

ESERCIZIO 1.1

La seguente tabella riporta i dati concernenti 4 caratteri rilevati su 30 abitazioni situate nella città di Roma. I caratteri rilevati riguardano il quartiere, il valore stimato del terreno (in migliaia di Euro), il valore stimato delle ristrutturazioni (in migliaia di Euro) ed il prezzo di vendita (in migliaia di Euro).

ID	Quartiere	Valore stimato del terreno	Valore stimato delle ristrutturazioni	Prezzo di vendita
1	B	44,01	13,62	175,00
2	B	30,28	24,34	187,50
3	B	31,68	97,94	173,50
4	B	15,06	58,84	81,00
5	B	47,73	81,46	125,50
6	B	66,24	25,66	243,00
7	C	15,44	70,31	105,00
8	C	12,22	52,90	87,50
9	C	12,04	34,65	60,00
10	C	30,41	73,21	140,00
11	C	26,52	38,72	84,50
12	D	16,10	66,43	85,00
13	D	55,93	40,92	243,80
14	D	14,59	52,79	82,00
15	D	16,20	78,68	86,80
16	E	59,80	86,12	205,00
17	E	20,80	19,64	78,00
18	E	26,60	6,39	58,00
19	E	10,70	57,66	74,20
20	F	16,93	16,94	39,00
21	F	25,52	37,81	59,90
22	F	14,49	14,95	49,50
23	G	17,95	31,95	61,50
24	G	10,13	10,84	25,40
25	G	5,27	41,27	37,70
26	G	7,55	18,73	36,00
27	G	7,74	18,47	37,20
28	G	7,12	20,13	54,30
29	G	8,37	21,35	45,00
30	G	7,15	16,19	22,40

1) Definire, per ciascuno dei caratteri riportati in tabella, il tipo elencando l'insieme delle possibili modalità/intensità.

2) Costruire la distribuzione di frequenza per il carattere QUARTIERE riportando le frequenze assolute e le frequenze relative. Rappresentare graficamente la distribuzione ottenuta.

3) E' sempre possibile, partendo da una qualunque delle due colonne contenenti le frequenze, risalire all'altra? (rispondere brevemente motivando la risposta).

4) Considerando separatamente le abitazioni situate nei quartieri "B", "C" e "D" e quelle situate nei quartieri "E", "F", "G", costruire le due distribuzioni di frequenza per il carattere PREZZO DI VENDITA (prezzo di vendita nei quartieri "B", "C" e "D" e prezzo di vendita nei quartieri "E", "F" e "G"), in modo da favorire il confronto tra i due gruppi.

4.1) Suddividere ciascuna distribuzione in 3 classi equiampie ed effettuare il confronto tra distribuzioni riportando le frequenze (assolute e relative) e la funzione di ripartizione empirica.

Rappresentare graficamente le due distribuzioni attraverso l'istogramma normalizzato.

Commentare brevemente i risultati ottenuti.

4.2) Suddividere ciascuna distribuzione in 3 classi equifrequenti ed effettuare il confronto tra distribuzioni riportando le frequenze (assolute e relative) e la funzione di ripartizione empirica.

Rappresentare graficamente le due distribuzioni attraverso l'istogramma normalizzato.

Commentare brevemente i risultati ottenuti.

5) Calcolare media aritmetica, moda e mediana per il carattere VALORE DEL TERRENO a partire dalla:

5.1) distribuzione semplice del carattere;

5.2) distribuzione ottenuta suddividendo le osservazioni in 6 classi equiampie;

5.3) distribuzione ottenuta suddividendo le osservazioni in 6 classi equifrequenti.

Soluzione

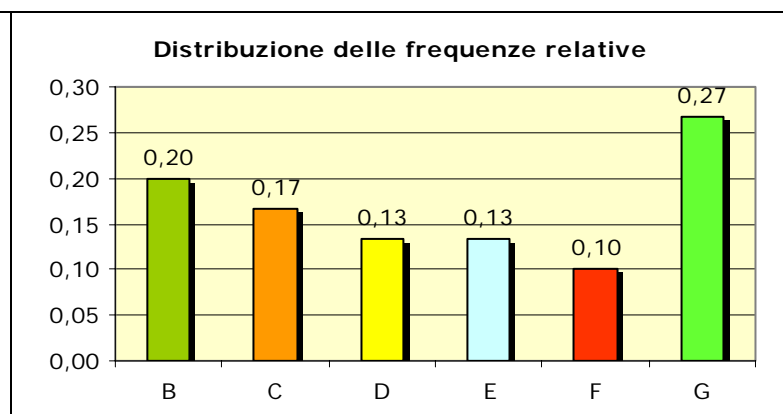
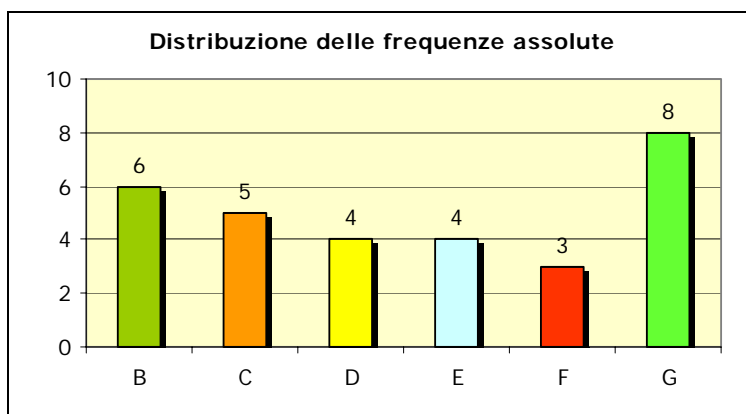
1)

La descrizione di ciascuno dei 4 caratteri per tipo e modalità assunte può essere sintetizzata come segue:

Carattere	Tipo	Modalità
Quartiere	qualitativo	{B, C, D, E, F, G}
Valore stimato del terreno	quantitativo	[5,27; 66,24]
Valore stimato delle ristrutturazioni	quantitativo	[6,39; 97,94]
Prezzo di vendita	quantitativo	[22,40; 243,80]

2)

QUARTIERE	Frequenze assolute n_i	Frequenze relative f_i
B	6	0,20
C	5	0,17
D	4	0,13
E	4	0,13
F	3	0,10
G	8	0,27
Totale complessivo	30	1



3)

Dalle frequenze assolute n_i è sempre possibile derivare le frequenze

relative f_i perché $f_i = \frac{n_i}{N}$, dove $N = \sum_{i=1}^n n_i$.

È possibile, invece, risalire dalle frequenze relative f_i a quelle assolute n_i solo se si conosce il valore di N , perché $n_i = N \times f_i$.

4)

Le due distribuzioni semplici del carattere PREZZO separate per i due gruppi di quartieri B-C-D ed E-F-G sono le seguenti:

Quartiere	Prezzo di vendita
B	81,00
B	125,50
B	173,50
B	175,00
B	187,50
B	243,00
C	60,00
C	84,50
C	87,50
C	105,00
C	140,00
D	82,00
D	85,00
D	86,80
D	243,80

Quartiere	Prezzo di vendita
E	58,00
E	74,20
E	78,00
E	205,00
F	39,00
F	49,50
F	59,90
G	22,40
G	25,40
G	36,00
G	37,20
G	37,70
G	45,00
G	54,30
G	61,50

Al fine di effettuare un confronto tra le due distribuzioni, è necessario innanzitutto ordinare i prezzi in ordine crescente come segue:

Prezzo B-C-D
60,00
81,00
82,00
84,50
85,00
86,80
87,50
105,00
125,50
140,00
173,50
175,00
187,50
243,00
243,80

Prezzo E-F-G
22,40
25,40
36,00
37,20
37,70
39,00
45,00
49,50
54,30
58,00
59,90
61,50
74,20
78,00
205,00

4.1)

Per ottenere l'ampiezza costante delle tre classi basta dividere il range (= valore massimo – valore minimo) per 3:

$$\text{range (PREZZO BCD)} = 243,80 - 60,00 = 183,80$$

L'ampiezza della prima distribuzione di frequenza è pari a:

$$a_{BCD} = \frac{\text{range}}{3} = \frac{183,80}{3} = 61,27$$

Le 3 classi sono, dunque, delimitate dai seguenti estremi:

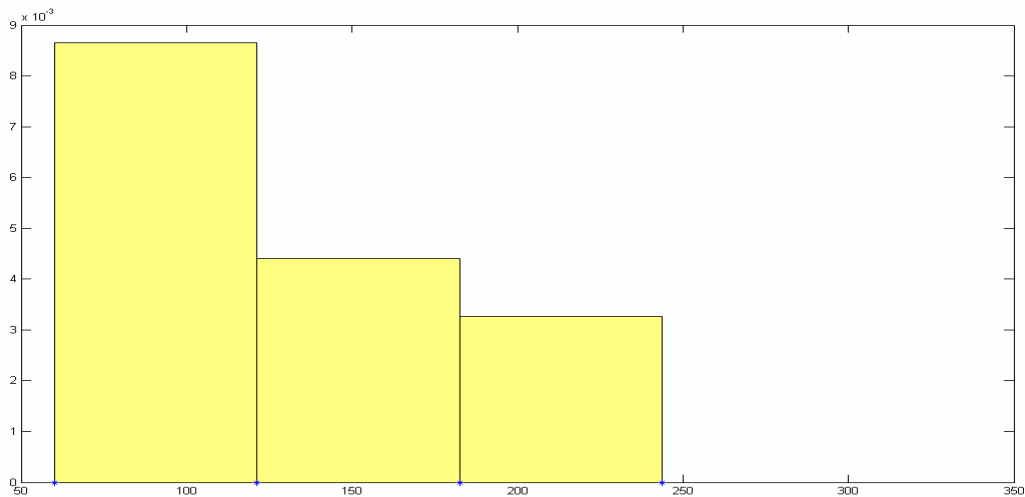
$$C_1 = [60,00; 121,27]$$

$$C_2 =] 121,27; 182,53]$$

$$C_3 =]182,53; 243,80]$$

La relativa distribuzione di frequenza è:

C_i	n_i	f_i	N_i	F_i	a_i	d_i
[60,00; 121,27]	8	0,53	8	0,53	61,27	0,00865
] 121,27; 182,53]	4	0,27	12	0,80	61,27	0,00441
]182,53; 243,80]	3	0,20	15	1,00	61,27	0,00326
Totali	15	1,00				



L'istogramma normalizzato mostra sulle ascisse gli estremi delle classi e sulle ordinate la densità di frequenza, calcolata come rapporto tra la frequenza relativa di ciascuna classe e l'ampiezza della classe stessa:

$$d_i = \frac{f_i}{a_i}$$

N.B.: essendo le classi di ampiezza costante l'istogramma normalizzato presenta la stessa forma dell'istogramma che ha sulle ascisse le frequenze relative o le assolute.

Per quanto riguarda il secondo gruppo:

$$\text{range (PREZZO EFG)} = 205,00 - 22,40 = 182,60$$

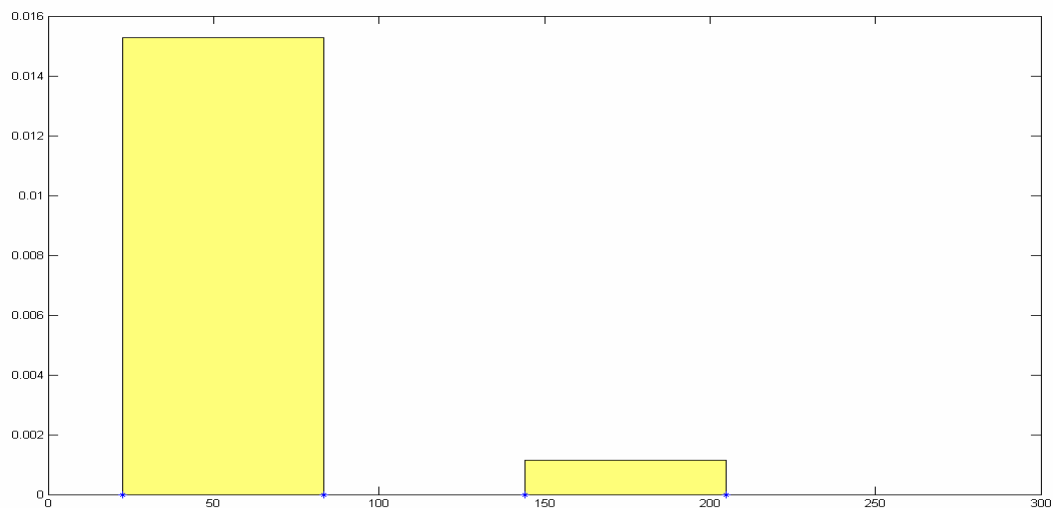
$$a_{\text{EFG}} = \frac{\text{range}}{3} = \frac{182,60}{3} = 60,86667$$

$$C_1 = [22,40; 83,27]$$

$$C_2 =] 83,27; 144,13]$$

$$C_3 =]144,13; 205,00]$$

C_i	n_i	f_i	N_i	F_i	a_i	dn_i
[22,40; 83,27]	14	0,93	14	0,93	60,867	0,01528
] 83,27; 144,13]	0	0,00	14	0,93	60,867	0,00
]144,13; 205,00]	1	0,07	15	1,00	60,867	0,00115
Totali	15	1,00				



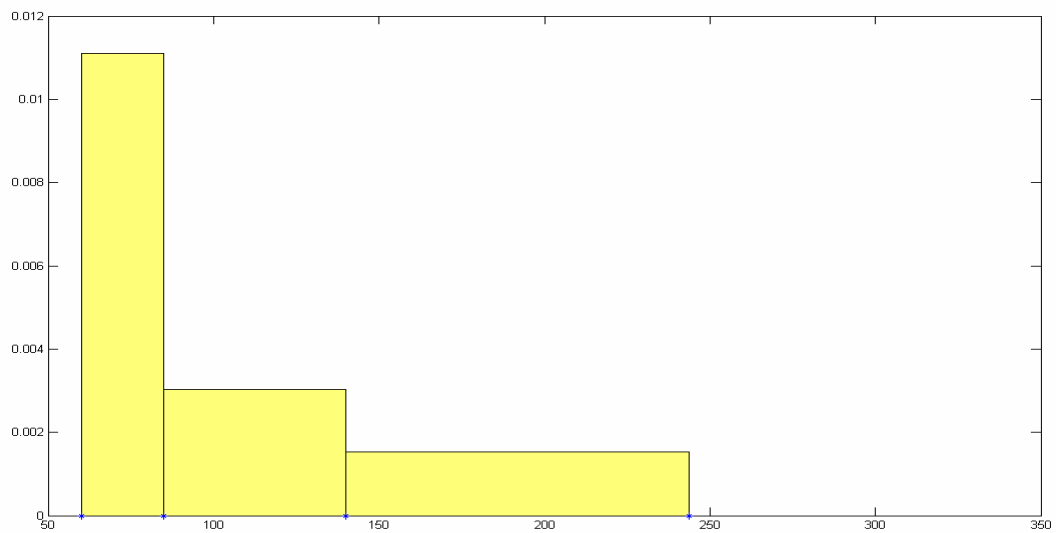
4.2)

Le classi equifrequenti sono costituite dallo stesso numero di elementi. Basta quindi dividere la numerosità totale di ciascuna distribuzione semplice, pari in entrambi i casi a 15, per il numero di classi desiderate, ossia 3.

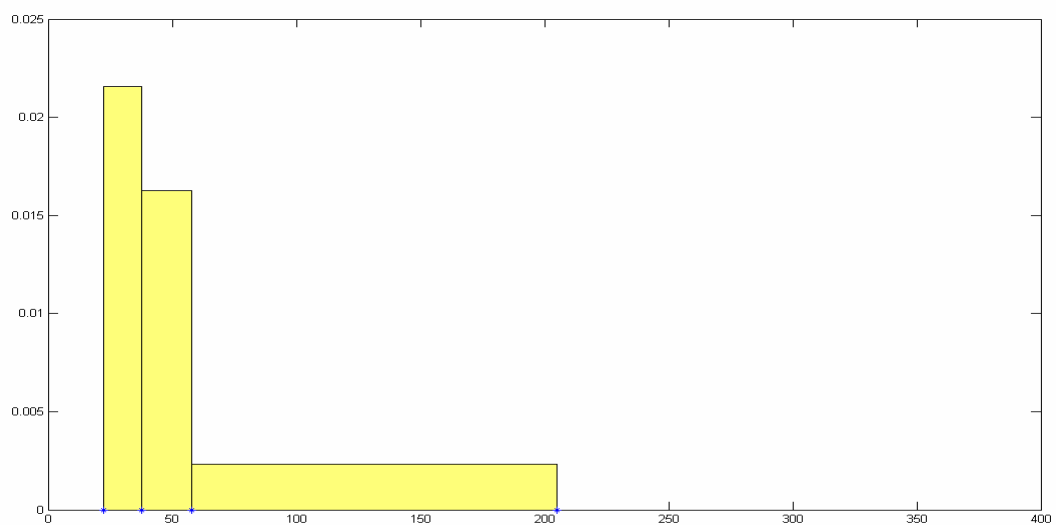
La frequenza costante all'interno delle classi sarà, dunque, pari a 5.

Le due distribuzioni di frequenza risultanti sono le seguenti:

C_i	n_i	f_i	N_i	F_i	a_i	dn_i
[60,00; 85,00]	5	0,33	5	0,33	15,00	0,01111
] 85,00; 140,00]	5	0,33	10	0,66	55,00	0,00303
]140,00; 243,80]	5	0,34	15	1,00	100,80	0,00153
Totali	15	1,00				



C_i	n_i	f_i	N_i	F_i	a_i	dn_i
[22,40; 37,70]	5	0,33	5	0,33	15,3	0,021568
] 37,70; 58,00]	5	0,33	10	0,66	20,3	0,016256
]58,00; 205,00]	5	0,34	15	1,00	147,0	0,002313
Totali	15	1,00				



5)

5.1)

- Media aritmetica

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{N} = \frac{680,57}{30} = 22,69$$

- Mediana

Innanzitutto è necessario ordinare i valori:

	Valore stimato del terreno
1	5,27
2	7,12
3	7,15
4	7,55
5	7,74
6	8,37
7	10,13
8	10,70
9	12,04
10	12,22
11	14,49
12	14,59
13	15,06
14	15,44
15	→ 16,10
16	→ 16,20
17	16,93
18	17,95
19	20,80
20	25,52
21	26,52
22	26,60
23	30,28
24	30,41
25	31,68
26	44,01
27	47,73
28	55,93
29	59,80
30	66,24

Essendo $N = 30$ un numero pari, la mediana è data dalla media aritmetica dei due elementi di posto $N/2$ ed $(N/2) + 1$, ossia:

$$Me = \frac{x_{(N/2)} + x_{(\frac{N}{2}+1)}}{2} = \frac{x_{(15)} + x_{(16)}}{2} = \frac{16,10 + 16,20}{2} = 16,15$$

- Moda

Non essendovi alcun valore che si presenta più di una volta, la moda non è determinabile.

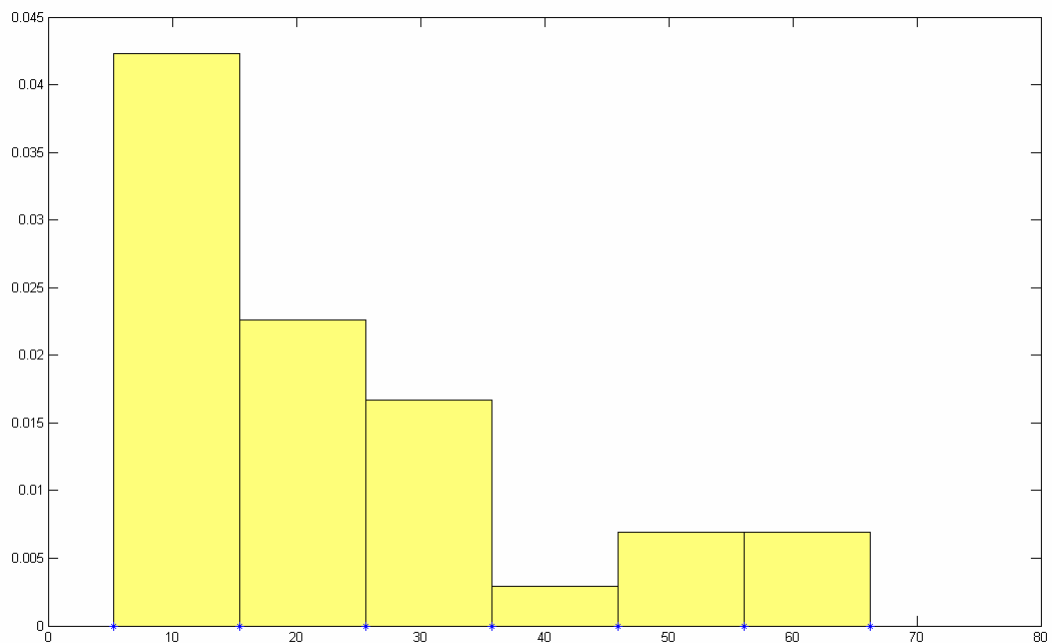
5.2) L'ampiezza costante delle 6 classi derivanti dalla distribuzione semplice del carattere si ottiene come:

$$\text{range (VALORE DEL TERRENO)} = 66,24 - 5,27 = 60,97$$

$$a = \frac{\text{range}}{6} = \frac{60,97}{6} = 10,16167$$

La distribuzione di frequenza ottenuta suddividendo le osservazioni in 6 classi equiampie è:

C_i	c_i	n_i	f_i	N_i	F_i	a	d
[5,27; 15,43]	10,35	13	0,43	13	0,43	10,16	0,0423
]15,43; 25,59]	20,51	7	0,23	20	0,67	10,16	0,0226
]25,59; 35,76]	30,67	5	0,17	25	0,83	10,16	0,0167
]35,76; 45,92]	40,84	1	0,03	26	0,87	10,16	0,0030
]45,92; 56,08]	51,00	2	0,07	28	0,93	10,16	0,0069
]56,08; 66,24]	61,16	2	0,07	30	1,00	10,16	0,0069
Totali		30	1,00				



- Media aritmetica

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^n c_i n_i}{N} = \\ &= \frac{(10,35 \times 13) + (20,51 \times 7) + (30,67 \times 5) + 40,84 + (51 \times 2) + (61,16 \times 2)}{30} = \\ &= \frac{696,67}{30} = 23,22\end{aligned}$$

- Moda

Essendo le classi equiampie, la classe modale è semplicemente quella con la frequenza più elevata, ossia la classe di estremi $Mo^- = 5,27$ ed $Mo^+ = 15,43$.

- Mediana

Alla mediana corrisponde una frequenza relativa cumulata F_{Me} pari a 0,5. In base alle frequenze cumulate si evince che la classe mediana è la seconda.

All'interno di questa, il valore della mediana viene calcolato come segue:

$$Me = x_{Me-1} + \frac{x_{Me} - x_{Me-1}}{F_{Me} - F_{Me-1}} (0,5 - F_{Me-1})$$

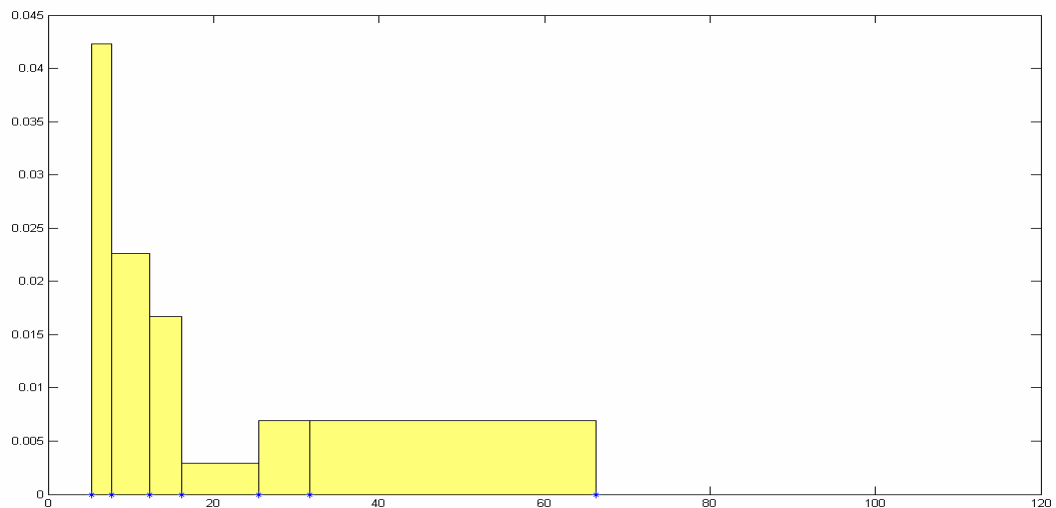
Sostituendo:

$$\begin{aligned}Me &= 15,43 + \frac{25,59 - 15,43}{0,67 - 0,43} \times (0,5 - 0,43) = \\ &= 15,43 + \frac{10,16}{0,24} \times 0,07 = 15,43 + 2,9 = 18,33\end{aligned}$$

5.3)

La distribuzione di frequenze ottenuta suddividendo le osservazioni in 6 classi equifrequenti è:

C_i	c_i	n_i	f_i	N_c	F_c	a_i	d_i
[5,27; 7,74]	6,51	5	0,1 $\bar{6}$	5	0,17	2,47	2,02
]7,74; 12,22]	10,29	5	0,1 $\bar{6}$	10	0,33	3,85	1,30
]12,22; 16,10]	15,29	5	0,1 $\bar{6}$	15	0,50	1,61	3,11
]16,10; 25,52]	20,86	5	0,1 $\bar{6}$	20	0,67	9,32	0,54
]25,52; 31,68]	29,10	5	0,1 $\bar{6}$	25	0,83	5,16	0,97
]31,68; 66,24]	55,13	5	0,1 $\bar{6}$	30	1	22,23	0,22
Totali		30	1,00				



▪ Media

$$\bar{x} = \sum_{i=1}^n c_i f_i =$$

$$= 6,51 \times 0,1\bar{6} + 10,29 \times 0,1\bar{6} + 15,29 \times 0,1\bar{6} + 20,86 \times 0,1\bar{6} + 29,10 \times 0,1\bar{6} + 55,13 \times 0,1\bar{6} = 22,86$$

▪ Moda

In base alla densità di frequenza, la classe modale è la terza, ossia la classe]12,22; 16,10]

▪ Mediana

Essendo la frequenza relativa cumulata F_3 delle prime 3 classi pari esattamente a 0,50, la mediana è rappresentata dall'estremo superiore della classe 3, ossia dal valore 16,10