

## Indici sintetici di dimensione

- ✘ Obiettivo dell'analisi dei dati : studiare i fenomeni collettivi che si manifestano in modo diverso da unità a unità → *variabilità*
- ✘ Sintesi della distribuzione del carattere con una sola modalità (valore)
  - *Indici sintetici di dimensione: medie*
  - La media dà un'idea sintetica della distribuzione del carattere sul collettivo in esame

## Le medie

- ✘ Esistono molti tipi di *media* a seconda dell'informazione che si vuole fornire e del tipo di situazione che si sta analizzando
- ✘ La *media* deve essere
  - un valore omogeneo con i dati osservati
  - *compreso tra le modalità della distribuzione* (tra le modalità minima e massima se si è in presenza di una mutabile ordinale o di una variabile (principio di Cauchy))
- ✘ Si distinguono:
  - ✓ *medie di posizione*
  - ✓ *medie analitiche*

## Medie di Posizione

- ✘ *Moda e Mediana*
- ✘ Si possono calcolare sia per caratteri qualitativi che quantitativi
- ✘ Modalità che occupano particolari posizioni all'interno della distribuzione del carattere
  - Non necessariamente cambiano se cambiano i dati della distribuzione
  - Sfruttano solo parzialmente l'informazione disponibile

## Moda

- ✘ Si può calcolare sia per *caratteri qualitativi* che *quantitativi*
- ✘ Modalità cui corrisponde la max frequenza (assoluta, relativa o percentuale)
- ✘ Può non essere unica
- ✘ Se il carattere è diviso in classi
  - **Classe modale**: classe di modalità a cui corrisponde la max densità media di frequenza

## Esempio: qual è la moda?

- ✘ Distribuzione dei laureati di SDC nell'a.a. 2003/2004 per Corso di Laurea

CDL	$n_i$
STC	48
SCPO	71
COOP	6
<b>Totale</b>	<b>125</b>

La moda è  
"SCPO"

- ✘ Distribuzione degli studenti di SDC frequentanti la facoltà nell'a.a. 2001/2002 secondo la Soddisfazione della Scelta Universitaria

Soddisfazione	$n_i$	$f_i$	$p_i$
Per nulla	4	0.014	1.4
Poco	34	0.120	12.0
Abbastanza	185	0.654	65.4
Pienamente	60	0.212	21.2
<b>Totale</b>	<b>283</b>	<b>1.000</b>	<b>100.0</b>

La moda è  
"ABBASTANZA"

## Esempio: qual è la moda?

- ✘ Distribuzione degli studenti di SDC frequentanti la facoltà nell'a.a. 2001/2002 per Num. Corsi Frequentati

Num. Corsi Freq.	$n_i$
1	15
2	43
3	103
4	80
5	32
6	8
7	2
<b>Totale</b>	<b>283</b>

Num. Corsi Freq.	$n_i$
1	15
2	43
3	103
4	80
5	32
6	8
7	2
<b>Totale</b>	<b>283</b>

La moda è  
"3"

## Esempio: distribuzione multimodale

- × Distribuzione degli studenti di SDC frequentanti la facoltà nell'a.a. 2001/2002 per Tipo di Maturità

Tipo di Maturità	$n_i$	$f_i$	$p_i$
Classica	12	0,300	30,0
Scientifica	12	0,300	30,0
Tecnica	7	0,175	17,5
Altro	9	0,225	22,5
<b>Totale</b>	<b>40</b>	<b>1.000</b>	<b>100,0</b>

Tipo di Maturità	$n_i$	$f_i$	$p_i$
Classica	12	0,300	30,0
Scientifica	12	0,300	30,0
Tecnica	7	0,175	17,5
Altro	9	0,225	22,5
<b>Totale</b>	<b>40</b>	<b>1.000</b>	<b>100,0</b>

La distribuzione è **bimodale**  
Le mode sono "Classica" e "Scientifica"

## Esempio: classe modale

- × Distribuzione di laureati di SDC nell'a.a. 2003/2004 per Voto di Laurea

Voto Laurea	$n_i$
[87-98]	18
(98-102]	27
(102-105]	25
(105-109]	35
(109-111]	25
<b>Totale</b>	<b>130</b>

Qual è la classe modale?

Voto Laurea	$n_i$	$a_i$	$d_i$
[87-98]	18	11	1,64
(98-102]	27	4	6,75
(102-105]	25	3	8,33
(105-109]	35	4	8,75
(109-111]	25	2	12,50
<b>Totale</b>	<b>130</b>		

La classe modale è "(109-111]"

# Mediana

✗ Si può calcolare per *caratteri qualitativi ordinati* e per *caratteri quantitativi*

## Definizione

modalità che bipartisce la graduatoria (crescente o decrescente) delle osservazioni

U.S.	Alice	Marco	Elisa	Lucia	Fabio
Modalità	Basso	Basso	Medio	Medio	Alto

↑  
Mediana

## Calcolo della mediana per una variabile: 1

Collettivo di  $n=5$  unità

Variabile osservata  $X = \text{Altezza}$



165 cm



150 cm



180 cm



50 cm

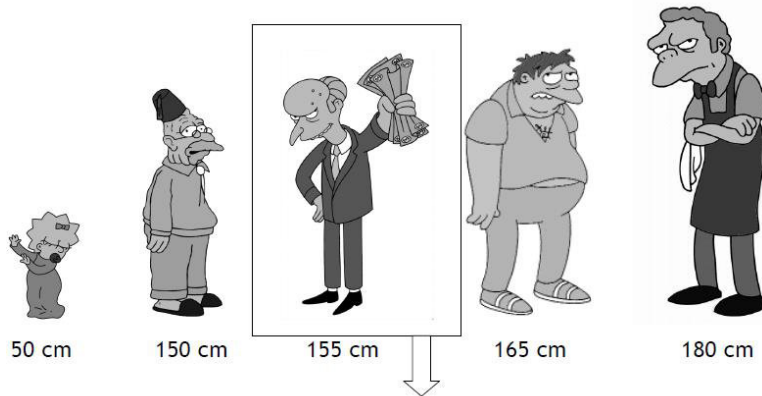


155 cm

1. Ordino le unità secondo un ordine crescente di *Altezza*

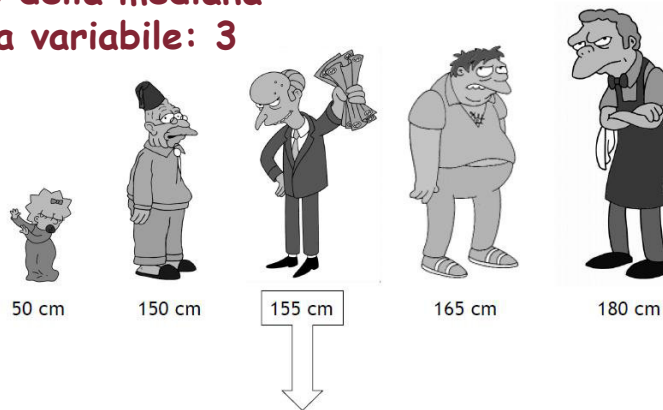
## Calcolo della mediana per una variabile: 2

Collettivo di  $n=5$  unità  
Variabile osservata  $X =$  Altezza



2. Identifico l'unità centrale nella serie ordinata dei dati

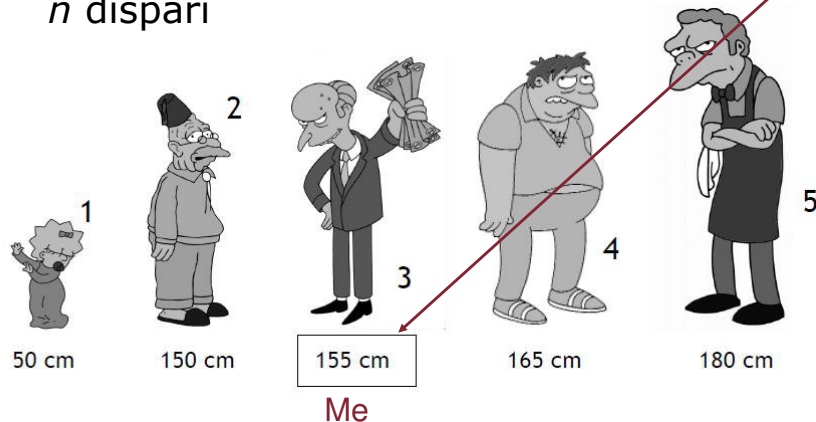
## Calcolo della mediana per una variabile: 3



**Me=155 cm**

3. La mediana è il valore che la variabile *Altezza* assume sull'unità che divide il collettivo in due parti numericamente uguali

Formalmente  $\longrightarrow$   $Me = x_{[(n+1)/2]} = x_{[(5+1)/2]} = x_3$   
 $n$  dispari



**N.B.** Le misure di posizione sono VALORI,  
NON FREQUENZE!!

Pagina 47

## Mediana: il calcolo

- 1) Ordinare in senso crescente (decescente) le u.s. rispetto alle modalità su di esse osservate del carattere in esame
- 2) Individuare l'unità che occupa il posto centrale

➤  $n$  dispari

★ Il posto centrale è  $\frac{n+1}{2}$

➤  $n$  pari

★ Ci sono due posti centrali :  $\frac{n}{2}$  e  $\frac{n}{2} + 1$

Pagina 48

## Mediana: il calcolo

3) Calcolare la Mediana: è la modalità presentata dall'unità individuata al punto 2)

➤  $n$  dispari

★ La mediana è la modalità dell'u.s. che occupa il posto  $\frac{n+1}{2}$  cioè  $Me = x_{\frac{n+1}{2}}$

➤  $n$  pari

★ La mediana è rappresentata dalla coppia di modalità delle u.s. che occupano i posti  $n/2$  e  $(n/2)+1$  cioè

$$Me = x_{(n/2)} \text{ e } Me = x_{(n/2)+1}$$

★ Se il carattere è quantitativo, la mediana è la semisomma delle 2 modalità individuate cioè

$$Me = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

## Esempio

✗ Distribuzione unitaria dei giudizi di 5 studenti

Senza ordine	
U.S.	Giudizio
U1	Buono
U2	Insuf.
U3	Discreto
U4	Suff.
U5	Ottimo

$$n = 5 \quad \frac{n+1}{2} = 3$$

Ordine crescente		
Posto	U.S.	Giudizio
1	U2	Insuf.
2	U4	Suff.
3	U3	Discreto
4	U1	Buono
5	U5	Ottimo

$$x_{\frac{n+1}{2}} = x_3 = \textit{Discreto}$$

La mediana è "Discreto"



## Esempio

✗ Distribuzione unitaria dei voti di 8 studenti

Senza ordine		
U.S.	Voto	
U1	22	
U2	30	
U3	28	
U4	18	
U5	27	
U6	20	
U7	25	
U8	28	

$$n = 8 \qquad \frac{n}{2} = 4 \qquad \frac{n}{2} + 1 = 5$$

$$x_{\frac{n}{2}} = x_4 = 25 \qquad x_{\frac{n}{2}+1} = x_5 = 27$$

Ordine crescente		
Posto	U.S.	Voto
1	U4	18
2	U6	20
3	U1	22
4	U7	25
5	U5	27
6	U3	28
7	U8	28
8	U2	30

Le mediane sono  
"25" e "27"

$$Me = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} = \frac{x_4 + x_5}{2} = \frac{25 + 27}{2} = 26$$

## Distribuzione di frequenza

Carattere $X$	Frequenza Assoluta $n_i$	Frequenza Relativa $f_i = n_i/n$	Frequenza Percentuale $p_j = f_i \times 100$
$x_1$	$n_1$	$f_1$	$p_1$
$x_2$	$n_2$	$f_2$	$p_2$
⋮	⋮	⋮	⋮
$x_i$	$n_i$	$f_i$	$p_i$
⋮	⋮	⋮	⋮
$x_K$	$n_K$	$f_K$	$p_K$
Totale	$n$	1	100

## Frequenze Cumulate

- ✗ Hanno senso solo se il carattere in esame è almeno un carattere qualitativo ordinale
- ✗ La frequenza cumulata associata alla modalità  $x_i$  del carattere rappresenta il numero di u.s che presentano una modalità non superiore a  $x_i$
- ✗ Frequenze Cumulate
  - si ottengono sommando le frequenze assolute (relative o percentuali) associate alle modalità inferiori o uguali alla modalità per la quale si sta calcolando la frequenza cumulata

## Osservazione

- ✗ Sommatoria  $\sum x$  : indica in modo sintetico la somma di un insieme di numeri  $x$ 
  - Sia  $x_1=3, x_2=5, x_3=1, x_4=9, x_5=15$

$$\sum_{i=1}^5 x_i = x_1 + x_2 + x_3 + x_4 + x_5 = 3 + 5 + 19 + 15 = 33$$

## Calcolo delle frequenze cumulate

- ✘ Frequenza Assoluta Cumulata della modalità  $x_i$

$$N_i = \sum_{h=1}^i n_h = n_1 + n_2 + \dots + n_i$$

- ✘ Frequenza Relativa Cumulata della modalità  $x_i$

$$F_i = \sum_{h=1}^i f_h = f_1 + f_2 + \dots + f_i$$

- ✘ Frequenza Percentuale Cumulata della modalità  $x_i$

$$P_i = \sum_{h=1}^i p_h = p_1 + p_2 + \dots + p_i$$

## Calcolo delle frequenze cumulate

- ✘ Ovviamente in corrispondenza della modalità più grande  $x_K$  si avrà
  - Frequenza assoluta cumulata  $N_K = n$
  - Frequenza relativa cumulata  $F_K = 1$
  - Frequenza percentuale cumulata  $P_K = 100$

## In sintesi

Carattere $X$	Frequenza Assoluta $n_i$	Frequenza Cumulata $N_i$
$x_1$	$n_1$	$N_1=n_1$
$x_2$	$n_2$	$N_2=n_1+n_2$
$\vdots$	$\vdots$	$\vdots$
$x_i$	$n_i$	$N_i=n_1+n_2+\dots+n_i$
$\vdots$	$\vdots$	$\vdots$
$x_K$	$n_K$	$N_K=n$
<b>Totale</b>	<b><math>n</math></b>	

Pagina 57

## Esempio: mediana per una distribuzione di frequenza

- × Distribuzione degli studenti di SDC frequentanti la facoltà nell'a.a. 2001/2002 secondo la Soddisfazione della Scelta universitaria

Soddisfazione	$n_i$	$N_i$	Posti in graduatoria
Per nulla	4	4	Da 1 a 4
Poco	24	38	Da 5 a 38
Abbastanza	185	223	Da 39 a 223
Pienamente	60	283	Da 224 a 283
<b>Totale</b>	<b>283</b>		

La mediana è  
"Abbastanza"

$$n = 283 \quad \frac{n+1}{2} = 142 \quad x_{\frac{n+1}{2}} = x_{142} = \text{Abbastanza}$$

Pagina 58

## Proprietà della Mediana

- 1) E' sempre compresa tra la modalità minima  $x_1$  e la modalità massima  $x_k$  del carattere
- 2) E' robusta cioè è poco sensibile ai cambiamenti che possono avvenire sulle modalità estreme della distribuzione del carattere