

Presentazione della conferenza

"Inferenza Statistica e Indagini Campionarie"

dott. Susy Osti Dipartimento Studi ISTAT Veneto ITIS Vito Volterra

> *Maggio 2008 Classi 5B e 5C*



L'incontro di oggi

si articola in quattro parti

- 1. il sito www.istat.it
- 2. le principali indagini campionarie sulle famiglie
- 3. i campioni utilizzati e la rilevazione sul campo
- 4. alcuni risultati



Istat_it http://www.istat.it

domenica 25 maggio 2008, ore 10:18

Home

L'Istituto

Sala stampa

Dati e prodotti

Servizi

Censimenti

Strumenti

enalish

mappa | quida | contatti |

newsletter | link utili | Archivio documenti ₩

cerca



l'Istituto

- ▶Organizzazione
- Il Presidente
- ▶Rilevazioni
- Attività internazionale
- Audizioni
- ▶Convegni e seminari

sala stampa

- Comunicati
- ▶Appuntamenti
- ▶Per i giornalisti

dati e prodotti

- Banche dati
- ▶Tavole di dati
- File di microdati
- Catalogo
- Pubblicazioni scientifiche

servizi

- · Supporto agli utenti
- Abbonamenti
- Biblioteca
- Archivio storico

e novità



Commercio estero. Ad aprile -1.782 mln di euro il saldo con l'extra-



Pil. Nel I trimestre 2008 +0.4% sul IV trim. 2007, +0,2% in un anno



Vendite, A marzo -0,5% su febbraio e -1,0% in un anno



Industria, A marzo -4,3% il fatturato e -3,7% ali ordini in un anno

Commercio estero UE e Mondo Marzo 2008 Prezzi al consumo (definitivi) Aprile 2008 Produzione industriale Marzo 2008

Archivio M

ultimi dati

	mum	ual
	periodo	%*
Pil	3T/07	+1,9
Prezzi al consumo	04/08	+3,3
Prezzi alla produzione	03/08	+6,3
Produzione industriale	03/08	-2,5
Vendite al dettaglio	03/08	-1,0
Occupati	4T/07	+1,3
Retribuzioni	03/08	+2,8

*variazioni tendenziali

Drossimi appuntamenti

- 27 maggio Contratti, retribuzioni, conflitti Aprile 2008
- 28 maggio Presentazione del Rapporto Annuale Anno 2007
- 29 maggio Lavoro e retribuzioni grandi imprese Marzo 2008

statistiche per argomento

Popolazione

Struttura demografica Dinamica demografica Stranieri

Famiglia e società

Strutture familiari

Industria e servizi

Struttura e attività delle imprese Industria e costruzioni Vendite Turismo e trasporti Sviluppo e innovazione

MOBILE. RSS

cont@ct centre

▶ Per informazioni e dati statistici e servizi di diffusione

tutto SU

PREZZI

- Indice per aggiornare affitti, assegni familiari, TFR
- · Informazioni, dati e



che informazione c'è?

- □ Dati & prodotti
 - banche dati / datawarehouse
 - informazione "strutturata"
 - strumenti per cercare ed estrarre informazioni
 - pubblicazioni on-line
 - copia elettronica dei volumi cartacei
 - dati e commenti
- Metadati & strumenti



Per una ricerca mirata...

- ▶Banche dati.
- ▶Tavole di dati
- File di microdati
- ▶Catalogo
- Pubblicazioni scientifiche

dafie prodotti Catalogo editoriale

Il catalogo contiene l'elenco completo delle pubblicazioni edite dall'Istat a partire dall'anno 2000 (le pubblicazioni precedenti sono disponibili presso la biblioteca).

Ad ogni pubblicazione corrisponde una scheda che ne riporta le informazioni bibliografiche e una breve descrizione contenutistica. La gran parte dei volumi presenti in questo catalogo è accessibile e scaricabile gratuitamente.

Le edizioni cartacee sono disponibili presso i centri di informazione statistica presenti in ogni regione e provincia autonoma e presso le librerie con cui l'Istat collabora. I volumi possono anche essere richiesti compilando l'apposito modulo di acquisto.

Ricerca per caratteristiche editoriali

Testo:	Cellulare
Settori editoriali: (descrizione dei settori)	Agricoltura e zootecnia Ambiente e territorio Commercio estero Conti economici Cultura Famiglia e società
Collane: (descrizione delle collane)	Annali di statistica Annuari Argomenti Censimenti Documentazione Essays Seleziona tutto Deseleziona tutto
Numero:	



I risultati della ricerca...

Risultati della ricerca di pubblicazioni

1 prodotto trovato

Parametri di ricerca:

Testo cercato: Cellulare

Legenda Prodotto cartaceo

CD-ROM allegato (8)



Prodotto scaricabile

Se il risultato non è soddisfacente, esegui una nuova ricerca oppure prova con la ricerca avanzata. Se vuoi stampare i risultati, utilizza la versione stampabile.





Rilevanza: 100% I

L'uso dei media e del cellulare in Italia

Indagine multiscopo sulle famiglie "I cittadini e il tempo libero" -

Anno 2006

Settori: Famiglia e società, Cultura

Periodo dei dati: Anno 2006 Collana: Informazioni, n. 2

Anno di edizione: 2008

In preparazione









. . .

dafi e prodotti

- ▶Banche dati
- ▶Tavole di dati
- File di microdati
- ▶Catalogo
- ▶Pubblicazioni scientifiche

<u>argomenti</u>

- Famiglia e società
- Comportamenti e stili di vita
- « Archivio

L'uso dei media e del cellulare in Italia

Indagine multiscopo sulle famiglie "I cittadini e il tempo libero" – Anno 2006

Periodo di riferimento: Anno 2006

Diffuso il: 29 aprile 2008

Settori: Famiglia e società Periodo dei dati: Anno 2006 Collana: Informazioni, n. 2 Anno di edizione: 2008

Supporti: **[[[]]**Prezzo: **28.00** €
In preparazione

Cod. ISBN: 9788845815799

Cod. SIGED: 21012008002000005

Il volume presenta i risultati dell'indagine multiscopo "I cittadini e il tempo libero" 2006, relativi all'uso dei media e del cellulare in Italia. Vengono presi in considerazione i dati sull'uso della televisione, di videogiochi, videocassette e/o dvd, sull'ascolto della radio, la lettura di quotidiani, settimanali e riviste non settimanali, l'utilizzo dei telefoni cellulari e il ricorso all'edicola.

Sono state raggiunte quasi 20 mila famiglie, per un totale di oltre 50 mila individui. L'analisi è stata condotta prendendo in considerazione le caratteristiche anagrafiche, sociali e territoriali degli individui, in modo da evidenziare il rapporto e il grado di familiarizzazione degli italiani con i media e le tecnologie dell'informazione e della comunicazione.

download

- ▶Testo integrale PDF (3408 kbyte)
- ▶Tavole ZIP (1598 kbvte)
- ▶Indice delle tavole PDF (263 kbyte)

per informazioni

Servizio Struttura e dinamica sociale Tel. 0646734841 adolfo.morrone@istat.it

Centro diffusione dati tel. 06 4673.3102-3-5-6 fax 06 4673.3101-7 Cont@ct centre

Commercializzazione e marketing tel. 06 4673.3278-79-80 fax 06 4673.3477

marketing@istat.it



i metadati

sono le informazioni sui dati

- come sono stati raccolti?
 - Metodologia
- secondo quali criteri?
 - Definizioni
 - Classificazioni



è essenziale capire da dove viene un numero, soprattutto quando si fanno dei confronti



Home: Strumenti: Metodi e software: Linee guida





metodi e software

- Linee guida
- ▶Software
- Destagionalizzazione
- Indici a catena
- ▶Pubblicazioni

definizioni e. classificazioni

gualità delle

per i rispondenti

▶Indagini: questionari e informazioni

<u>Linee guida metodologiche per rilevazioni statistiche</u>

Sono disponibili nozioni metodologiche di base e pratiche consigliate per rilevazioni statistiche dirette o basate su fonti amministrative.

- Premessa
- Progettare l'indagine
- Obiettivi, definizioni, classificazioni
- Disegno di indagine
- Indagini trasversali e longitudinali
- Indagini totali e campionarie
- Archivi di base
- Strategia di campionamento
- Tecniche di indagine
- Questionario [da Istat (1989) vol.2]
- Tempi e costi
- Sistema dei controlli di qualità
- Gruppo di progettazione
- Documento di progettazione
- Sperimentazioni (della fase progettuale)
- Fasi operative
- Rilevazione
- Registrazione su supporto magnetico
- Revisione automatica.
- Codifica dei quesiti aperti
- Elaborazioni statistiche [da Statistics Canada (1987)]
- Validazione
- Diffusione
- Indagini Amministrative
- Uso dei dati amministrativi per fini statistici
- Raccolta dei dati amministrativi
- Reperimento delle pratiche amministrative
- Trasposizione del dato amministrativo in informazione statistica
- Spedizione dei dati amministrativi all'ente statistico

Gli strumenti per la produzione statistica sono le metodologie adottate per il disegno delle indagini, l'acquisizione dei dati e la loro codifica, e dalle procedure e i software per la produzione delle informazioni



metodi e software

- +Linee guida
- FSoftware
- Destagionalizzazione
- ▶Indici a catena
- Pubblicazioni

definizioni e classificazioni

qualità delle indagini

per i **rispondenti**

 Indagini: questionari e informazioni Home: Strumenti: Definizioni e classificazioni

Definizioni e classificazioni

definizioni stat

Glossario dei termini statistici

altre definizioni internazionali

- **→ UN Glossary of Classification Terms**
- OECD Glossary of Statistical Terms
- ISI Glossary of Statistical Terms

classificazioni adottate a livello nazionale e internazionale agevola la comparabilità delle informazioni statistiche

L'uso di definizioni e

prodotte da organismi diversi

classificazioni adottate dall'Istat

- Classificazione delle attività economiche ATECO
- Classificazione delle forme giuridiche delle unità legali
- Classificazione delle professioni
- Classificazione dei titoli di studio



un percorso un po' nascosto ...

... ma interessante

Per gli studenti
Università
Scuola

servizi

- Supporto agli utenti
- Abbonamenti
- → Biblioteca
- Archivio storico
- Per gli studenti
- Concorsi e selezioni
- ▶Gare e appalti



cosa ci si può trovare ...

- □ lezioni di statistica serie, ma non troppo ←
- excel per la statistica
- □ L'ipertesto "valore dei dati"
- esperienze di altre scuole
- □...

Nota introduttiva



Statistica per esempi è composto di 11 capitoli e di due appendici che illustrano il questionario per la rilevazione dei dati e le rappresentazioni grafiche. Il **menù**, posizionato in alto a sinistra, permette l'accesso alla nota introduttiva, ai capitoli e al glossario che fornisce le definizioni dei termini, dei concetti e delle formule usate.

Il primo capitolo ha per protagonista Paolo, un tredicenne che gioca a ritrovarsi nelle statistiche ufficiali. Per permettere di seguire i passi di Paolo nel gioco é stata predisposta anche una visualizzazione interattiva. All'interazione si può accedere anche attraverso la guida*.

Le tavole

Le tavole sono tratte dall'Annuario statistico italiano 2002.

Dal 12 dicembre 2007 su www.istat.it è disponibile l'Annuario statistico italiano 2007.

Per agevolarne la lettura, alcune tavole non sono riportate integralmente, data la loro complessità, ma soltanto per le parti trattate nel testo. In questi casi, chiaramente indicati, un link permette di visualizzare la tavola completa.

I capitoli

Per consentire a ciascun lettore di utilizzare i testi secondo le proprie esigenze i capitoli sono rilasciati anche in formato word e pdf.

Il carattere del testo può essere ingrandito utilizzando l'opzione **Visualizza Carattere** (molto grande, grande, medio, piccolo) nel menù del browser.

Indice del libro

- 1. Un gioco: ritrovarsi nelle statistiche ufficiali (22 dicembre 2003) W e
- 2. <u>Una lettura statistica dei giornali</u> (27 febbraio 2004) W e 📆
- 3. Uomini, donne, sport (26 maggio 2004) W e 📆
- 4. Quanti eravamo, quanti siamo, quanti saremo (16 giugno 2004) 👿 e 🏗
- 5. Il campionato di calcio italiano di serie A (28 giugno 2004) W e
- 6. La pesca sportiva in acque dolci (28 settembre 2005) W e

L'uso di Excel per la statistica

Esercitazione











🖹 🗐 2.Lavorare con Excel



2.1.L'ambiente di lavoro



2.3.Installazione di funzioni avanzate



☐ 3.Rappresentazione grafica dei dati



3.1 Frequenza assoluta e relativa



3.2.Tabella di frequenza



📵 3.3.Diagramma a barre



📵 3.4.Istogramma



億 3.5.Diagramma degli estremi e quartili





🖹 词 4. Valori di posizione



4.1.Media



📳 4.3.Mediana





🖹 词 5.Misure di dispersione



🗐 5.1.Varianza



📵 2.2.Operazioni base nel foglio di lavoro. Excel per la statistica fa parte del set di materiali didattici per gli studenti e gli insegnanti delle scuole medie inferiori e superiori che vogliono. capire, studiare, utilizzare la statistica. I contenuti sono : Che cos'è il censimento, Che cos'è la demografia, Excel per la statistica, il Glossario dei termini usati, le Faq (domande più frequenti). Excel per la statistica e il Glossario sono i primi rilasci.

> Questo lavoro è stato predisposto tenendo conto dell'esperienza dell'Istituto di statistica del Portogallo (www.ine.pt). Gli studenti e i docenti sono invitati a far pervenire osservazioni e commenti.

Consultazione-download (7.198 KB, formato zip).

Requisiti: Netscape (versioni 4.0 e successive), Explorer (versioni 4.0 e successive), Excel (versioni 95 e successive).

a cura di binariodieci binariodieci@istat.it per informazioni rivolgersi a Danila Fulgenzio fulgenzi@istat.it





"Il valore dei dati" è un ipertesto concepito per aiutare i lettori non esperti a usare al meglio le statistiche. Si basa su materiali analoghi prodotti da istituzioni internazionali e da istituti di statistica di altri paesi ed è diretto a chiunque abbia un interesse verso l'uso dei numeri per descrivere i fenomeni della vita reale. Non è un corso vero e proprio, perché ogni scheda è concepita per essere autosufficiente e non richiede la lettura delle schede precedenti, ma può anche essere usato in modo sistematico da studenti e insegnanti interessati ad approfondire l'uso concreto della statistica. Chi cerca un approccio amichevole a informazioni statistiche su uno specifico settore potrà trovare un aiuto nella colonna sinistra di questo sito, mentre chi invece desiderasse esaminare altro materiale di divulgazione statistica disponibile sulla Rete può avvalersi dei link posti nella colonna destra.

I dati dell'Istat per le vostre ricerche: un accesso amichevole a temi generali

La popolazione

La congiuntura

Il lavoro

I prezzi

L'ambiente

"Il valore dei dati" è stato realizzato dall'Istituto nazionale di statistica ed è stato finanziato, in parte, dal Ministero dell'istruzione, dell'università e

Le statistiche: a che cosa servono, come si elaborano, come si presentano, quanto sono credibili

L'uso dei dati per valutare e decidere

Oggi più che mai i governi, le imprese, i cittadini hanno bisogno di informazioni statistiche attendibili. Capire i numeri è indispensabile per progredire. È attraverso concetti statistici che si misurano, per esempio, il benessere di una comunità, il grado di mantenimento degli obiettivi di stabilità nell'Europa dell'euro, ma anche scolarizzazione, livelli di povertà, degrado ambientale, cioè successi e insuccessi rispetto ai grandi obiettivi al 2015 che l'umanità si è data attraverso le Nazioni Unite. (Continua)

I fondamentali della statistica: raccolta, produzione, diffusione

I dati statistici possono provenire da diversi tipi di rilevazioni. Il processo di produzione di statistiche su un determinato fenomeno parte dall'analisi delle possibili fonti dei dati e si conclude con la pubblicazione delle informazioni. In questo capitolo sono descritte anche le diverse modalità di diffusione delle statistiche da parte dell'Istituto nazionale di statistica e di Eurostat... (Continua)

Come fidarsi della qualità dei dati

In tutti gli ordinamenti giuridici la statistica ufficiale è quella prodotta dagli istituti nazionali di statistica e dai sistemi a essi collegati. Le organizzazioni pubbliche di produzione statistica e gli organismi internazionali raccolgono, elaborano e diffondono i dati; concordano e stabiliscono le metodologie, le

Strumenti e collegamenti per conoscere meglio il mondo dei numeri

Scrivere una relazione utilizzando dati statistici

Glossario

Catalogo dei prodotti editoriali

Definizioni e classificazioni

Metodologie

Pubblicazioni scientifiche

Statistica per esempi

L'uso di excel per la statistica



la vita quotidiana in statistica

si possono trovare informazioni su molti aspetti del nostro vivere quotidiano

- ambiente
- famiglie
- lavoro
- educazione e comunicazione
- salute
- criminalità e violenza
- stili di vita e qualità della vita
- **.** . . .



la vita quotidiana in statistica

un esempio: il volume

Cultura socialità e tempo libero 🛑



- Mass media e letture
- L'uso del PC ed Internet
- Il tempo libero: spettacoli ed intrattenimenti
- Attività fisica e sportiva
- Le vacanze
- Chi va in vacanza e chi non ci va, perchè?
- La partecipazione sociale e la pratica religiosa
- La partecipazione politica
- La soddisfazione per l'anno trascorso
- Gli amici



RILEVAZIONE DEI DATI

COME SI REPERISCONO LE INFORMAZIONI?

DA CHI SONO DATI I "DATI"?

Prima di calcolare delle statistiche sui DATI è necessario conoscerne la **fonte** e capire da che **tipo di rilevazione** sono stati prodotti

Ci sono tecniche e metodi statistici diversi per le varie tipologie di DATI





rilev. SPERIMENTALE

LA PRODUZIONE DEI DATI E' **CONTROLLATA** DA CHI GUIDA L'ESPERIMENTO

sono sotto controllo:

- le condizioni dell'esperimento e

es. coltura batteri a temperatura di 18°

- le caratteristiche delle unità statistiche scelte

es. malati di epatite C, maschi

il controllo avviene:

- direttamente, tramite il disegno sperimentale o

es. esperimenti chimici in laboratorio

- indirettamente, tramite la randomizzazione

es. somministrazione di due farmaci a due gruppi di pazienti





rilev. OSSERVAZIONALE

NON SI POSSONO CONTROLLARE LE CONDIZIONI

IN CUI VENGONO "PRODOTTI I DATI"

es. n. di incidenti stradali a Padova nel 2001

NON SEMPRE SI POSSONO CONTROLLARE LE CARATTERISTICHE DELLE UNITA' STATISTICHE

l'osservatore ha un **ruolo passivo** si limita a osservare / misurare i fenomeni che accadono senza poterli modificare





INDAGINE STATISTICA

E' UNA SITUAZIONE OSSERVAZIONALE

OBIETTIVO: STUDIARE UN FENOMENO

es. propensione all'acquisto dei visitatori di un sito internet

PASSI:

- individuare la **popolazione** (chi, quando, dove, ...)

 es. visitatori del sito comunale nel mese di marzo 2004
- individuare il **carattere** d'interesse e le sue **modalità** *es. utilizzo di servizi on-line*
- individuare altri caratteri collegati

es. sesso, età, ...

- individuare il **metodo di misura** es. intervista CAWI (Computer Assisted Web Interview)

.... cioè PROGETTARE L'INDAGINE



INDAGINE STATISTICA

PUO' ESSERE:

 TOTALE O CENSUARIA
 si rilevano i caratteri di tutte le unità della popolazione

CAMPIONARIA

si rilevano i caratteri di un **sottoinsieme di unità** della popolazione

per **induzione** si ottengono informazioni su tutta la popolazione





LE DUE STATISTICHE

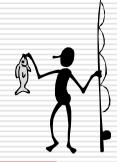
STATISTICA DESCRITTIVA

SINTETIZZA
IL MODO DI MANIFESTARSI DI UN FENOMENO
NELLA **POPOLAZIONE OSSERVATA**

STATISTICA INFERENZIALE

ATTRAVERSO L'OSSERVAZIONE DI UNA SOTTO-POPOLAZIONE (POCHE UNITA') CONSENTE DI OTTENERE INFORMAZIONI SU UNA **POPOLAZIONE PIU' AMPIA**





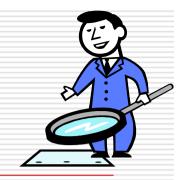
INDAGINE CAMPIONARIA

VANTAGGI:

- RIDUZIONE DI TEMPI E COSTI
- MIGLIOR CONTROLLO DELLA RILEVAZIONE
- INFORMAZIONI PIU' APPROFONDITE

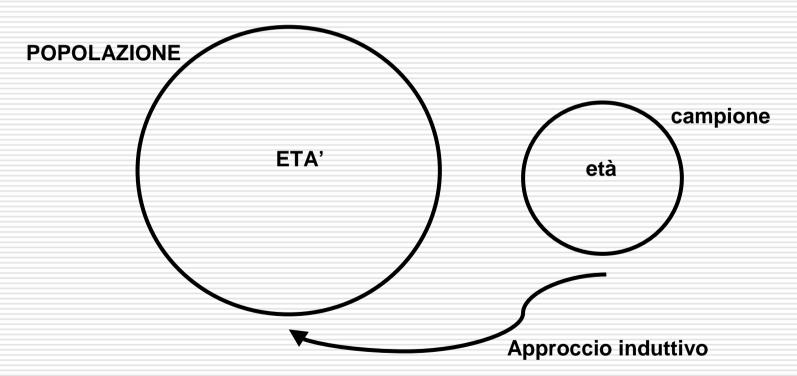
es. censimento della popolazione es. popolazioni con numerosità molto elevate es. indagini di mercato sui potenziali acquirenti di un bene es. casi in cui l'osservazione distrugge l'unità statistica es. fenomeni sociali, interviste lunghe e dettagliate, ...



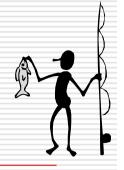


INFERENZA INDUTTIVA

è il principio su cui si basano le tecniche di campionamento e statistica inferenziale







LA CASUALITA'

UN CAMPIONE E' CASUALE QUANDO

LA SELEZIONE DELLE UNITA' E' AFFIDATA A REGOLE PROBABILISTICHE

IL "CASO" E' UN ALLEATO PREZIOSO:

- CONSENTE DI APPLICARE LE TECNICHE DI INFERENZA STATISTICA
- EVITA LE DISTORSIONI SISTEMATICHE CAUSATE DA SELEZIONI "SOGGETTIVE"



un campione CASUALE

- ciascuna unità della popolazione ha uguale probabilità di essere estratta
- può essere "Senza reinserimento", cioè le unità estratte non vengono riesaminate, non possono essere estratte di nuovo
 - man mano che si estraggono le unità campione, cambiano le probabilità di estrarre le altre
- può essere "CON reinserimento", cioè ogni elemento campionato ritorna a far parte della popolazione e può essere estratto di nuovo
 - la probabilità di estrarre un elemento rimane costante





alcuni esempi ...

- un campione di lampadine estratto da un processo di produzione per verificarne la qualità
- un campione di utenti della biblioteca per studiare il loro utilizzo e possesso di mezzi tecnologici moderni per la comunicazione (attenzione al reinserimento ...)
- un campione di cittadini aventi diritto al voto per prevedere i risultati delle prossime elezioni amministrative
- un campione di casalinghe per analizzare il fenomeno degli incidenti domestici





alcuni dei campioni più frequenti ... CAMPIONAMENTO STRATIFICATO

prima di estrarre casualmente le unità campionarie la popolazione viene divisa in gruppi (strati)

Lista di campionamento: elenco individui, raggruppati per sesso

Uomini U1 U2 U3 U4 U5 U6 U7 U8 U9 U10 U11 U12 ... Un

Donne D1 (D2) D3 (D4) D5 D6 D7 D8 D9 D10 (D11) D12 ... Dn

viene usato per estrarre i Comuni che faranno parte del campione per le indagini campionarie sociali

l'Italia viene tagliata in pezzi (strati) divisi per regione e tipo di comune, in base all'ampiezza demografica in modo da essere sicuri di includere nel campione comuni di tutte le regioni e di tutte le dimensioni





CAMPIONAMENTO A DUE STADI

L'estrazione delle unità di interesse avviene in due passi (stadi) successivi

Lista di campionamento di primo stadio: elenco comuni italiani raggruppati (stratificati)

per regione e tipo

Strato A C1 C2 C3 C4 C5 C6 C7 C8 C9 C10

dalla lista anagrafica del Comune 4, estratto al primo stadio vengono estratte le unità campionarie finali: le famiglie

viene usato per estrarre le famiglie all'interno dei Comuni che faranno parte del campione per le indagini campionarie sociali





CAMPIONAMENTO SISTEMATICO

Scelta casualmente la prima unità tutte le altre sono individuate automaticamente grazie al "passo" di campionamento

Lista di campionamento

U1 (U2) U3 U4 (U5) U6 U7 (U8) U9 U10 (U11) U12 U13 ... Un

primo numero casuale = 2 passo = 3

viene usato nei comuni per estrarre dalla lista anagrafica le famiglie da intervistare per le indagini campionarie sociali





la STIMA

 consiste nell'attribuire il valore "più appropriato" a un parametro sconosciuto della popolazione

es. si vuole conoscere il voto medio alla maturità degli studenti di ingegneria, attraverso un campione degli iscritti all'A.a. 2006-2007

- la si calcola
 - attraverso i risultati campionari
 - con prefissata probabilità di errore



Ia STIMA PUNTUALE

Il parametro (caratteristica) sconosciuto della popolazione viene stimato con il valore puntuale espresso dalla statistica campionaria

es. dato un campione di 1.000 nati nel 2007 stimo che il peso medio dei nati nel 2007 sia 3.100 gr pari al peso medio dei 1.000 nati campione



la STIMA PUNTUALE

es. si vuole conoscere la percentuale di utenti soddisfatti del trasporto ferroviario attraverso un campione dei viaggiatori nel mese di dicembre 2007

- la si calcola
 - attraverso i risultati campionari es. % di "soddisfatti" tra i viaggiatori estratti nel campione
 - nel rispetto di certi criteri di ottimalità es. devo calcolare una % precisa al secondo valore decimale?
 - con prefissata probabilità di errore es. data la numerosità campionaria e il tipo di fenomeno studiato (variabilità, ...) mi potrò sbagliare di circa un tot %



Ia STIMA INTERVALLARE

L'informazione sulla caratteristica sconosciuta della popolazione viene rappresentata con due statistiche campionarie:

limite inferiore e limite superiore

dell'intervallo di confidenza all'interno del quale si stima sia contenuto il vero valore

> es. dato un campione di 100 iscritti nel 2007 stimo che l'altezza media degli studenti di ingegneria sia compresa tra 172 e 176 cm

l'ampiezza dell'intervallo determina la precisione di stima: quanto minore è l'ampiezza tanto maggiore è la PRECISIONE DI STIMA

mi è di qualche utilità sapere che l'altezza media degli studenti di ingegneria è compresa tra 10 cm e 10 metri? questo è un intervallo di confidenza al 100%!



Le principali indagini Istat

campionarie presso le famiglie

- □Indagini Multiscopo
- Consumi delle famiglie
- □Forze di Lavoro
- □Indagini sui diplomati e i laureati

Sistema "Multiscopo" di indagini sulle famiglie

- □ dal 1993 (con qualche precedente nel 1987-91):
 - vita quotidiana
 - qualità della vita
 - statistica sociale
 - punto di vista dei cittadini
 - aspetti quantitativi e qualitativi della vista di tutti i giorni delle famiglie italiane



Le indagini Multiscopo

- ogni anno a partire dal 1993
 - "Aspetti della vita quotidiana"
- ogni cinque anni, a rotazione
 - "Salute e ricorso ai servizi sanitari
 - "Tempo libero e cultura"
 - "Sicurezza dei cittadini"
 - "Famiglia e soggetti sociali"
 - "Uso del tempo"
- □ dal 1997 "Viaggi e vacanze" (ogni trimestre)



- □disegno campionario
 - campionamento a 2 stadi
 - con stratificazione delle unità di 1° stadio
 - 1° stadio = Comuni campione, estratti dall'Istat
 - 2º stadio = Famiglie campione, estratte dalle anagrafi comunali
- numerosità campionaria
 - circa 800 comuni (circa 50 in Veneto)
 - circa 23.000 famiglie (circa 1.400 in Veneto)

Prospetto 11.1 - Distribuzione regionale dei comuni, delle famiglie e degli individui nell'universo e nel campione per l'indagine Annuale anno 2003 (dati in migliaia)

REGIONI	Comuni		Famiglie		Individul	
	Universo	Campione	Universo (a)	Campione	Universo	Campione
Plemonte	1.206	60	1.820.161	1.535	4.236.628	3.642
Valle d'Aosta	74	20	52.533	517	119.329	1.180
Lombardia	1.546	79	3.670.047	1.824	9.072.928	4.527
Bolzano Bozen	115	22	173.546	604	460.034	1.659
Trento	223	24	196.704	577	473.872	1.406
Veneto	581	49	1.695.018	1.134	4.506.136	3.064
Friuli-Venezia Giulia	219	31	507.714	777	1.178.292	1.844
Liguria	235	26	724.007	831	1.599.770	1.835
Emilia-Romagna	341	45	1.688.525	1.170	3.992.148	2.826
Toscana	287	49	1.412.899	1.167	3.519.918	2.908
Umbria	92	22	318.129	634	835.596	1.703
Marche	246	36	559.525	864	1.463.300	2.275
Lazio	377	33	2.194.208	1.261	5.268.803	3.069
Abruzzo	305	37	471.252	885	1.275.503	2.412
Molise	136	23	124.590	599	324.783	1.516
Campania	551	54	1.922.896	1.408	5.758.610	4.204
Puglia	258	48	1.417.826	1.163	4.066.859	3.338
Basilicata	131	24	209.786	586	600.454	1.684
Calabria	409	41	718.691	969	2.026.616	2.777
Sidila	390	48	1.818.778	1.265	5.039.128	3.560
Sardegna	377	38	574.566	804	1.632.290	2.279
Italia	8.100	809	22.271.401	20.574	57.450.997	53.708



- Criteri di stratificazione al primo stadio
 - Strato = gruppo di Comuni tra loro "omogenei"
 - vicinanza geografica, dimensione demografica simile



questionari di rilevazione

sono di due tipi per ogni indagine

- 1. Questionario per intervista diretta
- 2. Questionario per autocompilazione



- modalità di esecuzione della rilevazione sul campo (PAPI*)
 - invio della lettera di preavviso da parte del Comune
 - il rilevatore intervista la famiglia ed i singoli componenti, lascia in consegna i questionari per autocompilazione per passare poi a ritirarli dopo qualche giorno
 - non è prevista la sostituzione della famiglia in caso di "caduta"

* Paper and pencil interview



- □ Rappresentatività dei dati
 - a livello nazionale
 - a livello regionale
 - per tipologie di Comuni
 - □Comuni centro dell'area metropolitana nel loro insieme (TO, MI, VE, GE, BO, FI, Roma, NA, BA, PA, CT, CA)
 - □Comuni intorno alle aree metropolitane
 - \square Comuni < o = 2.000 abitanti
 - □Comuni > 2.000 e < 10.000
 - □Comuni > 10.000 e < 50.000
 - □Comuni > 50.000



- □Diffusione dei dati
 - ■Sito www.istat.it
 - Editoria
 - Note Rapide
 - □ Volumi del settore "Famiglia e società"
 - □Volumi della serie "Argomenti"
 - □ Volumetti tascabili serie "Il Mulino"
 - □ Comunicati stampa
 - П...



Indagine sulla *sicurezza dei cittadini e* su *viaggi e vacanze*

- □ intervista telefonica
- □ la lista di selezione delle unità è l'elenco degli abbonati Telecom per il settore famiglie
 - sovracopertura tutti i numeri telefonici che corrispondono a seconde case o ad attività professionali
 - sottocopertura tutte le famiglie non abbonate al telefono o intestatarie di numeri riservati
- disegno di campionamento probabilistico, piuttosto che il campionamento per quote

Perché un campionamento Probabilistico?

- assegna probabilità di inclusione note e positive a tutte le unità della popolazione appartenenti alla lista
- consente una migliore confrontabilità dei risultati con quelli provenienti dalle altre indagini condotte dall'Istat sulle famiglie
- permette un miglior controllo delle operazioni di rilevazione svolte dalla società preposta alla rilevazione nel caso di svolgimento dell'intervista telefonica da parte di una società esterna,



Disegno di campionamento

- a uno stadio stratificato
 - Unità di primo stadio gli indirizzi telefonici per stratificazione territoriale.
- la lista degli abbonati al telefono contiene unicamente variabili di tipo territoriale è integrata con altre informazioni territoriali a livello comunale, tipologia socio-demografica dei comuni
- la stratificazione viene effettuata secondo l'incrocio delle modalità della regione e delle sei tipologie comunali



Indagine **su viaggi e vacanze**

- □ Il disegno di campionamento utilizzato è del tipo stratificato a grappoli, in cui ciascun grappolo è un indirizzo telefonico
- □ Tutti gli individui appartenenti ai grappoli campione, cioè appartenenti alle famiglie a cui corrispondono i numeri selezionati, vengono inclusi nel campione.
- È ammessa l'intervista proxi, in quanto si ritiene che la persona di riferimento (che risponde al telefono) può dare risposte corrette sia sui viaggi e le vacanze della famiglia, che sui viaggi e le vacanze dei singoli componenti



Indagine sulla sicurezza dei cittadini

- Disegno di campionamento a due stadi con stratificazione delle unità di primo stadio
 - Le unità di primo stadio sono gli indirizzi telefonici
 - Le unità di secondo stadio sono gli individui: per ciascuna famiglia si estrae, mediante metodo di selezione casuale, un individuo campione tra i componenti della famiglia aventi età >14 anni
- Agli individui campione vengono posti sia quesiti riguardanti i fatti delittuosi subiti dalla famiglia di appartenenza, sia quesiti relativi ai fatti delittuosi da lui subiti



Numerosità campionarie

Prospetto 11.2 - Distribuzione regionale del campione dell'indagine sulla sicurezza dei cittadini e del campione trimestrale dell'indagine su viaggi e vacanze

Regioni	Indagine sicurezza	Indagine viaggi e vacanze	
Piemonte	3.796	265	
Valle d'Aosta	2.528	92	
Lombardia	4.622	330	
Bolzano Bozen	2.434	-	
Trento	2.414	-	
Trentino - Alto Adige	-	206	
Veneto	2.939	201	
Friuli - Venezia Giulia	2.687	139	
Liguria	2.773	152	
Emilia romagna	3.047	197	
Toscana	3.024	213	
Umbria	2.472	103	
Marche	2.561	141	
Lazio	3.138	219	
Abruzzo	2.656	144	
Molise	2.297	96	
Campania	3.036	196	
Puglia	2.862	187	
Basilicata	2.426	99	
Calabria	2.669	158	
Sicilia	2.986	213	
Sardegna	2.617	150	
Italia	60.000	3.501	



□Indagine di tipo socio-economico

□dal 1959 produce le statistiche principali sull'offerta di lavoro

(es. tasso di disoccupazione)



- ■disegno campionario
 - campionamento a 2 stadi
 - con stratificazione delle unità di 1º stadio
 - 1° stadio = Comuni campione, estratti dall'Istat
 - 2º stadio = Famiglie campione, estratte dalle anagrafi comunali
- numerosità campionaria
 - circa 1.250 comuni (72 in Veneto)
 - circa 77.000 famiglie (4.000 in Veneto)
 - □ogni famiglia viene intervistata per 4 volte a intervalli di tempo predefiniti



- modalità di esecuzione della rilevazione sul campo (CAPI*)
 - invio della lettera di preavviso da parte del Comune
 - il rilevatore intervista la famiglia ed i singoli componenti inserendo i dati in un personal computer

* Computer Assisted Personal Interview



- sostituzione
 - è prevista già in fase di campionamento
 - ☐ si estrae la famiglia base e
 - ☐ tre potenziali famiglie "sostitute" appartenenti alla stessa zona (contesto socio-economico)
- periodicità
 - sono "coperte" con interviste tutte le settimane dell'anno
 - per tenere sottocontrollo la stagionalità del fenomeno



- reinterviste alle famiglie secondo la tecnica d'indagine PANEL
 - per misurare le probabilità di transizione da uno stato all'altro, cioè vedere come muta nel tempo la condizione dei cittadini nel mercato del lavoro
 - si utilizza il sistema CATI per le reinterviste telefoniche (2°, 3° e 4° intervista)



- □Rappresentatività dei dati
 - a livello nazionale
 - regionale
 - provinciale
 - per tipologie di Comuni
- Diffusione
 - ■Sito <u>www.istat.it</u>
 - Editoria



Indagine sugli sbocchi professionali dei laureati

- popolazione universo
 - laureati nell'anno solare 1998 in tutti gli atenei italiani (129.307)
- predisposizione della lista campionaria
 - censimento dei laureati di tutti gli atenei italiani



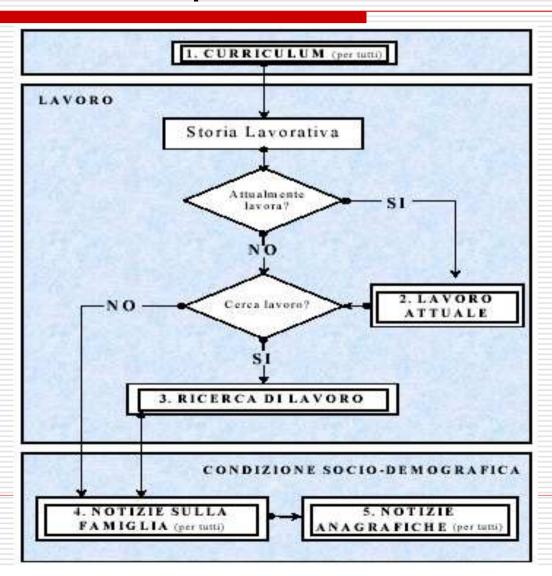
Indagine sugli sbocchi professionali dei laureati

- disegno campionario
 - campionamento a 1 stadio
 - stratificazione per sesso, sede universitaria e corso di laurea
- numerosità campionaria
 - campione teorico 30.052 laureati
 - campione effettivo 20.844 laureati



Indagine sugli sbocchi professionali dei laureati:

il questionario





Indagine sugli sbocchi professionali dei laureati

- modalità di esecuzione della rilevazione sul campo (CATI*)
 - invio della lettera di preavviso da parte dell'Istat
 - il rilevatore i laureati
 - non è prevista la sostituzione della famiglia in caso di "caduta"
 - * Computer Assisted Telephone Interview



Indagine sugli sbocchi professionali dei laureati

- □Rappresentatività dei dati
 - a livello nazionale
 - per ripartizione geografica (nord-ovest, nord-est, centro, sud, isole)
 - per gruppi di corsi di laurea, incrociati per sesso



COEFFICIENTI DI RIPORTO ALL'UNIVERSO





ogni unità campionaria rappresenta se stessa e altre unità del suo strato che non sono entrate a far parte del campione quindi

le risposte date dalle unità campionarie devono essere RIPORTATE all'UNIVERSO in modo che descrivano tutte le unità della popolazione

cioè si moltiplicano i risultati per i PESI CAMPIONARI detti anche coefficienti di riporto all'universo



Cosa significa moltiplicare per i pesi?

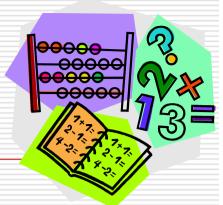


significa immaginare che le unità che non sono state estratte siano ben rappresentate da quelle campionarie

ogni unità campione PESA tanto quante sono le unità universo che è chiamata a rappresentare







i pesi sono pari a l'INVERSO DELLA PROBABILITA' DI ESTRAZIONE

	Unità	Unità	Probabilità	Coefficiente
	universo	campionarie	di	di riporto
Strato	N	n	estrazione	all'universo
1	1.200	140	0,117	8,571
2	1.600	160	0,100	10,000
3	2.700	280	0,104	9,643
		•••		•••
	• • •		•••	•••
k	800	68	0,085	11,765





perché si calcolano così?

perché devono essere inversamente proporzionali alla numerosità campionaria:

più sono le unità campionarie meno devo pesarle per far loro rappresentare le unità non intervistate

Esempio:

se da un gruppo di 1.000 persone
estraggo un campione di 100
(prob. estraz. = 100 / 1.000 = 1 / 10)
ciascuna unità campionaria ne rappresenta 10
(se stessa + altre 9 non intervistate)
quindi, le risposte delle 100 unità campionarie devono essere
moltiplicate per un peso pari a 10



la triste realtà degli statistici: non esistono solo gli errori campionari, purtroppo

Le formule matematiche della teoria dei campioni ci permettono di calcolare gli ERRORI CAMPIONARI cioè quelli legati al solo fatto che studiamo una "parte" per trarre informazioni sul "tutto"

ma, nel fare indagini ed esperimenti si compiono molti

ERRORI NON CAMPIONARI

e per questi la teoria della probabilità ci aiuta poco

Un consiglio da zia: prima di prendere in considerazione dati elaborati con formule complicate, controllatene la qualità, verificate che siano stati raccolti con metodi statistici seri, che siano state rispettate le regole della casualità e che le misure siano state fatte in modo corretto

altrimenti: potreste perdere il vostro tempo e prendere decisioni errate



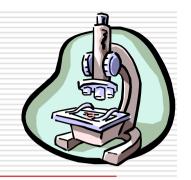
alcuni esempi ...

Se le bilance per pesare le unità di interesse sono mal tarate e indicano sempre qualche grammo in meno del peso effettivo, si otterrà sicuramente una stima distorta per difetto

Se la metà delle persone estratte in un campione non rispondono ai quesiti che si è progettato di porre loro ci sarà un bel da fare a far calcoli, ma io non mi fiderei di quelle stime, qualsiasi sia il numero di decimali usato nei calcoli

Se si farà un campione di confezioni di sale prodotto da una catena per stimarne il peso medio proprio in un giorno in cui c'è un elevato tasso di umidità si sarà indotti a pensare che il processo di produzione produca sacchetti troppo pieni e pesanti





I'ACCURATEZZA

dipende dalla vicinanza tra la statistica prodotta e il valore vero

quant'è la popolazione italiana ? quanto accurato è il valore ottenuto dai dati di censimento ?

ha due componenti

sistematica

gli stranieri irregolari saranno sicuramente sfuggiti spesso alle operazioni censuarie

casuale

chi ha compilato il questionario avrà commesso qualche errore "di sbaglio"



gli ERRORI NON CAMPIONARI

sono una componente importante dell'errore totale

dipendono dall'operato sul campo di tutte le persone coinvolte in un'indagine

e riguardano tutte le fasi in cui si articola un'indagine statistica

i principali riguardano

- copertura, selezione del campione
- mancata risposta totale e parziale
- misurazione
- trattamento dati





COPERTURA e CAMPIONAMENTO

* non sempre la lista da cui si estrae il campione contiene tutte le unità della popolazione da studiare (universo)

nell'elenco anagrafico comunale manca qualche famiglia e ne compare qualche altra che ormai non vi risiede più il problema è la distanza tra la situazione di fatto e quella "de jure"

* la selezione delle unità campionarie deve rispettare i criteri di casualità, altrimenti l'immagine prodotta dal campione non rispecchia la popolazione da cui proviene

non è facile per chi contatta le famiglie campione affrontare casi "difficili":
famiglie problematiche, stranieri, individui con problemi di salute, ...
ma l'Italia è fatta anche di queste persone
chi modifica artificiosamente le liste di campionamento
eliminando le unità "scomode"
distorce profondamente la qualità della rilevazione
appiattendo il risultato verso una situazione
"mediamente felice e tranquilla",



mancata risposta TOTALE

si ha quando un'unità campionaria non può essere raggiunta e "misurata" o si rifiuta di collaborare all'indagine

di solito, a sfuggire sono le unità "estreme": sono selezionate e diverse dalle unità rispondenti e questo comporta una distorsione nelle stime

> le famiglie non raggiunte nelle indagini campionarie generalmente appartengono a strati sociali disagiati o, al contrario, molto agiati

le indagini sulle famiglie Istat campionano, di solito una famiglia italiana su 1.000 se ne "cade" una è come se si perdessero le risposte di 1.000 famiglie italiane



mancata risposta PARZIALE

si ha quando un'unità campionaria che accetta di rispondere salta un singolo quesito per errore o perché si rifiuta di fornire qualche risposta delicata

questi errori dipendono anche dal modo in cui sono costruiti gli strumenti di rilevazione: i questionari

i quesiti delicati sul reddito vengono sempre posti alla fine del questionario perché rischiano di compromettere il rapporto sereno di comunicazione tra intervistatore e intervistato e di compromettere la qualità delle altre risposte



l'errore di MISURA

nelle indagini presso le famiglie lo strumento di misura non è il metro ma il questionario

deve essere usato con sapienza e precisione rispettandolo in ogni suo elemento senza stravolgerlo o personalizzarlo

perché nelle intenzioni dei ricercatori ogni quesito ha un suo senso e significato esattamente nel modo in cui è formulato e nella posizione in cui si trova



il QUESTIONARIO

è uno strumento di misura e di comunicazione

influenza notevolmente la qualità dei risultati

di solito si cerca di prevenirne gli errori progettando i questionari in modo da:

- stabilire accuratamente l'ordine dei quesiti
- utilizzare un linguaggio facilmente comprensibile
- porre le domande in modo neutro senza influenzare le risposte con giudizi
- evitare eccessivo ricorso a domande aperte

l'effetto ricordo ... l'effetto "telescoping" ... la desiderabilità sociale delle risposte ...



e ora qualche risultato ...





secondo un'indagine Istat svolta nel 2007

Quasi la metà della popolazione di 3 anni e più (41,7%) utilizza il personal computer Mentre il 36,9% della popolazione di 6 anni e più naviga su Internet

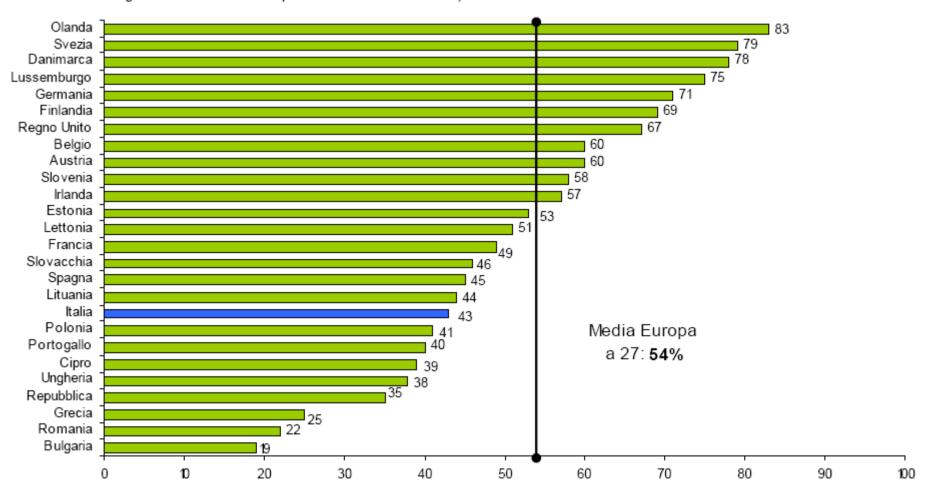
Il picco di utilizzo del personal computer si ha tra i 15 e i 19 anni (oltre il 77%) e per Internet tra i 15 e i 24 anni (oltre il 68%)

Possesso di Internet a casa:

l'Italia è indietro in Europa

Fonte: Eurostat, Indagine comunitaria sulla diffusione delle ICT presso le famiglie e gli individui

Famiglie con almeno un componente tra i 16 e i 64 anni che possiedono un accesso ad Internet da casa. Anno 2007 (per 100 famiglie con almeno un componente tra i 15 e i 64 anni)







secondo un'indagine Istat del 2006 il 77,4 % delle persone di 6 anni e più utilizza il telefono cellulare.

L'uso del cellulare è influenzato dall'età.

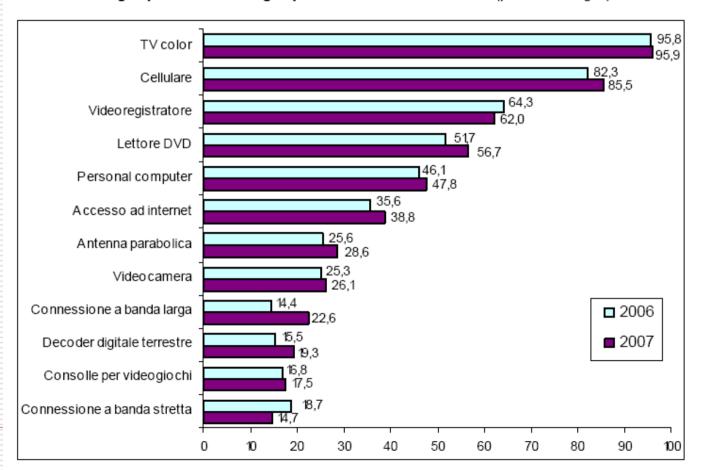
Le persone che lo usano sono, infatti, il 28,7 % dai 6 ai 10 anni, raggiungono l'80,0 % già tra dagli 11 ai 14 anni, per poi superare il 93 % dai 15 ai 44 anni





Aumentano i beni tecnologici posseduti dalle famiglie

Famiglie per beni tecnologici posseduti. Anni 2006 e 2007 (per 100 famiglie)







secondo un'indagine Istat del 2006 sono 39 milioni hanno ascoltato la radio almeno una volta nell'anno, pari al 68,8% delle persone di 3 anni

Le trasmissioni più seguite alla radio sono

la musica leggera 63,3 %

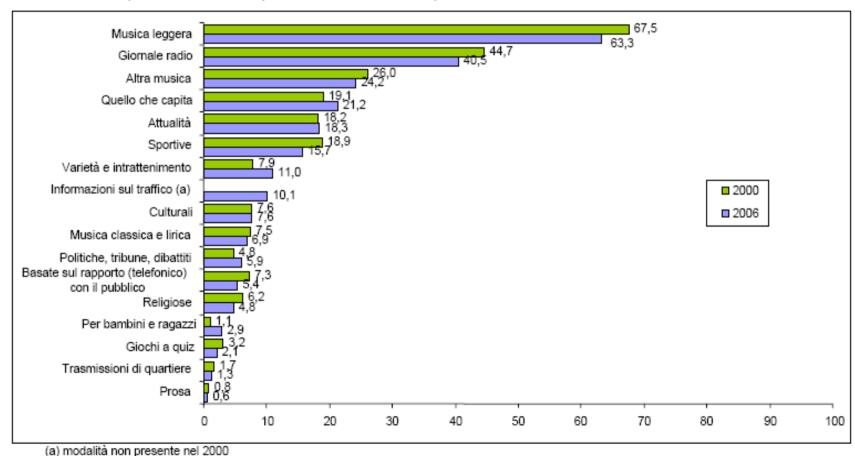
il giornale radio 40,5 %

gli altri tipi di musica 24,2 %...

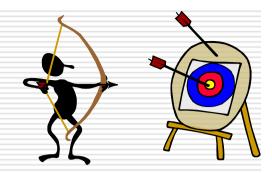




Persone di 3 anni e più che ascoltano la radio per tipo di trasmissione seguita - Anni 2000 e 2006 (per 100 persone di 3 anni e più che ascoltano la radio)







secondo un'indagine Istat svolta nel 2006

il 30,2% della popolazione di 3 anni e più pratica sport con continuità

il 10,1% lo pratica saltuariamente

il Nord-est è la ripartizione geografica con la quota più elevata di persone che praticano sport con continuità (36,5%)





Persone di 6 anni e più secondo il linguaggio abitualmente usato in diversi contesti relazionali. Italia e Veneto. Anno 2000 (valori percentuali)

In famiglia	solo o prevalentemente italiano solo o prevalentemente dialetto sia italiano sia dialetto altra lingua	Italia 44,1 19,1 32,9 3,0	Veneto 22,6 42,6 ★ 29,8 3,9
Con gli amici	solo o prevalentemente italiano solo o prevalentemente dialetto sia italiano sia dialetto altra lingua	48,0 16,0 32,7 2,4	23,7 38,2 ** 34,4 2,7
Con gli estranei	solo o prevalentemente italiano solo o prevalentemente dialetto sia italiano sia dialetto altra lingua	72,7 6,8 18,6 0,8	52,4 14,2 ** 32,0 0,2



ci sono anche alcune statistiche che vi possono aiutare a orientarvi nelle vostre scelte formative

- percorsi di studio e di lavoro dei diplomati

inserimento professionale dei laureati



■ il giudizio dei cittadini sulla scuola





. . .

grazie per l'attenzione!