

STATISTICA SOCIALE

corso di laurea in
SCIENZE TURISTICHE

A.A. 2007/2008

Docente: *Carla Rampichini*

carla@ds.unifi.it

Programma del corso

- ✶ La statistica e i suoi obiettivi.
- ✶ Rilevazione dei dati.
- ✶ Variabili statistiche univariate.
- ✶ Il sistema statistico del turismo.
- ✶ Principali indicatori del fenomeno turistico.
- ✶ L'indagine ad hoc.
- ✶ Indagini campionarie ISTAT in materia di viaggi e vacanze.

Testi consigliati

Pasetti P. (2002) *Statistica del turismo*, Roma: Carocci

Testi di utile consultazione

- ✶ Bailey K.D. (1995) *Metodi della ricerca sociale*, Bologna: Il Mulino. Capitoli V-VIII e XIII-XV.
- ✶ Fortini M. (2000), *Linee guida metodologiche per rilevazioni statistiche*, ISTAT.
- ✶ Lavine D. M. et al. (2002) *Statistica*, Milano: Apogeo.
- ✶ Tassinari G. (1997), *Appunti di Statistica del turismo*, Bologna: CLUEB.

Esame e prove intermedie

Esame: prova scritta + prova orale

Per gli studenti frequentanti le **prove intermedie** costituiscono parte integrante dell'esame di profitto.

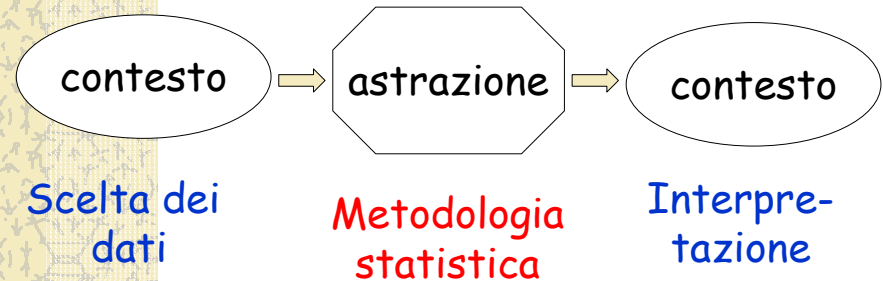
Pagina WEB del corso:

www.ds.unifi.it/rampi/statistica_sociale2006_07.htm

Cos'è la Statistica?

- ✶ Trattamento dell'informazione in forma quantitativa.
- ✶ Analisi e interpretazione dei dati attraverso concetti e strumenti di tipo analitico.

Cos'è la Statistica?

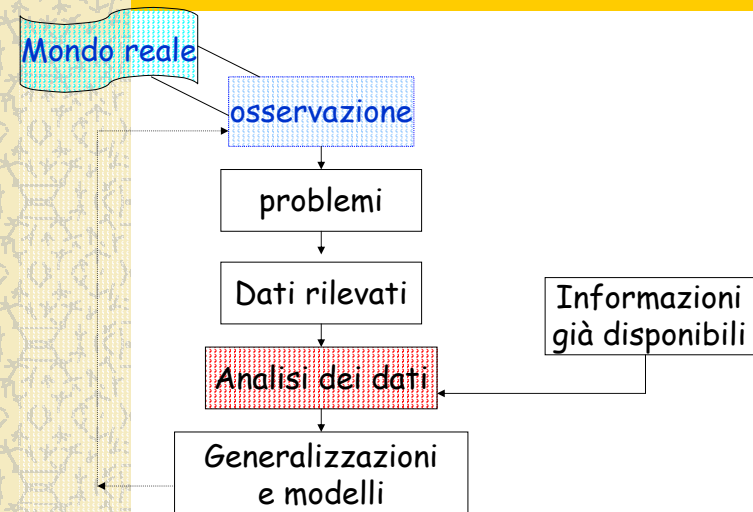


Apprendimento e valutazione del metodo prescindono dal tipo di applicazione

STATISTICA

- ✶ sec. XVII e XVIII: misurazione “ricchezza” e “potenza” dello **Stato**;
- ✶ Nasce anche Calcolo delle Probabilità.
- ✶ **Statistica descrittiva (Censimenti) verso la seconda metà del XIX secolo;**
- ✶ Inferenza Statistica inizi '900 (Fisher 1920), sviluppi teorici dopo 2a guerra mondiale.

Procedimento di ricerca



Statistica e analisi dei dati

Punto di partenza: matrice dei dati

Controllo dei dati

- ✎ Valori anomali o errati;
- ✎ Dati mancanti.

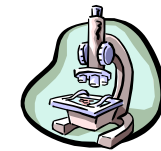
Elaborazione delle informazioni

- ✎ Relazioni tra caratteri;
- ✎ Somiglianza tra caratteri e/o unità;
- ✎ Rappresentazione caratteri e/o unità in uno spazio ridotto.

STATISTICA

Matematica applicata volta allo studio di **dati** ottenuti per:

(a) sperimentazione;



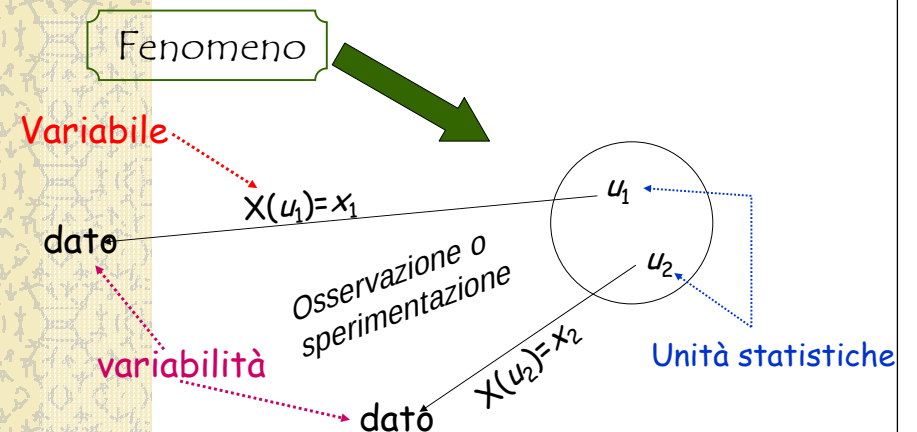
(b) osservazione.



Alcuni termini 'statistici'

- ✎ **Popolazione di riferimento**: insieme di individui che interessa studiare
- ✎ **Unità statistiche**: elementi della popolazione di riferimento
- ✎ **Dati**: risultato della rilevazione-misurazione di caratteristiche delle unità statistiche
- ✎ **Variabili**: caratteristiche rilevate/misurate sulle unità statistiche
- ✎ **Modalità**: valori distinti assunti da una variabile
- ✎ **Campione**: sottoinsieme della popolazione oggetto della rilevazione

Unità statistiche, popolazioni, variabili



Esempio: analisi del comportamento dei turisti stranieri

- Popolazione oggetto di indagine: turisti stranieri presenti nel periodo giugno-agosto 2004
- Variabili di interesse:
 - Giorni di permanenza secondo l'esercizio ricettivo
 - Km percorsi
 - Spesa turistica complessiva



Tipologie di dati

- Dati di **stato** (stock): si riferiscono ad un preciso istante di tempo (p.e. popolazione censita al 31-12-2004)
- Dati di **flusso**: riferiti ad un intervallo di tempo (p.e. presenze turistiche a Firenze durante il 2004). Sono i più frequenti nelle statistiche del turismo.

Dati di stock: esercizi ricettivi

Prospetto 1 - Capacità degli esercizi ricettivi (a) - Anni 1999-2002 (valori assoluti)

ANNI	Esercizi alberghieri		Esercizi complementari Bed and breakfast (b)		Totale esercizi	
	Numero	Letti	Numero	Letti	Numero	Letti
1999	33.341	1.807.275	35.856	1.816.616	69.197	3.623.891
2000	33.361	1.854.101	83.858	2.055.897	117.219	3.909.998
2001	33.421	1.891.281	94.860	2.133.049	128.281	4.024.330
2002	33.411	1.929.544	75.966	2.150.643	4.338 113.715	4.099.585

(a) (b) I dati si riferiscono alla consistenza di fine anno 2002. A partire dalla rilevazione relativa all'anno 2002 è stata inserita una nuova sezione "Alloggi privati in affitto" nell'ambito della quale sono contemplati i bed and breakfast.

I FLUSSI NELLE STRUTTURE RICETTIVE

Nel 2004 gli alberghi e le strutture complementari operanti in Italia hanno registrato **82 milioni e 968 mila arrivi**, per un totale di **336 milioni e 843 mila presenze**. Rispetto al 2003 gli arrivi sono aumentati dello 0,3% mentre le giornate di presenza sono diminuite del 2,2%.

Tabella 1 – Arrivi e presenze di italiani e stranieri negli esercizi ricettivi. Anni 2003-2004
(valori in migliaia e variazioni percentuali)

		2003		2004	
		Numero	Variaz. % su 2002	Numero	Variaz. % su 2003
MOVIMENTO CLIENTI					
	ITALIANI				
	Arrivi	47.719	4,5	47.498	-0,5
	Presenze	204.760	2,5	199.618	-2,5
	Perman. media	4,29	-1,8	4,20	-2,1
STRANIERI	Arrivi	35.006	-3,7	35.470	1,3
	Presenze	139.653	-4,1	137.225	-1,7
	Perman. media	3,99	-0,3	3,87	-3,0
TOTALE	Arrivi	82.725	0,8	82.968	0,3
	Presenze	344.413	-0,2	336.843	-2,2
	Perman. media	4,16	-1,2	4,06	-2,4

Fonte: Istat, Rilevazione sul movimento dei clienti negli esercizi ricettivi – Dati 2004 provvisori.

Come si ottengono le informazioni?

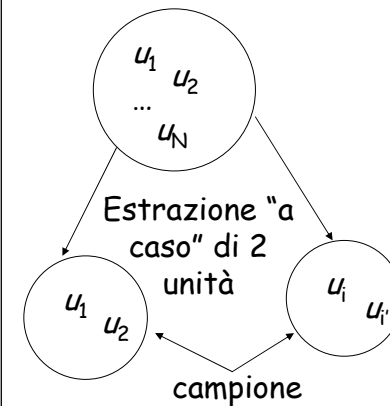
- Consultazione di esperti (es. metodo Delphi)
- Analisi di dati esistenti (fonti statistiche)
- Rilevazioni statistiche *ad hoc*

Esperimenti di campionamento e campioni

Popolazione reale
N "grande"

Esperimento di campionamento casuale:

- senza reinserimento
- semplice (con reinserimento)



Il campione non è una riproduzione fedele della popolazione



Inferenza statistica

Esempio di esperimento campionario

Quante persone scrivono con la mano sinistra?

N molto grande; M= n. di "mancini"

?? $p=M/N$??

Campione di dimensione $n=100$

$$\hat{p} = 4/100 = 0.04$$

$p \neq \hat{p}$ per errore di campionamento

Quanto $p \neq \hat{p}$? Inferenza statistica

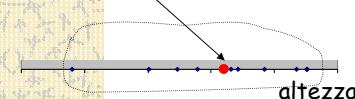
$p \in [0.01; 0.08]$

Modelli statistici descrittivi e inferenziali

La statistica studia la variabilità attraverso 'modelli' matematici

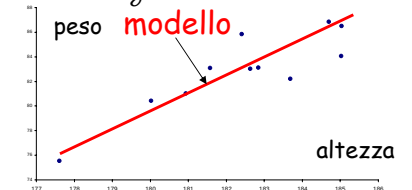
1 variabile (es. altezza)

Quanto ci si discosta da un valore 'tipico' modello



2 variabili (altezza-peso)

Quanto ci si discosta da una relazione sistematica



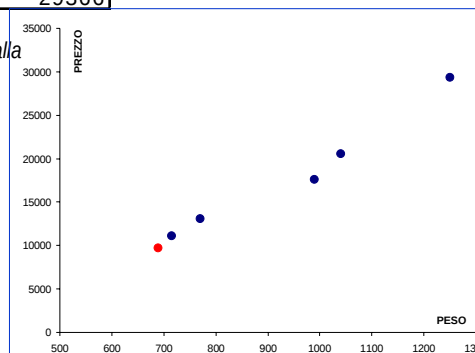
Esempio: le automobili si vendono a peso?

Peso e prezzo dei modelli base della FIAT, marzo 1993

Autovettura	PESO (kg)	PREZZO (mil lire)
Cinquecento 700 ED	690	9697
Panda 1.0 i.e. L.	715	11071
Uno Fire 1.0 i.e. 3P	770	13041
Tipo 1.4 i.e.	990	17580
Tempra 1.4 e.e.	1040	20549
Croma 2.0	1250	29366

Fonte: Gente Motori, febbraio 1993

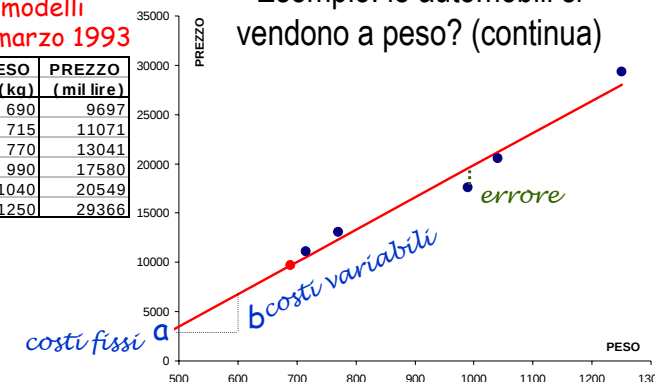
PACE L. e SALVAN A., 1996, *Introduzione alla Statistica - I Statistica Descrittiva*, CEDAM, Padova



Peso e prezzo dei modelli base della FIAT, marzo 1993

Autovettura	PESO (kg)	PREZZO (mil lire)
Cinquecento 700 ED	690	9697
Panda 1.0 i.e. L.	715	11071
Uno Fire 1.0 i.e. 3P	770	13041
Tipo 1.4 i.e.	990	17580
Tempra 1.4 e.e.	1040	20549
Croma 2.0	1250	29366

Esempio: le automobili si vendono a peso? (continua)



$$\text{PREZZO} = a + b \cdot \text{PESO} + \text{ERRORE}$$

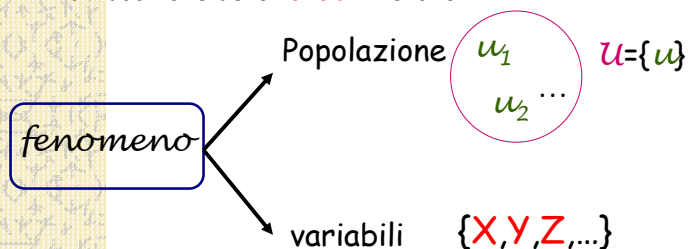
parte sistematica

scostamento dovuto ad altri 'fattori'

Definizione e rilevazione dei dati (1)

Definizione dei dati

- Individuazione del fenomeno
- Individuazione della popolazione di riferimento e delle unità che la compongono
- Individuazione delle variabili rilevanti



Rilevazione dei dati

Piano di rilevazione

$U = \{u\}$ nota

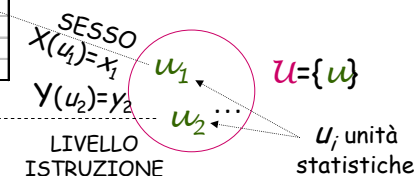
- Predisposizione lista unità da rilevare
- Predisposizione del questionario
- Istruzione dei rilevatori (intervistatori)

Rilevazione materiale

unità u	etichetta	SESSO	ETA' (a.c.)	LIVISTR	DIST(KM)
Alpio	1	M	28	2	5
Caio	2	M	17	4	7.5
Prima	3	F	20	4	12
Velio	4	M	32	2	3.2
Rufa	5	F	16	1	-
Sesto	6	M	34	2	12.3
Beowulf	7	M	18	1	25
Sebaste	8	F	25	2	7.7

SPOGLIO

variabili



Esempio matrice dei dati

Etichette di identificazione

VARIABILI

unità u	etichetta	SESSO	ETA' (a.c.)	LIVISTR	DIST(KM)
Alpio	1	M	28	2	5
Caio	2	M	17	4	7.5
Prima	3	F	20	4	12
Velio	4	M	32	2	3.2
Rufa	5	F	16	1	-
Sesto	6	M	34	2	12.3
Beowulf	7	M	18	1	25
Sebaste	8	F	25	2	7.7

Modalità delle variabili

LIVISTR: 1=LIC. Elem., 2=Lic. Media,
3=Diploma, 4=Laurea

Ogni riga della matrice rappresenta una
unità statistica

segue esempio matrice dei dati

unità u	etichetta	SESSO	ETA' (a.c.)	LIVISTR	DIST(KM)
Alpio	1	M	28	2	5
Caio	2	M	17	4	7.5
Prima	3	F	20	4	12
Velio	4	M	32	2	3.2
Rufa	5	F	16	1	-
Sesto	6	M	34	2	12.3
Beowulf	7	M	18	1	25
Sebaste	8	F	25	2	7.7

Importanza
delle
etichette

Controllare
coerenza dei
dati!

Dato mancante
(missing)

Attenzione alla qualità dei dati!!

Dal questionario alla matrice dei dati

anno	gener	scuola	anno_cc	CFU	soddis	vaca	organ	durata	stato	mezzo	tipo	allogg	spesa
82	1	1	2	30	3	1	2	7	1	3	1	12	500
85	2	5	2	33	3	1	2	7	2	1	1	1	650
85	2	5	2	33	3	1	2	7	2	1	1	1	650
69	2	6	.	27	3	1	2	14	2	1	1	1	1500

La matrice dei dati

$$X = \{x_{ij}\}$$

Raccoglie i risultati codificati di una rilevazione statistica

p colonne $\rightarrow$ variabili $x_j, j=1,2,\dots,p$

n righe $\rightarrow$ unità statistiche $x_i, i=1,2,\dots,n$

x_{ij} valore della variabile x_j misurata sull'unità x_i

Variabili e modalità

Variabile

caratteristica delle unità statistiche che al variare delle unità può assumere almeno due valori

Modalità

Valori assumibili da una variabile (generalmente note a priori). L'insieme di tali valori è detto INSIEME delle MODALITÀ della variabile

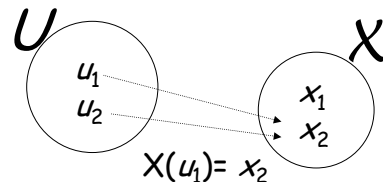
Notazione

$X, Y, Z \dots$
 $X_1, X_2, \dots, X_p$

$x, y, z, \dots$ modalità

X, Y, Z insieme delle modalità

$$X(u): U \rightarrow X$$



CLASSIFICAZIONE DELLE VARIABILI

Variabile qualitativa

Modalità espresse da nomi, aggettivi, attributi

Variabile quantitativa

Modalità espresse da numeri



SCALE DI MISURAZIONE DEI CARATTERI

• SCALA NOMINALE

$$x_i = x_j, \quad x_i \neq x_j$$

• SCALA ORDINALE

$$x_i = x_j, \quad x_i < x_j, \quad x_i > x_j$$

$$x_i > x_j \text{ e } x_j > x_k \Rightarrow x_i > x_k, \quad \text{proprietà transitiva}$$

• SCALA D'INTERVALLI

Fissare unità di misura e **origine** del sistema

$$(x_i - x_j) = (x_k - x_h), \quad (x_i - x_j) > (x_k - x_h), \quad (x_i - x_j) < (x_k - x_h)$$

• SCALA DI RAPPORTI

Fissare unità di misura, 0=assenza fenomeno

$$(x_i / x_j) = (x_k / x_h), \quad (x_i / x_j) > (x_k / x_h), \quad (x_i / x_j) < (x_k / x_h)$$

Trasformazione di variabili quantitative

I numeri corrispondenti alle modalità di variabili quantitative possono essere **trasformati**: x_i equivalenti a x_i^*

- Trasformazione lineare crescente per **scala di intervalli**

$$x_i^* = \underset{\text{origine}}{a} + \underset{\text{unità di misura}}{b} x_i$$

- Trasformazione lineare omogenea crescente per **scala di rapporti**

$$x_i^* = \underset{\text{unità di misura}}{b} x_i$$

Proprietà delle trasformazioni

Additività (vale solo per scale di rapporti!)

✶ Scala di rapporti

$$x_1^* + x_2^* = bx_1 + bx_2 = b(x_1 + x_2) = (x_1 + x_2)^*$$

✶ Scala di intervalli

$$x_1^* + x_2^* = (a + bx_1) + (a + bx_2) = 2a + b(x_1 + x_2) \neq (x_1 + x_2)^*$$

Invarianza

✶ Scala di rapporti

$$\frac{x_1^*}{x_2^*} = \frac{bx_1}{bx_2} = \frac{x_1}{x_2}$$

✶ Scala di intervalli

$$\frac{x_1^* - x_2^*}{x_3^* - x_4^*} = \frac{a + bx_1 - a - bx_2}{a + bx_3 - a - bx_4} = \frac{x_1 - x_2}{x_3 - x_4}$$

Trasformazioni e ricodifiche

Quantitativa continua $Y=(100,210, 210)$ $\longleftrightarrow$ Quantitativa discreta $X=(101, 102, \dots)$

$X=101$ se Y in $(100,101]$

$\downarrow$

$Z=\text{basso}$ se Y in $(100,150]$
 $Z=\text{medio}$ se Y in $(150,175]$
 $Z=\text{alto}$ se Y in $(175,210]$

Attenzione: perdita di informazione!!!

Qualitativa ordinale $Z=(\text{basso}, \text{medio}, \text{alto})$ $\longleftrightarrow$ Qualitativa nominale $W=(\text{normale}, \text{altro})$

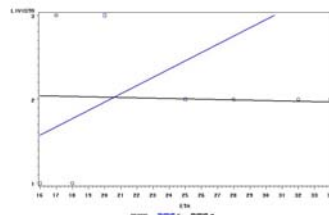
$W=\text{normale}$ se $Z=\text{medio}$
 $W=\text{altro}$ se $Z=(\text{alto}, \text{basso})$

Le analisi statistiche

Univariate

studio della **variabilità marginale**
l'età media è 23.75 anni...

unità u	etichetta	SESSO	ETA' (a.c.)	LIVISTR	DIST(KM)
Alpio	1	M	28	2	5
Caio	2	M	17	4	7.5
Prima	3	F	20	4	12
Velio	4	M	32	2	3.2
Rufa	5	F	16	1	-
Sesto	6	M	34	2	12.3
Beowulf	7	M	18	1	25
Sebaste	8	F	25	2	7.7



Bivariate

l'età media dei maschi è più elevata di quella delle femmine

Multivariate

Al crescere dell'età il livello di istruzione aumenta per le femmine e diminuisce per i maschi

studio della **variabilità congiunta**

Tabelle di frequenza

Un esempio di matrice dei dati

Unità u	Etichetta	SESSO	ETA'	LIVISTR	DIST
Alpio	1	M	28	2	5
Caio	2	M	17	4	7.5
Prima	3	F	20	4	12
Velio	4	M	32	2	3.2
Rufa	5	F	16	1	-
Sesto	6	M	34	2	12.3
Beowulf	7	M	18	1	25
Sebaste	8	F	25	2	7.7

Dati in forma grezza

M, M, F, M, F, M, M, F

Tabelle di frequenza

Distribuzione per sesso

SESSO	frequenza
M	5
F	3
Totale	8

Distribuzione per livello di istruzione

LIVISTR	frequenza
1 Lic. elementare	2
2 Lic. Media	4
3 Maturità	0
4 Laurea	2
Totale	8

Definizione Se X è una variabile e x_j è una modalità di X si dice **frequenza assoluta** di x_j il n. di volte che x_j risulta osservata. Si indica con n_j

Definizione L'insieme delle coppie (x_j, n_j) si dice **distribuzione di frequenza assoluta** o **serie statistica**

Indagine sul turismo montano

Alcune delle caratteristiche rilevate

- ✎ Et  in anni compiuti alla data dell'intervista
- ✎ Sesso
- ✎   la prima volta che viene in questa localit ?
(N=no, S=sì)
- ✎ Tipo di alloggio
- ✎ Numero di pernottamenti
- ✎ Pratica dello sci
- ✎ Spesa complessiva pro-capite in migliaia di lire

Dal questionario alle tabelle

Matrice dei dati da indagine sul turismo montano (944 turisti)

Obs	ETA	SESSO	PRIMA	ALLOGGIO	PERNOTTO	SCI	SPESA
1	32	F	S	3	4	2	200
2	38	M	S	3	10	4	400
3	30	F	N	2	3	3	250
4	33	F	N	11	2	1	30
5	40	F	N	6	7	4	300
6	35	M	N	7	2	3	100
7	25	M	N	6	2	3	100
8	23	F	N	3	7	3	300
9	51	M	N	2	6	1	600
10	23	F	S	11	3	3	70
...							
943	28	M	N	3	3	4	0
944	29	M	S	3	14	3	900

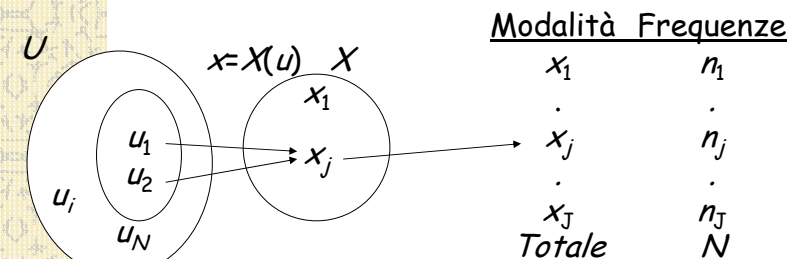
Operazioni di spoglio: dalla matrice dei dati alle tabelle

Turisti per sesso

SESSO	Frequenza	Freq. %
F	470	49.79
M	474	50.21

Distribuzioni di frequenza

n_j numero di unit  statistiche che presentano la modalit  $X = x_j$



$|U| = N = \sum n_j$ *numerosit  della popolazione*

Frequenze assolute e relative

Si dice FREQUENZA RELATIVA di una modalità x_j , o di una classe di modalità $(x_{j-1}; x_j)$ e si indica con f_j , la frazione o proporzione di u.s. che presentano tale modalità.

$$f_j = \frac{n_j}{\sum_{j=1}^J n_j} = \frac{n_j}{N} \quad j = 1, 2, \dots, J$$

$$0 \leq f_j \leq 1 \quad j = 1, 2, \dots, J$$

Proprietà:

$$\sum_{j=1}^J f_j = 1$$

Finalità delle frequenze relative

Facilitare la percezione del **PESO** delle modalità

Sesso	Freq. Assoluta	Freq. Relativa	Freq. Rel. %
M	1750	0.583	58.3
F	1250	0.417	41.7
Totale	3000	1	100

Facilitare **CONFRONTI** tra popolazioni

Sesso	Freq. Assoluta		Freq. Rel. %	
	Pop. A	Pop. B	Pop. A	Pop. B
M	1750	850	58.3	85.0
F	1250	150	41.7	15.0
Totale	3000	1000	100	100

Pratica dello sci: decodifica delle modalità

SCI	Frequenza	Freq. %
SOLO ALPINO	540	57.82
SOLO FONDO	88	9.42
ENTRAMBI	134	14.35
NESSUNO	172	18.42

Valori mancanti = 10

Per una migliore lettura: decodifica delle modalità!

Tipo di alloggio

ALLOGGIO	Frequenza	Freq. %
1	46	4.89
2	86	9.15
3	166	17.66
4	80	8.51
5	14	1.49
6	182	19.36
7	200	21.28
9	4	0.43
10	2	0.21
11	56	5.96
12	30	3.19
13	2	0.21
14	18	1.91
15	52	5.53
20	1	0.11
40	1	0.11

Molte modalità: per una migliore lettura decodifica e raggruppamento delle modalità!

Valori mancanti = 4

Tipo di alloggio

ALLOGGIO	Frequenza	Freq. %
ALBERGHIERO	378	40.21
EXTRA - ALBERGHIERI	230	24.47
CASA PROPRIETA'	260	27.66
ALTRO	18	1.91
PENDOLARI	52	5.53
	20	0.11
	40	0.11

20 e 40 codici errati!!

Frequency Missing = 4

Decodifica e raggruppamento delle modalità

NB: decodifica consente di individuare codici errati.

Frequenze cumulate e funzione di ripartizione

🔧 **Frequenze cumulate:** somma delle frequenze sino alla modalità considerata

🔧 **Funzione di ripartizione** della v.s. quantitativa X
 $F(X_j)$ = proporzione di u.s. con valori di $X \leq X_j$

$$F(x_j) = \sum_{k=1}^j f_k = \Pr\{X \leq x_j\}$$

Età all'intervista

ETA	Frequenza	Freq. %	Freq. Cumulata	Freq. % Cumulata
5	2	0.21	2	0.21
14	2	0.21	4	0.42
15	6	0.64	10	1.06
16	18	1.91	28	2.97
17	16	1.69	44	4.66
18	24	2.54	68	7.20
19	24	2.54	92	9.75
...				
66	2	0.21	938	99.36
72	2	0.21	940	99.58
73	4	0.42	944	100.00

Per una migliore lettura: definizione delle classi di modalità!

Età all'intervista: raggruppamento in classi

ETA	Frequenza	Freq. %	Frequenza Cumulata	Freq. % Cumulata
0 - 13	2	0.21	2	0.21
14 - 29	358	37.92	360	38.14
30 - 59	560	59.32	920	97.46
≥ 60	24	2.54	944	100.00

Per una migliore lettura: definizione delle classi di modalità!

Età studenti a.a. 2007/08

età a.c.

ETA	Frequency	Percent	Cumulative Frequency	Cumulative Percent
20	54	53.47	54	53.47
21	21	20.79	75	74.26
22	17	16.83	92	91.09
23	3	2.97	95	94.06
25	1	0.99	96	95.05
26	1	0.99	97	96.04
29	1	0.99	98	97.03
39	1	0.99	99	98.02
40	1	0.99	100	99.01
46	1	0.99	101	100.00

Frequency Missing = 1

Soddisfazione studenti a.a. 07/08

sei soddisfatto dei corsi seguiti l'anno scorso?

Soddisfatto?	Frequenza	Freq. %	Freq.Cum.	Freq.cum. %
Decisamente no	1	1.00	1	1.00
Più no che sì	19	19.00	20	20.00
Più sì che no	56	56.00	76	76.00
Decisamente sì	24	24.00	100	100.00

Frequency Missing = 2

Crediti studenti a.a. 07/08

Crediti al 19/9/07

crediti	Freq	Freq. %	Freq.Cum.	Freq.cum. %
0-24	14	15.91	14	15.91
25-40	34	38.64	48	54.55
41-50	19	21.59	67	76.14
>50	21	23.86	88	100.00

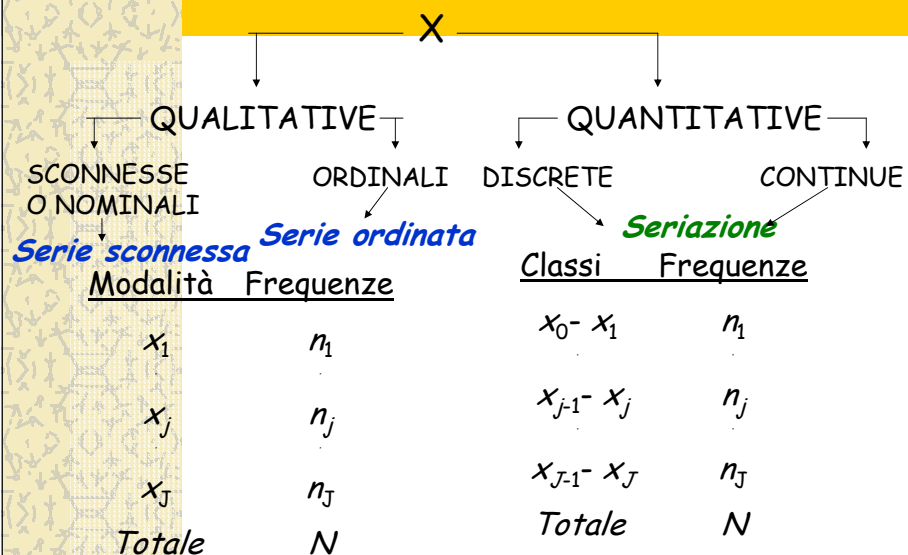
Frequency Missing = 14

[Altre tabelle](#)

Tabella di frequenza con cumulate

Modalità	Freq.	Fr.rel.	Freq.cum	F(x)
x_1	n_1	f_1	n_1	f_1
x_2	n_2	f_2	$n_1 + n_2$	$f_1 + f_2$
...	...	...	...	...
x_j	n_j	f_j	$n_1 + n_2 + \dots + n_j$	$f_1 + f_2 + \dots + f_j$
...	...	...	...	...
x_J	n_J	f_J	N	1
Totale	N	1		

Classificazione delle distribuzioni



Struttura generale

Serie univariata

Modalità	Freq.
x_1	n_1
...	...
x_j	n_j
...	...
x_J	n_J
Totale	N

Seriazione univariata

Classi	Freq.
$x_0 - x_1$	n_1
...	...
$x_{j-1} - x_j$	n_j
...	...
$x_{J-1} - x_J$	n_J
Totale	N

$j=1, 2, \dots, J$ modalità; $N=n$ unità statistiche

Tabelle di frequenza per analisi bivariate e multivariate

Dati in forma grezza

$(M,2), (M,4), (F,4), \dots, (F,2)$

Un esempio di matrice dei dati

Unità u	Etichetta	SESSO	ETA'	LIVIST	DIST
Alpio	1	M	28	2	5
Caio	2	M	17	4	7.5
Prima	3	F	20	4	12
Velio	4	M	32	2	3.2
Rufa	5	F	16	1	-
Sesto	6	M	34	2	12.3
Beowful	7	M	18	1	25
Sebaste	8	F	25	2	7.7

Tabelle di frequenza

Distribuzione per sesso e livello di istruzione

SESSO	LIVELLO DI ISTRUZIONE			TOTALE
	Lic. Elem.	Lic. Media	Laure	
M	1			5
F	1	1	1	3
TOTALE	2	4	2	8

Tabella a doppia entrata o serie bivariata

Tabelle a doppia entrata (bivariate)

Classificazione congiunta delle u.s. in base a coppie di caratteri

n_{ij} frequenze congiunte: nr u.s. con $X=x_i$ e $Y=y_j$

$X Y$	y_1	...	y_j	...	y_c	Tot
x_1	n_{11}	...	n_{1j}	...	n_{1c}	$n_{1.}$
...	...	...	...	...	...	...
x_i	n_{i1}	...	n_{ij}	...	n_{ic}	$n_{i.}$
...	...	...	...	...	...	...
x_r	n_{r1}	...	n_{rj}	...	n_{rc}	$n_{r.}$
Tot	$n_{.1}$	...	$n_{.j}$	...	$n_{.c}$	N

Tabella bivariata: pratica dello sci per titolo di studio

SCI(pratica dello sci)		STUDIO(titolo di studio)				
Frequenza	nessuno	element.	media	diploma	laurea	Totale
SOLO ALPINO	0	16	104	296	124	540
SOLO FONDO	0	0	18	42	28	88
ENTRAMBI	2	2	14	70	46	134
NESSUNO	0	13	37	88	34	172
Totale	2	31	173	496	232	934

Missing = 10

Tabella bivariata: tipo di alloggio per classe di età

ALLOGGIO	ETA				Totale
	0-13	14-29	30-59	60 E +	
ALBERGHIERO	0	138	238	2	378
EXTRA-ALBERG.	0	96	130	4	230
CASA PROPRIETA'	2	100	144	14	260
ALTRO	0	8	8	2	18
PENDOLARI	0	14	36	2	52
Totale	2	356	556	24	938

Missing = 4

Frequenze relative congiunte pratica dello sci per titolo di studio

SCI(pratica dello sci)		STUDIO(titolo di studio)				
Freq. Rel. %	nessuno	element.	media	diploma	laurea	Totale
SOLO ALPINO	0.00	1.71	11.13	31.69	13.28	57.82
SOLO FONDO	0.00	0.00	1.93	4.50	3.00	9.42
ENTRAMBI	0.21	0.21	1.50	7.49	4.93	14.35
NESSUNO	0.00	1.39	3.96	9.42	3.64	18.42
Totale	0.21	3.32	18.52	53.10	24.84	100.00

Frequency Missing = 10

Frequenze relative congiunte: $f_{ij} = n_{ij}/N = Pr(X=x_i, Y=y_j)$

Soddisfazione e voto medio esami studenti a.a 07/08

Soddisf	voto medio esami				Total
	<24	24-<27	27-28	29-30	
Decisamente no	0	1	0	0	1
Più no che sì	4	9	2	0	15
Più sì che no	7	33	11	4	55
Decisamente sì	1	7	10	5	23
Total	12	50	23	9	94

Anno di corso e CFU studenti a.a 07/08

Anno	cfu(Crediti al 19/9/07)				Total
	0-24	25-40	41-50	>50	
2	12	31	18	12	73
	14.63	37.80	21.95	14.63	89.02
3	2	1	0	6	9
	2.44	1.22	0.00	7.32	10.98
Total	14	32	18	18	82
	17.07	39.02	21.95	21.95	100.00

Frequency Missing = 20

La tabella contiene le frequenze congiunte in valore assoluto
e percentuale, p.e. $14.63 = 100 \times 12 / 82 = 100 \times f_{ij}$