociale dell'East End
- fu un lavoro pionieristico che combinava metodi statistici e metodi qualitativi (come l'etnografia) e, al di là dei limiti metodologici, ha dato il via ad altre importanti indagini statistiche



Gli studi successivi: Poverty, A Study of Town Life

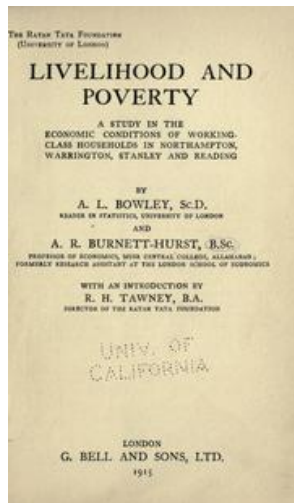


Collezione dell'Università di Harvard. Digitalizzato da Google Inc.

- Dopo l'opera di Booth seguirono numerosi studi analoghi. Tra i più importanti "Poverty, A Study of Town Life" di **Seebohm Rowntree** (1871-1954), uno studio sulla povertà nella città di York
- Rowntree realizza un'indagine statistica che può essere considerata il primo studio scientifico sulla povertà. Introdusse importanti innovazioni metodologiche
- La *prima innovazione* fu il metodo delle interviste dirette. Anzichè raccogliere dati da persone "informate", intervistò direttamente le famiglie ($n=46000$)
- La *seconda* fu il metodo di stima applicato per derivare la soglia di povertà
- Sulla base dei dati Rowntree stimò che circa il 30% della popolazione di York poteva essere considerata povera ▶

Gli studi successivi: Livelihood and Poverty

- Un altro studio fondamentale fu “Livelihood and Poverty” di **Arthur Lyon Bowley** (1869-1957), statistico economico e matematico
- Prosegue nello studio della povertà sulla falsariga di Booth e Rowntree.
- La più importante innovazione fu il primo sistematico utilizzo del **campionamento** nelle indagini statistiche in campo sociale.



Collezione dell'Università della California.
Digitalizzato da Google.

Indice

- 1 Le indagini statistiche: introduzione
- 2 Esempi di indagine statistica
- 3 I censimenti
- 4 Le poverty surveys
- 5 Le indagini multiscopo**
- 6 Rilevazione sulle forze lavoro

Il sistema delle indagini multiscopo

- È un sistema integrato di indagini statistiche campionarie
- Il Sistema di indagini multiscopo nasce nel 1993 con il duplice obiettivo di:
 - ▶ sistematizzare le precedenti esperienze di analisi (anni '80) delle problematiche sociali
 - ▶ soddisfare la domanda di informazione statistica in campo sociale
- Attualmente è composto da 7 indagini in campo sociale:
 - ▶ l'indagine annuale *Aspetti della vita quotidiana*
 - ▶ un'indagine trimestrale su *Viaggi e vacanze*
 - ▶ cinque indagini tematiche, effettuate a rotazione in un arco di tempo di cinque anni: *Condizioni di salute e ricorso ai servizi sanitari, I cittadini e il tempo libero, Sicurezza dei cittadini, Famiglia e soggetti sociali, Uso del tempo*
- l'indagine Aspetti della vita quotidiana fornisce ogni anno un set di indicatori sociali sulle principali aree tematiche, aree che vengono poi sviluppate e approfondite nelle indagini ad hoc a cadenza quinquennale (anche in questo caso i dati sono disponibili sul **datawarehouse** dell' ISTAT
<http://dati.istat.it/>)

Il sistema delle indagini multiscopo

Schema riassuntivo e cadenza temporale (fino al 2005)

TITOLI DELLE INDAGINI	Cadenza	Anni di rilevazione
Aspetti della vita quotidiana	Annuale	1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2005 (a)
Condizioni di salute e ricorso ai servizi sanitari	Quinquennale	1994, 2000 (b), 2005 (c)
I cittadini e il tempo libero (d)	Quinquennale	1995, 2000, 2006 (e)
Sicurezza dei cittadini	Quinquennale	1998 (f), 2002
Famiglia e soggetti sociali	Quinquennale	1998, 2003
Uso del tempo	Quinquennale	1988, 1989 (g), 2003 (h)
Viaggi e vacanze	Trimestrale	1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005

- (a) L'indagine prevista a novembre del 2004 è stata posticipata a febbraio 2005.
 (b) L'indagine è stata effettuata tra dicembre 1999 e settembre 2000.
 (c) L'indagine è stata effettuata tra dicembre 2004 e settembre 2005.
 (d) Nel 1995 l'indagine era denominata Tempo libero e cultura.
 (e) L'indagine prevista a dicembre del 2005 è stata posticipata ad aprile 2006.
 (f) L'indagine è stata effettuata tra ottobre 1997 e gennaio 1998.
 (g) L'indagine è stata effettuata tra giugno 1988 e maggio 1989.
 (h) L'indagine è stata effettuata tra aprile 2002 a marzo 2003.

Aspetti della vita quotidiana

- L'indagine annuale Aspetti della vita quotidiana rappresenta, a partire dal 1993 l'elemento portante ed unificante del sistema d'indagini multiscopo
- utilizzata per costruire le principali serie storiche su molteplici aspetti sociali, senza però entrare nel dettaglio delle singole aree tematiche, approfondite, nelle indagini quinquennali
- sostituisce la precedente indagine Multiscopo sulle famiglie (1987-1991)
 - ▶ Rispetto alla Multiscopo sulle famiglie, si ha un'estensione dei contenuti informativi con quesiti e temi utili anche ai fini della formulazione e del monitoraggio delle politiche sociali
- Oggetto dell'indagine è la vita quotidiana, articolata in vari aspetti:
 - ▶ scuola
 - ▶ lavoro
 - ▶ vita familiare e di relazione
 - ▶ abitazione e zona in cui si vive
 - ▶ tempo libero
 - ▶ partecipazione politica e sociale
 - ▶ salute e stili di vita
 - ▶ rapporto con i servizi

La rilevazione nelle indagini multiscopo

Sono 2 i metodi di rilevazione adottati nelle indagini multiscopo:

- **Intervista diretta** (metodo PAPI, Pencil And Paper Interview)

- ▶ l'indagine annuale sugli aspetti della vita quotidiana, l'indagine sull'uso del tempo, l'indagine sulle condizioni di salute, l'indagine su famiglia e soggetti sociali e l'indagine su tempo libero e cultura

- **Intervista telefonica**

- ▶ l'indagine sulla sicurezza dei cittadini e l'indagine su viaggi e vacanze

Esempio struttura questionario (Aspetti della vita quotidiana)

Il questionario (se ne somministrano 2, uno con intervista diretta e l'altro autocompilato), dell'indagine Aspetti della vita quotidiana è organizzato in quattro macroaree:

- A) Un primo gruppo di informazioni descrive gli ambiti di vita più vicini alla persona: la famiglia, l'abitazione e la zona in cui si vive.
- B) Un secondo insieme di informazioni riguarda le condizioni di salute e gli stili di vita in un'ottica di buona salute.
- C) Un terzo gruppo di informazioni è relativo alla cultura, alla socialità e alle attività del tempo libero
- D) Il quarto set di indicatori monitora la complessa interazione tra i cittadini ed i servizi di pubblica utilità, ne misura l'utilizzo e la soddisfazione

Campione e Universo (Indagini dirette)

- Le indagini del Sistema Multiscopo considerano come popolazione obiettivo – intesa come l'insieme delle unità statistiche oggetto di investigazione – le **famiglie residenti** in Italia e gli **individui ad esse appartenenti**
 - ▶ Per le famiglie, si adotta la definizione di **famiglia di fatto**, definita come insieme di persone coabitanti e legate da vincoli di matrimonio, parentela, affinità, adozione, tutela o affettivi
 - ▶ Per gli individui, la definizione della popolazione obiettivo dipende dalla singola indagine. Ad es. nell'indagine Sicurezza dei cittadini essa è costituita da tutti gli individui di età pari o superiore a 14 anni
- Il piano di campionamento della Multiscopo Aspetti della vita quotidiana (adottato anche per le altre indagini la cui rilevazione è diretta) è complesso: è a 2 stadi e doppio livello di stratificazione per comune:
 - ▶ PRIMO STADIO stratificazione in base a variabili di tipo territoriale e una successiva in base alla dimensione in termini di popolazione
 - ▶ Da ogni strato i -mo si selezionano n_i comuni
 - ▶ SECONDO STADIO: all'interno dei comuni si selezionano le famiglie
 - ▶ gli individui si selezionano mediante campionamento a grappolo (tutti i componenti di fatto appartenenti alla famiglia selezionata)

Campione e Universo (Indagini dirette)

Il prospetto illustra la numerosità campionaria e il corrispondente universo (popolazione) per l'indagine Aspetti della vita quotidiana.

NB: I dati dell'Universo sono espressi in migliaia

Prospetto 1 – Distribuzione regionale dei comuni, delle famiglie e degli individui nell'universo e nel campione – Anno 2013

REGIONI	Comuni		Famiglie		Individui	
	Campione	Universo	Campione	Universo (a)	Campione	Universo (a)
Piemonte / Valle d'Aosta - Vallée d'Aoste	84	1.280	1.837	2.031	4.043	4.563
Liguria	26	235	805	775	1.668	1.599
Lombardia	83	1.544	1.657	4.238	3.948	9.984
Trentino-Alto Adige	47	333	1.106	426	2.720	1.042
Veneto	54	581	1.146	2.003	2.842	4.925
Friuli-Venezia Giulia	32	218	701	538	1.637	1.223
Emilia-Romagna	47	348	1.044	1.891	2.426	4.444
Toscana	46	287	1.014	1.608	2.362	3.750
Umbria	22	92	571	370	1.371	905
Marche	34	239	703	610	1.795	1.561
Lazio	36	378	1.197	2.489	2.691	5.791
Abruzzo	37	305	794	537	1.962	1.341
Molise	23	136	601	127	1.491	317
Campania	55	551	1.359	2.082	3.738	5.813
Puglia	50	258	1.067	1.528	2.857	4.068
Basilicata	26	131	587	231	1.485	582
Calabria	43	409	946	792	2.339	1.996
Sicilia	50	390	1.219	1.972	3.086	5.019
Sardegna	39	377	797	705	1.854	1.665
Italia	834	8.092	19.151	24.952	46.315	60.588

(a) Stima Indagine multiscopo "Aspetti della vita quotidiana", dati in migliaia.

Indice

- 1 Le indagini statistiche: introduzione
- 2 Esempi di indagine statistica
- 3 I censimenti
- 4 Le poverty surveys
- 5 Le indagini multiscopo
- 6 Rilevazione sulle forze lavoro**

Rilevazione sulle forze lavoro

- principale fonte di informazione statistica sul mercato del lavoro italiano
- Le informazioni rilevate consentono di ottenere 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
- rientra tra le indagini comprese nel Programma statistico nazionale, che individua le rilevazioni statistiche di interesse pubblico.
- condotta fin dal 1959, con una periodicità trimestrale nei mesi di gennaio, aprile, luglio e ottobre
 - ▶ Da gennaio 2004 la rilevazione è continua, cioè le informazioni sono rilevate con riferimento a tutte le settimane di ciascun trimestre, mediante una distribuzione uniforme del campione nelle settimane

Rilevazione sulle forze lavoro (2)

- numerosità campionaria: Ogni anno viene intervistato un campione di oltre 250mila famiglie residenti in Italia (per un totale di circa 600 mila individui) distribuite in circa 1400 comuni italiani.

Rilevazione sulle forze lavoro (3)

- L'indagine prevede che la prima intervista venga effettuata di norma presso il domicilio della famiglia, da intervistatori che utilizzano un personal computer (metodo CAPI) per inserire le informazioni rilevate mediante **questionario**. Le interviste successive alla prima, per la famiglia di cui si dispone del numero di telefono, vengono effettuate di norma telefonicamente (metodo CATI)
- vengono intervistati tutti i componenti della famiglia con più di 15 anni
- Definizioni e dati
 - ▶ Gli **occupati** sono le persone di almeno 15 anni che al momento della rilevazione hanno dichiarato di possedere un'occupazione oppure che hanno effettuato nella settimana di riferimento almeno un'ora di lavoro
 - ▶ **In cerca di occupazione**: disoccupati o in cerca di prima occupazione
 - ▶ **Non forze lavoro**