Esercitazione 2

Statistica

Alfonso Iodice D'Enza
iodicede@unina.it

Università degli studi di Cassino

Outline

1. la media aritmetica

Outline

1. la media aritmetica
2. Proprietà della media aritmetica

Outline

Outline

la media aritmetica

Si consideri la funzione $f(.)$ additiva, vale a dire

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} x_i$$

Ricordando il principio di rappresentatitività (equazione **??**) si ha che

$$f(x_1, x_2, \ldots, x_n) = f(\mu, \mu, \ldots, \mu)$$

Poichè $f(.)$ è di tipo additivo la precedente uguaglianza si può riscrivere come

$$f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \mu$$

la media aritmetica

quindi

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \mu \Longleftrightarrow \sum_{i=1}^{n} x_i = n\mu \Longleftrightarrow \mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

la media aritmetica

- media semplice:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- media per dati organizzati in frequenze:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i n_i$$

- media per frequenze relative:

$$\mu = \sum_{i=1}^{n} x_i \frac{n_i}{n}$$

Calcolo della media aritmetica semplice: un esempio

La distribuzione unitaria della media voto di un collettivo di $n = 20$ studenti

Calcolo della media aritmetica semplice: un esempio

La distribuzione unitaria della media voto di un collettivo di $n = 20$ studenti

Unità statistiche	Voto medio
1	22.5
2	23
3	18.5
4	18.3
5	28
6	25.7
7	24.2
8	28.7
9	27.9
10	27
11	24.6
12	26.8
13	21.5
14	20.3
15	23.6
16	26.4
17	18.9
18	19.4
19	19.3
20	26

Calcolo della media aritmetica semplice: un esempio

La distribuzione unitaria della media voto di un collettivo di $n = 20$ studenti

Unità statistiche	Voto medio
1	22.5
2	23
3	18.5
4	18.3
5	28
6	25.7
7	24.2
8	28.7
9	27.9
10	27
11	24.6
12	26.8
13	21.5
14	20.3
15	23.6
16	26.4
17	18.9
18	19.4
19	19.3
20	26

$$\mu = \frac{1}{20} \times (22.5 + 23 + 18.5 + 18.3 + \\ + 28 + 25.7 + 24.2 + 28.7 + \\ + 27.9 + 27 + 24.6 + 26.8 + \\ + 21.5 + 20.3 + 23.6 + 26.4 + \\ + 18.9 + 19.4 + 19.3 + 26) = \\ = \frac{470.6}{20} = 23.53$$

Calcolo della media aritmetica per dati in frequenze assolute: un esempio

Considerando la distribuzione di frequenze rispetto $K = 4$ classi (intervalli) di voto: in questo caso si considerano i valori centrali delle classi.

Calcolo della media aritmetica per dati in frequenze assolute: un esempio

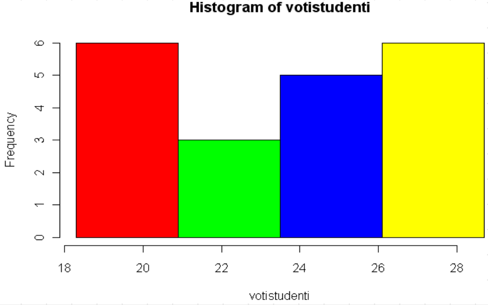
Considerando la distribuzione di frequenze rispetto $K = 4$ classi (intervalli) di voto: in questo caso si considerano i valori centrali delle classi.

Classi voto K=4	frequenza Assoluta
[18.3-20.9(	6
[20.9-23.5(	3
[23.5-26.1(	5
[26.1- 28.7]	6
Totale	*20*

Calcolo della media aritmetica per dati in frequenze assolute: un esempio

Considerando la distribuzione di frequenze rispetto $K = 4$ classi (intervalli) di voto: in questo caso si considerano i valori centrali delle classi.

Classi voto K=4	frequenza Assoluta
[18.3-20.9(	6
[20.9-23.5(	3
[23.5-26.1(	5
[26.1- 28.7]	6
Totale	*20*



Calcolo della media aritmetica per dati in frequenze assolute: un esempio

$$\mu = \frac{1}{20} \times (c_1 \times n_1 + c_2 \times n_2 +$$
$$+ c_3 \times n_3 + c_4 \times n_4) =$$
$$= \frac{1}{20} \times (19.6 \times 6 + 22.2 \times 3 +$$
$$+ (24.8 \times 5 + 27.4 \times 6) =$$
$$= \frac{472.6}{20} = 23.63$$

Calcolo della media aritmetica per dati in frequenze relative: un esempio

Considerando la distribuzione di frequenze relative rispetto $K = 4$ classi (intervalli) di voto rispetto ai valori centrali delle classi.

Calcolo della media aritmetica per dati in frequenze relative: un esempio

Considerando la distribuzione di frequenze relative rispetto $K = 4$ classi (intervalli) di voto rispetto ai valori centrali delle classi.

Classi voto K=4	**frequenza frelativa**
[18.3-20.9(	*0.3*
[20.9-23.5(	*0.15*
[23.5-26.1(	0.25
[26.1- 28.7]	0.3
Totale	***1***

Calcolo della media aritmetica per dati in frequenze relative: un esempio

Considerando la distribuzione di frequenze relative rispetto $K = 4$ classi (intervalli) di voto rispetto ai valori centrali delle classi.

Classi voto K=4	frequenza frelativa
[18.3-20.9(	0.3
[20.9-23.5(	0.15
[23.5-26.1(	0.25
[26.1- 28.7]	0.3
Totale	1

$$\mu = (c_1 \times f_1 + c_2 \times f_2 + \\ + c_3 \times f_3 + c_4 \times f_4) = \\ = (19.6 \times 0.3 + 22.2 \times 0.15 + \\ + 24.8 \times 0.25 + 27.4 \times 0.3) = \\ = (5.88 + 3.33 + 6.2 + 8.22) = 23.63$$

Proprietà della media aritmetica

1 criterio di internalità la media aritmetica è sempre compresa tra il minimo e massimo della distribuzione osservata:

$$\sum x_1 \leq \sum_{i=1}^{n} x_i \leq \sum x_n \Leftrightarrow nx_1 \leq \sum_{i=1}^{n} x_i \leq nx_n \Leftrightarrow$$

$$\Leftrightarrow x_1 \leq \frac{\sum_{i=1}^{n} x_i}{n} \leq x_n \Leftrightarrow x_1 \leq \mu \leq x_n$$

Proprietà della media aritmetica

2 media come baricentro la somma degli scarti dalla media è nulla:

$$\sum_{i=1}^{n}(x_i-\mu)=\sum_{i=1}^{n}x_i-n\mu=$$

$$=\sum_{i=1}^{n}x_i-n\left(\frac{\sum_{i=1}^{n}x_i}{n}\right)=\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}x_i=0$$

Proprietà della media aritmetica

Esercitazione 2

A. Iodice

la media aritmetica

Proprietà della media aritmetica

La media ponderata

Indici robusti

3 linearità della media aritmetica Sia X è una variabile con media μ, allora la variabile $Y = \alpha + \beta X$ avrà media

$$M(Y) = \frac{1}{n}\sum_{i=1}^{n}(\alpha + \beta x_i) = \frac{1}{n}\sum_{i=1}^{n}(\alpha) + \frac{1}{n}\sum_{i=1}^{n}(\beta x_i) =$$

$$= \frac{1}{n}(n\alpha) + \beta(\frac{1}{n}\sum_{i=1}^{n}(x_i)) =$$

$$= \alpha + \beta\mu$$

Proprietà della media aritmetica

Si consideri che $\alpha = -18/12$ e $\beta = \frac{1}{12}$

utilizzando α e β per normalizzare i voti degli studenti

$$Y = \alpha + X \times \beta = -\frac{18}{12} + X \times \frac{1}{12}$$

	[,1]	[,2]
[1,]	22.5	0.38
[2,]	23.0	0.42
[3,]	18.5	0.04
[4,]	18.3	0.03
[5,]	28.0	0.83
[6,]	25.7	0.64
[7,]	24.2	0.52
[8,]	28.7	0.89
[9,]	27.9	0.82
[10,]	27.0	0.75
[11,]	24.6	0.55
[12,]	26.8	0.73
[13,]	21.5	0.29
[14,]	20.3	0.19
[15,]	23.6	0.47
[16,]	26.4	0.70
[17,]	18.9	0.07
[18,]	19.4	0.12
[19,]	19.3	0.11
[20,]	26.0	0.67

X=Valori in trentesimi

Y=Valori normalizzati

Media X= 23.53
Media Y= 0.461

Media Y=a+b*(Media X)=
=-(18/12)+(1/12)*23.53= 0.461

4 proprietà associativa della media aritmetica

Sia X una variabile osservata su più gruppi: la media può essere ottenuta come media delle medie calcolate in ciascun gruppo - tenendo conto della differente numerosità dei gruppi . Il collettivo è suddiviso in $K gruppi$ di numerosita $n_1, n_2, \ldots, n_K$. La media del carattere X sul collettivo è μ.

Per la proprietà associativa si avrà che

$$\mu = \mu_1 \times \frac{n_1}{n} + \mu_2 \times \frac{n_2}{n} + \ldots + \mu_K \times \frac{n_K}{n}$$

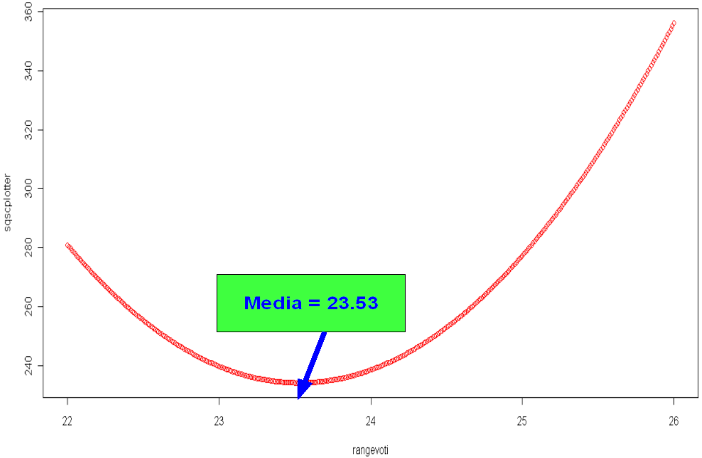
Proprietà della media aritmetica

5 minimizzazione dei quadrati degli scarti La media aritmetica μ rende minima la somma dei quadrati degli scarti X:

$$\sum_{i=1}^{n}(x_i-\mu)^2 = min$$

La media mininizza la somma dei quadrati degli scarti

Tornando ai voti degli studenti...



La media aritmetica ponderata

Nel calcolo della media aritmetica tutte le modalità e le unità statistiche hanno la stessa importanza: ciascuna modalità ha un *peso* pari a $\frac{1}{n}$ nel determinare il valore μ.

La media aritmetica ponderata

Nel calcolo della media aritmetica tutte le modalità e le unità statistiche hanno la stessa importanza: ciascuna modalità ha un *peso* pari a $\frac{1}{n}$ nel determinare il valore μ.

media aritmetica ponderata

Le modalità di un carattere possono tuttavia avere intrinsecamente una diversa importanza: in questi casi un indice appropriato la media aritmetica ponderata.

La media aritmetica ponderata

Nel calcolo della media aritmetica tutte le modalità e le unità statistiche hanno la stessa importanza: ciascuna modalità ha un *peso* pari a $\frac{1}{n}$ nel determinare il valore μ.

media aritmetica ponderata

Le modalità di un carattere possono tuttavia avere intrinsecamente una diversa importanza: in questi casi un indice appropriato la media aritmetica ponderata.
Siano ω_i i pesi di ciascuna modalità x_i, la media ponderata μ_ω sar data da

$$\mu_\omega = \frac{\sum_{i=1}^{K} x_i \omega_i}{\sum_{i=1}^{K} \omega_i}$$

La media aritmetica ponderata

Nel calcolo della media aritmetica tutte le modalità e le unità statistiche hanno la stessa importanza: ciascuna modalità ha un *peso* pari a $\frac{1}{n}$ nel determinare il valore μ.

media aritmetica ponderata

Le modalità di un carattere possono tuttavia avere intrinsecamente una diversa importanza: in questi casi un indice appropriato la media aritmetica ponderata.
Siano ω_i i pesi di ciascuna modalità x_i, la media ponderata μ_ω sar data da

$$\mu_\omega = \frac{\sum_{i=1}^{K} x_i \omega_i}{\sum_{i=1}^{K} \omega_i}$$

Calcolo della media aritmetica ponderata: un esempio

La distribuzione unitaria dei voti conseguiti da uno studente universitario: agli esami associato un numero di crediti proporzionali all'impegno richiesto.

Calcolo della media aritmetica ponderata: un esempio

La distribuzione unitaria dei voti conseguiti da uno studente universitario: agli esami associato un numero di crediti proporzionali all'impegno richiesto.

voto esame studenti A e B	credito studente A	credito studente B
26	4	4
24	6	12
24	4	4
22	5	5
28	12	6
30	8	5
30	9	5
25	9	7
27	7	9
28	5	8
24	10	10
27	5	9
28	10	10
23	6	10
30	10	6
30	4	4
23	4	12
30	12	4

Calcolo della media aritmetica ponderata: un esempio

La distribuzione unitaria dei voti conseguiti da uno studente universitario: agli esami associato un numero di crediti proporzionali all'impegno richiesto.

voto esame studenti A e B	credito studente A	credito studente B
26	4	4
24	6	12
24	4	4
22	5	5
28	12	6
30	8	5
30	9	5
25	9	7
27	7	9
28	5	8
24	10	10
27	5	9
28	10	10
23	6	10
30	10	6
30	4	4
23	4	12
30	12	4

media aritmetica semplice

$$\mu = \frac{26+24+24+\ldots+23+30}{18} =$$

$$= \frac{479}{18} = 26.61$$

Calcolo della media aritmetica ponderata: un esempio

La distribuzione unitaria dei voti conseguiti da uno studente universitario: agli esami associato un numero di crediti proporzionali all'impegno richiesto.

voto esame studenti A e B	credito studente A	credito studente B
26	4	4
24	6	12
24	4	4
22	5	5
28	12	6
30	8	5
30	9	5
25	9	7
27	7	9
28	5	8
24	10	10
27	5	9
28	10	10
23	6	10
30	10	6
30	4	4
23	4	12
30	12	4

media aritmetica ponderata studente A

$$\mu_\omega = \frac{(26\times 4)+(24\times 6)+(24\times 4)+\dots}{4+6+4+\dots}$$

$$\frac{\dots+(23\times 4)+(30\times 12)}{\dots+4+12} = \frac{3519}{130} = 27.07$$

Calcolo della media aritmetica ponderata: un esempio

La distribuzione unitaria dei voti conseguiti da uno studente universitario: agli esami associato un numero di crediti proporzionali all'impegno richiesto.

voto esame studenti A e B	credito studente A	credito studente B
26	4	4
24	6	12
24	4	4
22	5	5
28	12	6
30	8	5
30	9	5
25	9	7
27	7	9
28	5	8
24	10	10
27	5	9
28	10	10
23	6	10
30	10	6
30	4	4
23	4	12
30	12	4

media aritmetica ponderata studente B

$$\mu_\omega = \frac{(26\times 4)+(24\times 12)+(24\times 4)+\dots}{4+12+4+\dots}$$

$$\frac{\dots+(23\times 12)+(30\times 4)}{\dots+12+4} = \frac{3397}{130} = 26.14$$

Indici robusti (o di posizione): la moda

La moda di una distribuzione di frequenze si indica con Mo e corrisponde alla modalità con la frequenza assoluta (relativa) più alta

Indici robusti (o di posizione): la moda

La moda di una distribuzione di frequenze si indica con Mo e corrisponde alla modalità con la frequenza assoluta (relativa) più alta

Esempio

Dato un collettivo di 10 unità statistiche, si consideri la seguente serie di osservazioni:

$$\{1, 2, 3, 4, 4, 4, 4, 3, 4, 1\}$$

risulta $Mo = 4$, dal momento che la modalità 4 presente cinque volte nel collettivo.

Indici robusti (o di posizione): la moda

classe modale

E' possibile calcolare la moda anche per variabili quantitative continue raggruppate in classi: se le classi sono di ampiezza diversa si fa riferimento alla densità di frequenza e non alla frequenza di ciascuna classe.

Indici robusti (o di posizione): la mediana

La mediana Me di una distribuzione corrisponde alla modalità osservata sulla unità statistica centrale nella distribuzione ordinata delle osservazioni

Indici robusti (o di posizione): la mediana

La mediana Me di una distribuzione corrisponde alla modalità osservata sulla unità statistica centrale nella distribuzione ordinata delle osservazioni

variabile X discreta

Si considerino le osservazioni ordinate dal valore minimo x_1 al valore massimo x_n, la mediana Me sar data da

$$Me = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & \text{se } n \text{ dispari} \\ \frac{x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}}{2} & \text{se } n \text{ pari} \end{cases}$$

La mediana per variabili discrete (dati in distribuzione unitaria)

Esempio (n dispari)

Dato un collettivo di 15 unità statistiche, si consideri la seguente serie di osservazioni:

$$\{29, 7, 18, 15, 27, 23, 14, 1, 25, 13, 18, 24, 28, 22, 5\}$$

che ordinata in modo crescente diventa

$$\{1, 5, 7, 13, 14, 15, 18, 18, 22, 23, 24, 25, 27, 28, 29\}$$

la mediana è data dalla modalità che occupa la posizione $\frac{n+1}{2} = 16/2 = 8$, vale a dire $Me = 18$

La mediana per variabili discrete (dati in distribuzione unitaria)

Esempio (n pari)

Dato un collettivo di 12 unità statistiche, si consideri la seguente serie di osservazioni:

$$\{34, 42, 1, 34, 19, 42, 25, 35, 21, 15, 9, 10\}$$

che ordinata in modo crescente diventa

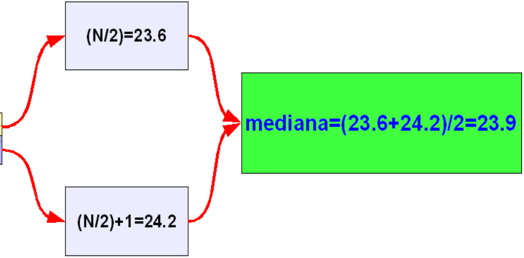
$$\{1, 9, 10, 15, 19, 21, 25, 34, 34, 35, 42, 42\}$$

la mediana data dalla semi somma delle modalità che occupano le posizioni $\frac{n}{2} = 12/2 = 6$ e $\frac{n}{2} + 1 = (12/2) + 1 = 7$, vale a dire $Me = \frac{21+25}{2} = \frac{46}{2} = 23$

La mediana per variabili discrete (dati in distribuzione unitaria)

Tornando ai voti degli studenti...

```
       [,1]
 [1,]  18.3
 [2,]  18.5
 [3,]  18.9
 [4,]  19.3
 [5,]  19.4
 [6,]  20.3
 [7,]  21.5
 [8,]  22.5
 [9,]  23.0
[10,]  23.6
[11,]  24.2
[12,]  24.6
[13,]  25.7
[14,]  26.0
[15,]  26.4
[16,]  26.8
[17,]  27.0
[18,]  27.9
[19,]  28.0
[20,]  28.7
```



La mediana per variabili discrete (dati organizzati in frequenze)

Se i dati della variabile X discreta sono noti mediante una distribuzione di frequenze (seriazione) l'individuazione della modalità assunta dall'unità statistica di posto centrale risulta agevole se, calcolate le frequenze relative cumulate, si considera la funzione di ripartizione $F(x)$ (supponendo di aver ordinato le modalità di X in modo crescente) .

La mediana per variabili discrete (dati organizzati in frequenze)

Se i dati della variabile X discreta sono noti mediante una distribuzione di frequenze (seriazione) l'individuazione della modalità assunta dall'unità statistica di posto centrale risulta agevole se, calcolate le frequenze relative cumulate, si considera la funzione di ripartizione $F(x)$ (supponendo di aver ordinato le modalità di X in modo crescente) .

individuare la modalità $x_{(i-1)}$ tale che

- $F(x_{(i-1)}) < 0.5$

La mediana per variabili discrete (dati organizzati in frequenze)

Se i dati della variabile X discreta sono noti mediante una distribuzione di frequenze (seriazione) l'individuazione della modalità assunta dall'unità statistica di posto centrale risulta agevole se, calcolate le frequenze relative cumulate, si considera la funzione di ripartizione $F(x)$ (supponendo di aver ordinato le modalità di X in modo crescente) .

individuare la modalità $x_{(i-1)}$ tale che

- $F(x_{(i-1)}) < 0.5$
- $F(x_{(i)}) \geq 0.5$

La mediana per variabili discrete (dati organizzati in frequenze)

Se i dati della variabile X discreta sono noti mediante una distribuzione di frequenze (seriazione) l'individuazione della modalità assunta dall'unità statistica di posto centrale risulta agevole se, calcolate le frequenze relative cumulate, si considera la funzione di ripartizione $F(x)$ (supponendo di aver ordinato le modalità di X in modo crescente) .

individuare la modalità $x_{(i-1)}$ tale che

- $F(x_{(i-1)}) < 0.5$
- $F(x_{(i)}) \geq 0.5$

La mediana data da $Me = x_i$: tra le n_i unità statistiche che presentano la modalità x_i ci saranno quelle di posizione centrale (o delle posizioni centrali se n dispari)

La mediana per variabili discrete (dati organizzati in frequenze)

Modalita di X	Frequenze assolute	Frequenze relative	F(x)
5	135	0.1482	0.1482
7	123	0.1350	0.2832
10	97	0.1065	0.3897
11	123	0.1350	0.5247
21	99	0.1087	0.6334
29	51	0.0560	0.6894
35	43	0.0472	0.7366
38	51	0.0560	0.7925
39	80	0.0878	0.8804
40	109	0.1196	1.0000

La mediana per variabili discrete (dati organizzati in frequenze)

Modalita di X	Frequenze assolute	Frequenze relative	F(x)
5	135	0.1482	0.1482
7	123	0.1350	0.2832
10	97	0.1065	0.3897
11	123	0.1350	0.5247
21	99	0.1087	0.6334
29	51	0.0560	0.6894
35	43	0.0472	0.7366
38	51	0.0560	0.7925
39	80	0.0878	0.8804
40	109	0.1196	1.0000

$F(x_{i-1})$

La mediana per variabili discrete (dati organizzati in frequenze)

Modalita di X	Frequenze assolute	Frequenze relative	F(x)
5	135	0.1482	0.1482
7	123	0.1350	0.2832
10	97	0.1065	0.3897
11	123	0.1350	0.5247
21	99	0.1087	0.6334
29	51	0.0560	0.6894
35	43	0.0472	0.7366
38	51	0.0560	0.7925
39	80	0.0878	0.8804
40	109	0.1196	1.0000

$F(x_{i-1})$

$F(x_i)$

La mediana per variabili discrete (dati organizzati in frequenze)

X_i=Me

Modalita di X	Frequenze assolute	Frequenze relative	F(x)
5	135	0.1482	0.1482
7	123	0.1350	0.2832
10	97	0.1065	0.3897
11	123	0.1350	0.5247
21	99	0.1087	0.6334
29	51	0.0560	0.6894
35	43	0.0472	0.7366
38	51	0.0560	0.7925
39	80	0.0878	0.8804
40	109	0.1196	1.0000

$F(x_{i-1})$

$F(x_i)$

La mediana per variabili continue (suddivise in classi)

Nel caso di variabili continue, la suddivisione in classi delle modalità del carattere consente l'identificazione della classe mediana: in tale classe cadrà la modalità assunta dall'unità statistica che occupa la posizione centrale della distribuzione ordinata delle modalità

La mediana per variabili continue (suddivise in classi)

Nel caso di variabili continue, la suddivisione in classi delle modalità del carattere consente l'identificazione della classe mediana: in tale classe cadrà la modalità assunta dall'unità statistica che occupa la posizione centrale della distribuzione ordinata delle modalità

approssimazione del valore Me (tramite $F(x)$ F. di ripartizione)

$$Me \simeq x_{i-1} + (x_i - x_{i-1}) \frac{0.5 - F(x_{i-1})}{F(x_i) - F(x_{i-1})}$$

- x_i la modalità per la quale $F(x_i) \geq 0.5$
- x_{i-1} la modalità per la quale $F(x_{i-1}) < 0.5$

La mediana per variabili continue (suddivise in classi)

Modalita di X Suddivisa in classi	Frequenze assolute	Frequenze relative	F(x)
[20,40]	135	0.1482	0.1482
)40,60]	123	0.1350	0.2832
)60,80]	97	0.1065	0.3897
)80,100]	123	0.1350	0.5247
)100,120]	99	0.1087	0.6334
)140,160]	51	0.0560	0.6894
)160,180]	43	0.0472	0.7366
)180,200]	51	0.0560	0.7925
)220,240]	80	0.0878	0.8804
)240,260]	109	0.1196	1.0000

La mediana per variabili continue (suddivise in classi)

Modalita di X Suddivisa in classi	*Frequenze assolute*	*Frequenze relative*	*F(x)*
[20,40]	135	0.1482	0.1482
)40,60]	123	0.1350	0.2832
)60,80]	97	0.1065	0.3897
)80,100]	123	0.1350	0.5247
)100,120]	99	0.1087	0.6334
)140,160]	51	0.0560	0.6894
)160,180]	43	0.0472	0.7366
)180,200]	51	0.0560	0.7925
)220,240]	80	0.0878	0.8804
)240,260]	109	0.1196	1.0000

$F(x_{i-1})$

$F(x_i)$

La mediana per variabili continue (suddivise in classi)

Modalita di X Suddivisa in classi	Frequenze assolute	Frequenze relative	F(x)
[20,40]	135	0.1482	0.1482
)40,60]	123	0.1350	0.2832
)60,80]	97	0.1065	0.3897
)80,100]	123	0.1350	0.5247
)100,120]	99	0.1087	0.6334
)140,160]	51	0.0560	0.6894
)160,180]	43	0.0472	0.7366
)180,200]	51	0.0560	0.7925
)220,240]	80	0.0878	0.8804
)240,260]	109	0.1196	1.0000

x_{i-1}

$F(x_{i-1})$

$F(x_i)$

La mediana per variabili continue (suddivise in classi)

Modalita di X Suddivisa in classi	*Frequenze assolute*	*Frequenze relative*	*F(x)*
[20,40]	135	0.1482	0.1482
[illegible]60]	123	0.1350	0.2832
]60,80]	97	0.1065	0.3897
)80,100]	123	0.1350	0.5247
[illegible]120]	99	0.1087	0.6334
[illegible]40,160]	51	0.0560	0.6894
)160,180]	43	0.0472	0.7366
)180,200]	51	0.0560	0.7925
)220,240]	80	0.0878	0.8804
)240,260]	109	0.1196	1.0000

x_{i-1}

x_i

$F(x_{i-1})$

$F(x_i)$

La mediana per variabili continue (suddivise in classi)

Trovati i valori x_i, x_{i-1}, $F(x_i)$ e $F(x_{i-1})$ possibile ricavare il valore approssimato di Me

La mediana per variabili continue (suddivise in classi)

Trovati i valori x_i, x_{i-1}, $F(x_i)$ e $F(x_{i-1})$ possibile ricavare il valore approssimato di Me

$$
\begin{aligned}
Me \simeq & x_{i-1} + (x_i - x_{i-1}) \frac{0.5 - F(x_{i-1})}{F(x_i) - F(x_{i-1})} = \\
& = 80 + (100 - 80) \frac{0.5 - 0.3897}{0.5247 - 0.3897} = \\
& = 80 + 20 \times 0.817 = 96.3407
\end{aligned}
$$

Propriet della mediana

- la mediana sempre un valore osservato su una delle unità statistiche oggetto di studio

Propriet della mediana

- la mediana sempre un valore osservato su una delle unità statistiche oggetto di studio
- la mediana minimizza la somma degli scarti in valore assoluto. Dato un valore c,

$$\sum_{i=1}^{n} |x_i - c|$$

minima se (*e solo se*) $c \equiv Me$.

Propriet della mediana

- la mediana sempre un valore osservato su una delle unità statistiche oggetto di studio
- la mediana minimizza la somma degli scarti in valore assoluto. Dato un valore c,

$$\sum_{i=1}^{n} |x_i - c|$$

  minima se (e *solo se*) $c \equiv Me$.
- resistenza della mediana ai valori estremi

Quantili di una distribuzione

La mediana corrisponde alla modalità assunta dall'unità statistica che bipartisce la distribuzione ordinata delle osservazioni Il 50% delle unità statistiche si trova alla sinistra della mediana, l'altro 50% alla sua destra. Analogamente possibile definire i seguenti quantili di una distribuzione.

Quantili di una distribuzione

La mediana corrisponde alla modalità assunta dall'unità statistica che bipartisce la distribuzione ordinata delle osservazioni Il 50% delle unità statistiche si trova alla sinistra della mediana, l'altro 50% alla sua destra. Analogamente possibile definire i seguenti quantili di una distribuzione.

- primo quartile Q_1: corrisponde alla modalità assunta dall'unità statistica per il quale $\frac{1}{4}n$ (25%) delle unità statistiche presenta valori ad esso inferiori
- secondo quartile Q_2: coincide con la mediana
- terzo quartile Q_3: corrisponde alla modalità assunta dall'unità statistica per il quale $\frac{3}{4}n$ (75%) delle unità statistiche presenta valori ad esso inferiori
- decili

Quartili in distribuzione unitaria)

Esempio (n dispari)

Dato un collettivo di 15 unità statistiche, si consideri la seguente serie di osservazioni:

$$\{29, 7, 18, 15, 27, 23, 14, 1, 25, 13, 18, 24, 28, 22, 5\}$$

che ordinata in modo crescente diventa

$$\{1, 5, 7, 13, 14, 15, 18, 18, 22, 23, 24, 25, 27, 28, 29\}$$

- la posizione di Q_1 $(n+1) * \frac{1}{4} = 16/4 = 4$, $Q_1 = 13$
- la posizione di Q_2 $(n+1) * \frac{2}{4} = 16 * 2/4 = 8$, $Q_2 = 18$
- la posizione di Q_3 $(n+1) * \frac{1}{4} = 16 * 3/4 = 12$, $Q_3 = 25$