

Elementi di Psicometria con Laboratorio di SPSS 1

4-Misure della tendenza centrale
(vers. 1.0b, 20 ottobre 2015)
versione per stampa

Germano Rossi¹

`germano.rossi@unimib.it`

¹Dipartimento di Psicologia, Università di Milano-Bicocca

2015-16

Tendenza centrale

- La tendenza centrale è un'indicazione generica di come sta andando la distribuzione della variabile
- Ci sono diversi indici che “misurano” la tendenza centrale, alcuni poco informativi, altri molto informativi
- Livello nominale: **Moda**
- Livello ordinale: **Mediana**
- Livello intervallo/rapporto: **Media**
- Ricordiamo che ogni livello “eredita” dai livelli precedenti

Tendenza centrale: Moda

- La **Moda** (Mo) è la frequenza più elevata di una distribuzione
- Se c'è una sola moda, la distribuzione si dice **Unimodale**
- Se sono 2, **Bimodale**
- Se sono più di 2, **Multimodale** (ma non si utilizza)

Esempio

$M=17, F=13 \Rightarrow$ **Maschi** perché ha frequenza 17

Se ci sono molte categorie, oppure poche categorie tutte con frequenze simili, la moda non ha molto senso.

Esempio

1 1 1 2 2 2 2 3 3 3 $Mo=2$ (ma non ha molto senso)

Spss: moda

■ Tramite

Analizza |
Statistiche descrittive |
Frequenze...
pulsante

Statistiche

possiamo far
calcolare la moda.

■ poi Continua e OK



Statistiche		
Genere		
→ N	Validi	30
	Mancanti	0
	Moda	1

Tendenza centrale: Mediana

La **mediana** (Mdn) divide la distribuzione a metà (corrisponde, ma non è sempre uguale, a Q_2)

- Se N è *dispari*, la Mdn è il **valore** in **posizione** centrale, corrispondente a $(N + 1)/2$

Esempio

Dati grezzi	5	2	1	3	5	1	4	4	3	1	5
Ordinati	1	1	1	2	3	3	4	4	5	5	5



$$(11 + 1)/2 = 6 \Rightarrow \text{Mdn}=3$$

Tendenza centrale: Mediana

- Se N è *pari*, la Mdn è il **valore** fra le 2 **posizioni centrali** (se esiste) cioè fra $N/2$ e $(N/2) + 1$
- Se i due valori sono *uguali*, quello è il valore della mediana

Esempio

Dati grezzi	5	2	1	3	5	1	4	4	3	1	
Ordinati	1	1	1	2	3	3	4	4	5	5	$N/2 = 10/2 = 5$ e
					↑	↑					
	$(N/2) + 1 = 5 + 1 = 6 \Rightarrow$										Mdn=3

Tendenza centrale: Mediana

- Se i due valori sono *diversi*
 - se la scala è **ordinale**: entrambi costituiscono la mediana
 - se è **quantitativa**: si fa la media fra i due valori

Esempio

Dati grezzi	5	2	1	4	5	1	4	4	3	1	
Ordinati	1	1	1	2	3	4	4	4	5	5	$N/2 = 10/2 = 5$ e
					↑	↑					
	$(N/2) + 1 = 5 + 1 = 6 \Rightarrow \text{Mdn}=3;4 \text{ (ORD)} \Rightarrow 3,5 \text{ (I/R)}$										

Attenzione

Spss (e la maggior parte dei software statistici) fanno **sempre** la media fra i due valori!

Tendenza centrale: Mediana

Esercizio

- 1 Mdn (3,5,7,9,11)
- 2 Mdn (2,3,5,7,9,11,12)
- 3 Mdn (3,4,5,5,6,7)
- 4 Mdn (3,4,5,6,7,8)
- 5 Mdn (4,5,7,9,13)
- 6 Mdn (1,5,7,9,25)

Soluzione

- 1 $N=5$; $\text{pos}=3$; $\text{Mdn}=7$
- 2 $N=7$; $\text{pos}=4$; $\text{Mdn}=7$
- 3 $N=6$; $\text{pos}=3$ e 4 ; $\text{Mdn}=5$
- 4 $N=6$; $\text{pos}=3$ e 4 ; $\text{Mdn}=5;6$ (5,5)
- 5 $N=5$; $\text{pos}=3$; $\text{Mdn}=7$
- 6 $N=5$; $\text{pos}=3$; $\text{Mdn}=7$

- Se aggiungiamo lo stesso numero di valori all'inizio e alla fine di una distribuzione, la Mdn non cambia
- Se cambiano i valori estremi della distribuzione, la Mdn non cambia

Tendenza centrale: media

- Se 4 amici escono a mangiare la pizza e poi pagano in parti uguali... stanno usando la **media**

- Ovvero:
 $(18 + 16.5 + 22 + 17.5)/4$

- Ovvero: $18.5 * 4$

- Qualcuno paga di più e qualcuno di meno, ma, alla fine, il “di più” si annulla con il “di meno”

pizza, bibita e dessert

Marco	18.0
Clara	16.5
Daniela	22.0
Andrea	17.5
Totale	74.0
a testa	18.5

Marco	$18.0 - 18.5 =$	0.5
Clara	$16.5 - 18.5 =$	2.0
Daniela	$22.0 - 18.5 =$	-3.5
Andrea	$17.5 - 18.5 =$	1.0
Totale	$74.0 - 74.0 =$	0

Tendenza centrale: Media [aritmetica]

La **media aritmetica** (\bar{X} , Md, M) è la somma (\sum) di tutti i valori di una distribuzione, divisa per la numerosità (N)

$$\bar{X} = \frac{\sum_{i=1}^N X_i}{N} = \frac{\sum X}{N}$$

Esempio

M(10, 15, 16, 18, 20, 24, 32, 35, 38, 40) = 24.8

$$\frac{10 + 15 + 16 + 18 + 20 + 24 + 32 + 35 + 38 + 40}{10} = \frac{248}{10}$$

Uso della distribuzione di frequenza

Quando si utilizzano delle distribuzioni di frequenza, si hanno i dati in un formato leggermente diverso. Anziché:

1 1 1 2 2 2 2 2 2 3 3 ($\bar{X} = 1.9$)

Possiamo usare una forma tabellare dove per ogni valore (x) indichiamo a fianco la frequenza (f) con cui compare:

Esempio

x	f	fx
1	3	3
2	6	12
3	2	6
Σ	11	21

$$\bar{X} = \frac{\sum fX}{N}$$

$$\bar{X} = \frac{21}{11} = 1.9$$

Tendenza centrale: Media [aritmetica]

Esercizio

1 $M(1,2,3,4,5)$

2 $M(3,4,5,6,7)$

3 $M(2,4,6,8,10)$

Soluzione

1 $(1+2+3+4+5)/5=15/5=3$

2 $(3+4+5+6+7)/5=25/5=5$

3 $(2+4+6+8+10)/5=30/5=6$

1 i numeri da 1 a 5

2 i numeri della prima serie sommati a 2

3 i numeri della prima serie moltiplicati per 2

Proprietà della media 1: Aggiungendo, sottraendo, moltiplicando o dividendo una costante a tutti i dati della distribuzione, anche la media subisce la stessa trasformazione

Proprietà della media 2: Gli scarti dalla media sommano a 0

Spss: media

La media viene visualizzata da Spss in molte procedure. Quelle specifiche sono:

- `Analizza | Statistiche descrittive | Frequenze...` (fra le varie statistiche che è possibile stampare vi è anche la media)
- `Analizza | Statistiche descrittive | Descrittive...` (è la procedura specifica per le statistiche descrittive)
- `Analizza | Statistiche descrittive | Esplora...` (stampa la media come una delle diverse statistiche per capire l'andamento e la distribuzione di una variabile)

Spss: Media con Frequenze...

- Dopo aver scelto le variabili, click-are su e selezionare **Media**
- Quindi, click-are su
- Con variabili quantitative conviene de-selezionare anche **Visualizza tabelle di frequenza**
- oppure in
 - Sopprimi le tabelle con più di n categorie
 - Numero massimo di categorie:
- Poi

Tendenza centrale

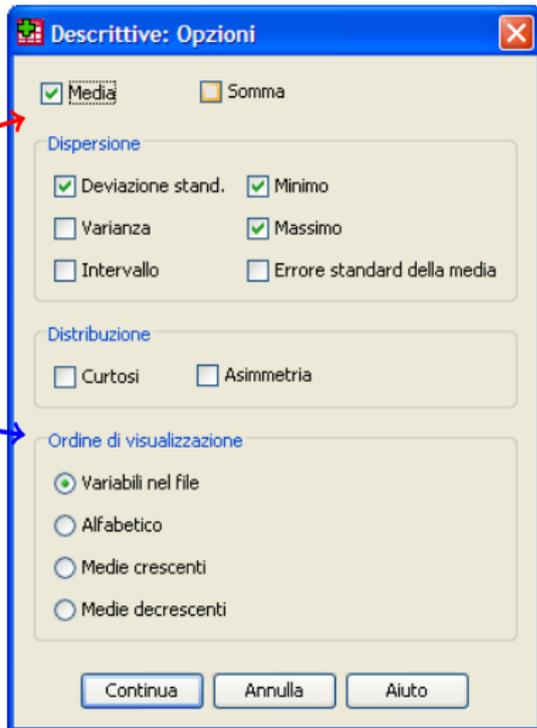
- Media
- Mediana
- Moda
- Somma

Spss: Media con Descrittive...

- Dopo aver scelto le variabili, click-are su 
- Normalmente **Media** è già selezionato
- Potete ordinare i risultati in vari modi
- Poi 

Statistiche descrittive

	N	Media
gruppo	100	,44
Antisemitismo	100	49,6300
Validi (listwise)	100	



Descrittive: Opzioni

Media Somma

Dispersione

Deviazione stand. Minimo

Varianza Massimo

Intervallo Errore standard della media

Distribuzione

Curtosi Asimmetria

Ordine di visualizzazione

Variabili nel file

Alfabetico

Medie crescenti

Medie decrescenti

Continua Annulla Aiuto

Tendenza centrale: Media con dati dicotomici

- Se una variabile è dicotomica (D) ed è stata categorizzata con 0 e 1, **la media di D equivale alla proporzione della categoria 1.**
- Infatti, possiamo pensare a \bar{D} come la somma di tutti gli 0 e la somma di tutti gli 1.

$$\bar{D} = \frac{\sum d_i}{N} = \frac{0 \cdot f_0 + 1 \cdot f_1}{N} = \frac{f_1}{N}$$

- Ma la somma degli 0 è 0 e la somma degli 1 è uguale alla frequenza degli 1.
- Quindi la media di una variabile dicotomica è $\bar{D} = \frac{f_1}{N}$
- L'equivalenza non vale se categorizziamo con numeri diversi da 0 e 1.

Confronto fra statistiche

- **Moda** (Nominale): è il peggior indice
- **Mediana** (Ordinale): non è per nulla sensibile ai valori estremi
- **Media** (Intervallo/Rapporto):
 - è il miglior indice di tendenza centrale
 - ma è molto sensibile ai valori estremi della distribuzione

- In una distribuzione simmetrica normale, *media, mediana e moda coincidono*
- Se la media è minore della mediana la distribuzione è asimmetrica a sinistra
- Se la media è maggiore della mediana, la distribuzione è asimmetrica a destra