



Argomenti della lezione

- La tabulazione dei dati
- Il concetto di frequenza
- Le distribuzioni di frequenza

RILEVAZIONE E SPOGLIO DEI DATI

- Dalla carta al computer
- La tabulazione dei dati

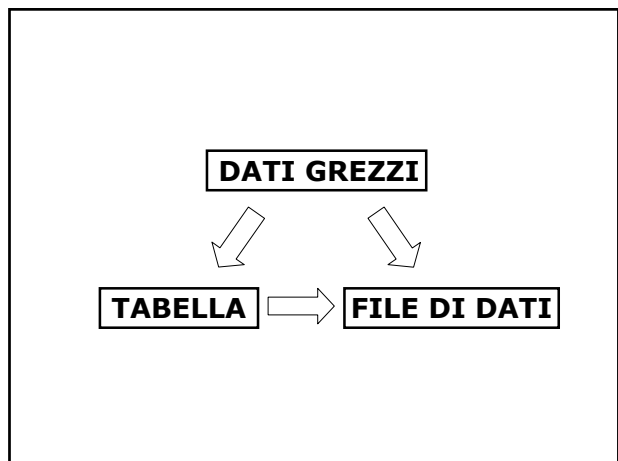
Dalla carta al computer

- Le risposte di ogni singolo soggetto ad un questionario di personalità
- Le sue risposte ad una scala di atteggiamento
- La categorizzazione del comportamento di ogni singolo soggetto

Dalla carta al computer

Costituiscono i cosiddetti **DATI GREZZI**

Per poter elaborare i "dati grezzi" è necessario organizzarli mediante un'operazione che si chiama **TABULAZIONE**



LA MISURA IN PSICOLOGIA

Prof. Anna Paola Ercolani (Università di Roma)

Lez. 8 - Rilevazione, spoglio dei dati e distribuzioni di frequenza

Operazioni preliminari per la tabulazione dei dati

Costruire un legame tra il questionario cartaceo e la tabella o il file di dati

Sigare ciascun questionario con un numero

Operazioni preliminari per la tabulazione dei dati

Stabilire l'ordine secondo il quale le risposte vengono trascritte nella tabella o inserite nel file

Seguire la numerazione che le domande hanno nel supporto cartaceo

Operazioni preliminari per la tabulazione dei dati

Scegliere un nome in codice per le variabili sulle quali si vuole lavorare

ETICHETTA

Se si tratta di un file di dati non può essere più lunga di otto caratteri

Operazioni preliminari per la tabulazione dei dati

Può essere necessario codificare numericamente alcune variabili

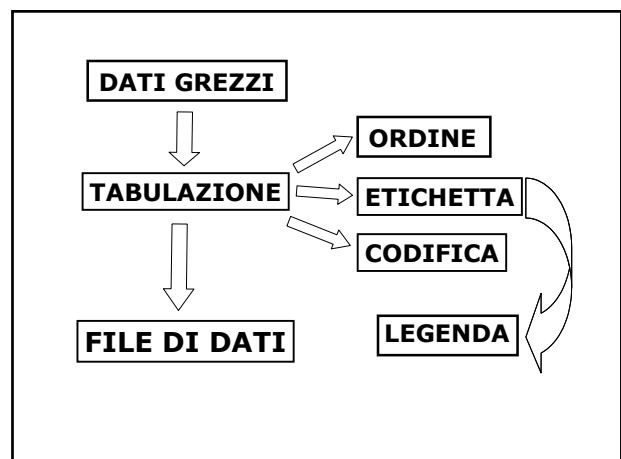
Di solito si attribuisce un numero della scala naturale (1, 2, 3, ...)
Se la variabile prevede solo due alternative conviene usare 0 e 1

Operazioni preliminari per la tabulazione dei dati

E' ovvio che il codice scelto è del tutto arbitrario

Perciò è necessario tenere memoria delle variabili originali e del relativo codice attribuito

LEGENDA



LA MISURA IN PSICOLOGIA

Prof. Anna Paola Ercolani (Università di Roma)
Lez. 8 - Rilevazione, spoglio dei dati e distribuzioni di frequenza

ESEMPIO

Supponiamo di aver somministrato a 10 soggetti una scala di atteggiamento composta da 5 affermazioni (con risposta si/no) e di aver rilevato anche il genere dei soggetti, la loro età e il loro titolo di studio

Numeriamo i questionari da 1 a 10

Decidiamo l'ordine delle variabili:

numero
genere
età
domande da 1 a 5

Stabiliamo un'etichetta per ciascuna variabile:

N.ord
Sex
Aa
D1, D2, ... D5

Stabiliamo un codice per le variabili

Sex (M=1;F=0)

D1 ... D5 (Sì=1; No = 0)

ESEMPIO

Supponiamo che un soggetto maschio di 33 anni, diplomato, abbia risposto alle 5 affermazioni: SI, SI, NO, NO, SI

Codifica

N.ord.	03	D1	1	D4	0
Sex	1	D2	1	D5	1
Aa	33	D3	0		

n. ord	sex	Aa	D1	D2	D3	D4	D5
01	1	33	1	1	1	0	0
02	0	27	0	1	0	1	0
03	1	33	1	1	0	0	1
04	0	40	0	0	1	1	1
05	0	33	1	0	1	1	0
06	1	28	0	0	0	0	1
07	1	40	1	1	1	1	0
08	0	28	0	1	0	1	1
09	1	28	1	1	0	1	1
10	1	30	1	0	0	0	0

LA MISURA IN PSICOLOGIA

Prof. Anna Paola Ercolani (Università di Roma)
Lez. 8 - Rilevazione, spoglio dei dati e distribuzioni di frequenza

Ss	X1	X2	X3	X4
01	15	33	18	15
02	21	27	18	15
03	21	33	22	19
04	10	40	20	22
05	12	33	18	25
06	19	28	30	24
07	14	40	22	24
08	20	28	27	22
09	11	28	22	24
10	11	30	12	25

		VARIABILI							
		1	2	3	...	j	...	p	
S O G G E T T I	1	X_{11}	X_{12}	X_{13}	...	X_{1j}	...	X_{1p}	$\sum_{j=1}^p X_{1j}$
	2	X_{21}	X_{22}	X_{23}	...	X_{2j}	...	X_{2p}	
	...								
	i	X_{i1}	X_{i2}	X_{i3}	...	X_{ij}	...	X_{ip}	$\sum_{j=1}^p X_{ij}$
	...								
	N	X_{N1}	X_{N2}	X_{N3}	...	X_{Nj}	...	X_{Np}	$\sum_{i=1}^N X_{ij}$
		$\sum_{i=1}^N X_{i1}$				$\sum_{i=1}^N X_{ij}$			$\sum_{i=1}^N \sum_{j=1}^p X_{ij}$

IL CONCETTO DI FREQUENZA

ESEMPIO

Ss	01	02	03	04	05	06	07	08	09	10
D3	1	0	0	1	1	0	1	0	0	0

Ss	01	02	03	04	05	06	07	08	09	10
X2	33	27	33	40	33	28	40	28	28	30

DEFINIZIONE

SI INTENDE PER FREQUENZA

Il numero di casi osservati per ciascuna modalità (categoria o valore) della variabile oggetto di studio qualsiasi sia la scala in cui sono stati raccolti

ESEMPIO

Ss	01	02	03	04	05	06	07	08	09	10
D3	1	0	0	1	1	0	1	0	0	0

I 10 soggetti rispondono Si (codice = 1) 4 volte
Quindi Frequenza = 4
f = 4

Il concetto di frequenza

Ss	01	02	03	04	05	06	07	08	09	10
X2	33	27	33	40	33	28	40	28	28	30

X2₁ = 27 ⇨ f₁ = 1
 X2₂ = 28 ⇨ f₂ = 3
 X2₃ = 30 ⇨ f₃ = 1
 X2₄ = 33 ⇨ f₄ = 3
 X2₅ = 40 ⇨ f₅ = 2

}	⇨	f ₁ = 1	}	DISTRIBUZIONE DI FREQUENZA
	⇨	f ₂ = 3		
	⇨	f ₃ = 1		
	⇨	f ₄ = 3		
	⇨	f ₅ = 2		
⇨				$\Sigma f_i = 10 = N$

DISTRIBUZIONI DI FREQUENZA

Gli indici 1, 2, 3, 4 e 5 delle frequenze f stanno a significare che la variabile assume 5 diversi valori

Possiamo indicare con k il numero dei valori diversi che può assumere una variabile e con N il numero dei soggetti

Allora potremo scrivere:

$$N = \sum_{i=1}^k f_i$$

ATTENZIONE

- X_i = valore generico della variabile (o codifica della variabile)
- f = numero delle volte che si presenta ciascun valore di X
- k = numero di valori di X

ESEMPIO

Ss	01	02	03	04	05	06	07	08	09	10
X	8	7	7	9	7	8	6	7	6	8

X₁ = 6 ⇨ f₁ = 2
 X₂ = 7 ⇨ f₂ = 4
 X₃ = 8 ⇨ f₃ = 3
 X₄ = 9 ⇨ f₄ = 1

$K = 4 \quad N = \sum_{i=1}^4 f_i$

FREQUENZE CUMULATE

Si ottengono sommando progressivamente le frequenze della distribuzione

FREQUENZE CUMULATE

f_i	f_{cum}
$f_1 = 2$	2
$f_2 = 4$	6 (2+4)
$f_3 = 3$	9 (6+3)
$f_4 = 1$	10 (9+1)

FREQUENZE RELATIVE O PROPORZIONI

Rapporto tra la frequenza di una modalità e il totale dei casi osservati

$$P_i = \frac{f_i}{N}$$

ESEMPIO

$$P_i = \frac{f_i}{N}$$

f_i	$N=42$
$f_1 = 12$	$P_1 = 12/42 = .2857$
$f_4 = 18$	$P_4 = 18/42 = .4285$

FREQUENZE PERCENTUALI

frequenza relativa moltiplicata per 100

$$P_{i(\%)} = \frac{f_i}{N} \cdot 100$$

ESEMPIO

$$P_{i(\%)} = \frac{f_i}{N} \cdot 100$$

$f_1 = 12$	$N = 42$
$P_1 = (12/42) \times 100 = 28.57 \cong 29\%$	
$f_4 = 18$	$N = 42$
$P_4 = (18/42) \times 100 = 42.85 \cong 43\%$	