

Compito 2di giugno 2013: risoluzione esercizi 1 e 2

1. Cinque monete vengono lanciate 1000 contemporaneamente: la seguente tabella riporta a sinistra il possibile numero di teste che possono uscire in un lancio di 5 monete e a destra il numero di volte che nell'esperimento si è avuto effettivamente quel risultato

n. teste	fr. ass
0	38
1	144
2	342
3	287
4	164
5	25
tot	1000

1.1. Si indichi qual è il carattere indagato e se ne discuta la natura.

1.2. Si costruiscano le distribuzioni di frequenza relative e frequenza cumulata.

1.3. Dire cosa indica la frequenza cumulata a 3.

1.4. Si calcoli la moda e la mediana. Si determinino i quartili.

1.5. Si calcoli il numero medio di teste e lo scarto quadratico medio da essa.

1.6. Si utilizzino due diverse rappresentazioni grafiche per le distribuzioni di frequenza assoluta e relativa.

1.7 dare una rappresentazione grafica della distribuzione di frequenze cumulate

1.1 Il carattere indagato è il numero di teste uscito in un lancio di 5 monete, ed è un carattere quantitativo discreto (i casi, o unità statistiche, sono i lanci di 5 monete)

1.2

n. teste	fr. ass	fr. rel	fr. ass. cum
0	38	0,038	38
1	144	0,144	182
2	342	0,342	524
3	287	0,287	811
4	164	0,164	975
5	25	0,025	1000
tot	1000	1	

1.3. La frequenza cumulata a 3 indica **il numero dei casi (lanci) in cui il numero di teste uscito non supera 3**; dalla tabella leggiamo che la frequenza cumulata a 3 vale **811**.

1.4.

Moda 2 (2 è il valore con la frequenza più alta)

Mediana 2 corrisponde alla prima frequenza cumulata (524) che è \geq dei numeri 500 e 501 e quindi comprende più della metà dei casi

Quartili:

$q_1 = 2$ corrisponde alla prima frequenza cumulata (524) che è \geq dei numeri 250 e 251 (posti $n/4$ e $(n/4) + 1$) e comprende più di $1/4$ dei casi

$q_2 = 2$ = **mediana** (vedi mediana)

$q_3 = 3$ corrisponde alla prima frequenza cumulata (811) che è \geq dei numeri 750 e 751 (posti $3n/4$ e $(3n/4) + 1$) e comprende più dei $3/4$ dei casi)

1.5.

Num. medio teste =

$$\bar{x} = \frac{38 \cdot 0 + 144 \cdot 1 + 342 \cdot 2 + 287 \cdot 3 + 164 \cdot 4 + 25 \cdot 5}{1000} = \frac{144 + 684 + 861 + 656 + 125}{1000} = \frac{2470}{1000} = 2,47$$

varianza

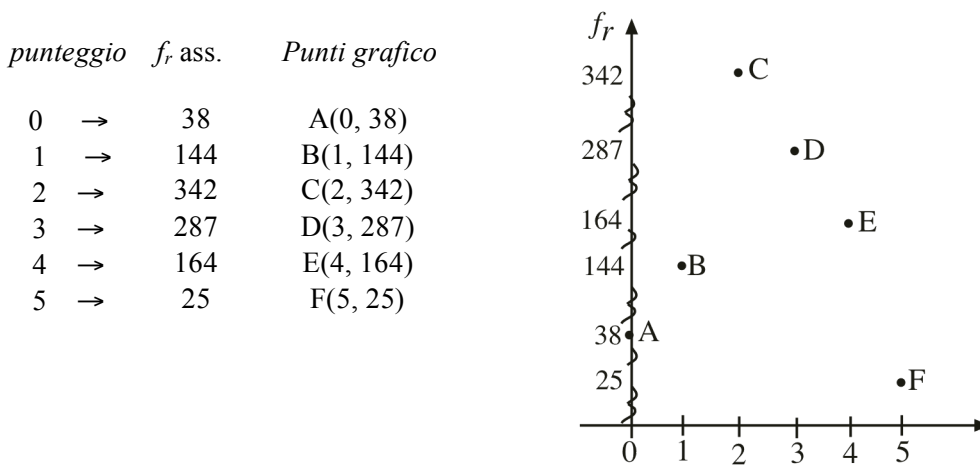
$$\sigma^2 = \frac{38 \cdot (0-2,47)^2 + 144 \cdot (1-2,47)^2 + 342 \cdot (2-2,47)^2 + 287 \cdot (3-2,47)^2 + 164 \cdot (4-2,47)^2 + 25 \cdot (5-2,47)^2}{1000}$$
$$= \frac{231,8342 + 311,1696 + 75,5478 + 80,6183 + 383,9076 + 160,0225}{1000} = \frac{1243,1}{1000} = 1,2431$$

scarto quadratico medio $\sigma = 1,1149 \dots$

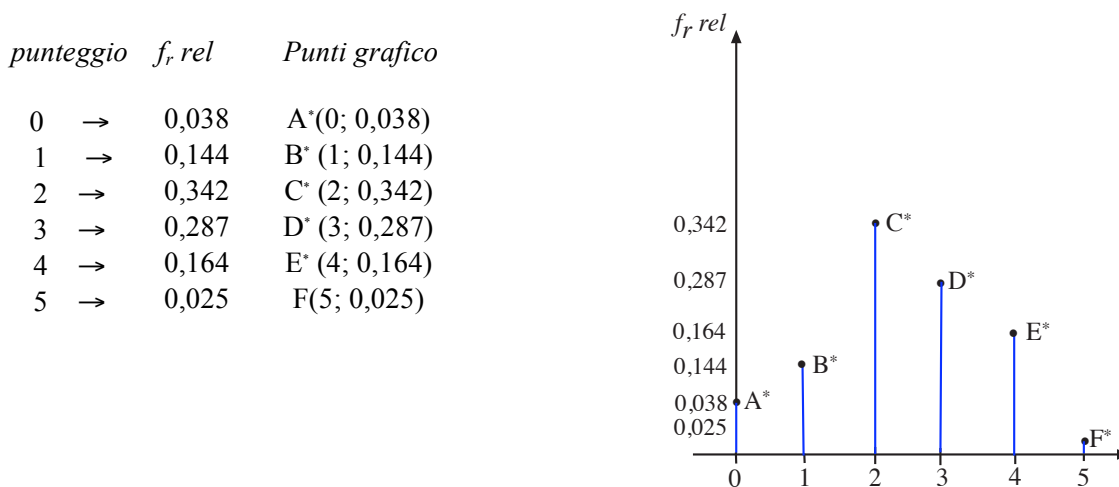
Coefficiente di variazione: $CV = \frac{\sigma}{\bar{x}} = \frac{1,1149 \dots}{2,47} \cong 0,4514$

Calcoliamo tale coefficiente perché ci serve per rispondere all'ultimo quesito del prossimo esercizio

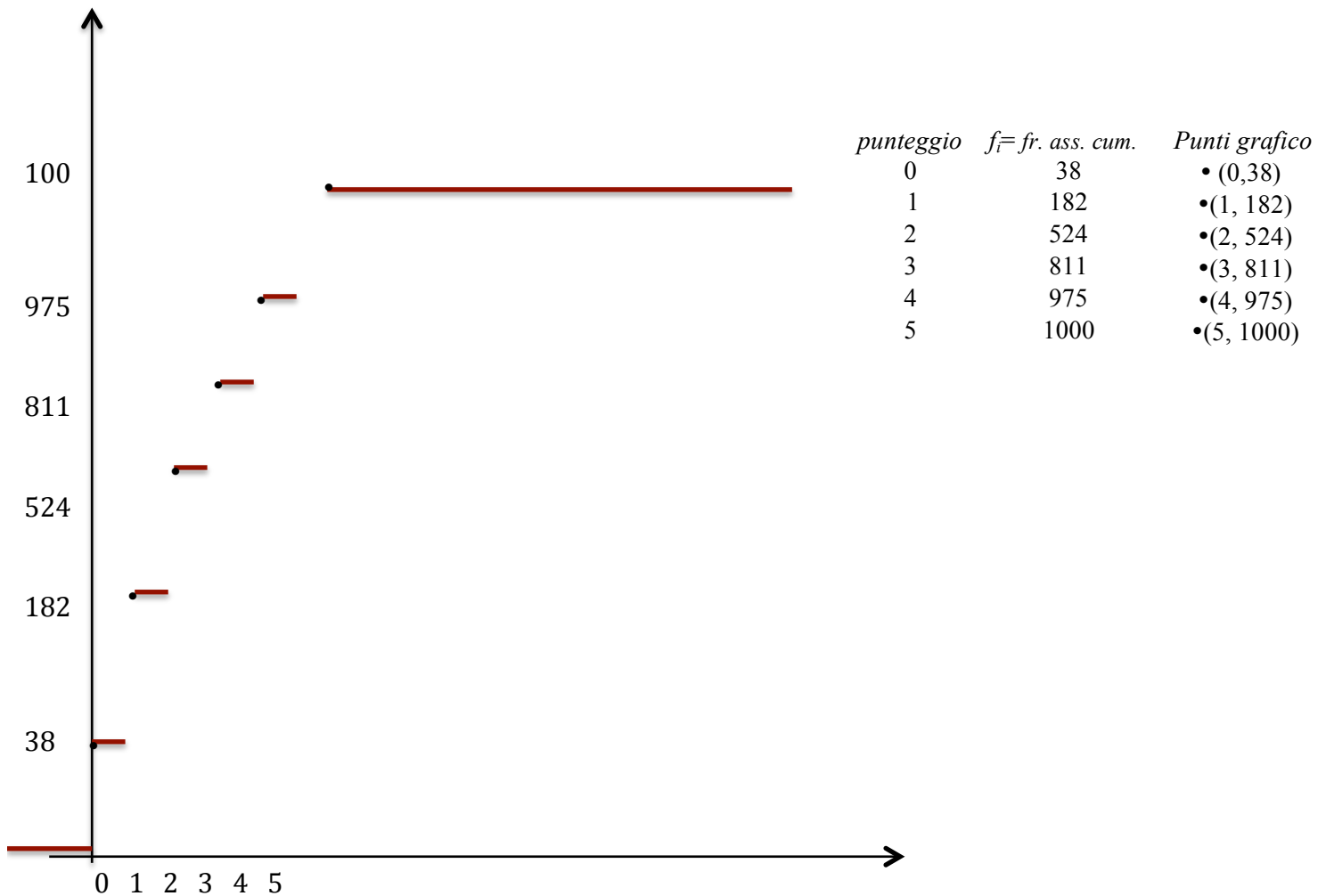
1.6 Grafico cartesiano per la distribuzione delle frequenze assolute, che è il grafico della funzione



Avendo utilizzato il grafico cartesiano semplice per la distribuzione delle frequenze assolute, possiamo utilizzare il grafico cartesiano a pettine per la distribuzione delle frequenze relative:



1.7 grafico della distribuzione di frequenza assoluta cumulata



2. La seguente tabella riporta, raggruppate in classi, le lunghezze in mm. di alcuni lombrichi

Classe di valori	$fr. ass$
$16 < x \leq 17$	1
$18 < x \leq 19$	8
$20 < x \leq 21$	5
$22 < x \leq 23$	3
$24 < x \leq 25$	3
Tot	20

- 2.1 Dire se le classi sono indicate con i limiti reali o formali e, se occorre, correggere le classi
- 2.2 Dare la rappresentazione grafica della distribuzione di frequenza e del poligono delle frequenze.
- 2.3 Determinare la distribuzione di frequenze relative, la funzione di ripartizione e dare di quest'ultima una rappresentazione grafica.
- 2.4 Calcolare la mediana e indicarla sull'ogiva.
- 2.5 Calcolare la media, e lo scarto quadratico medio.
- 2.6 Confrontare la variabilità di tale distribuzione in classi con quella del numero di teste di cui all'esercizio precedente.

2.1 Le classi sono indicate attraverso limiti formali, perché il secondo estremo di una classe non coincide con il primo estremo della classe successiva. Calcolata la distanza d tra secondo estremo di una classe e primo estremo della classe successiva, per passare ai limiti reali occorre aumentare ogni secondo estremo di $d/2$ e diminuire ogni primo estremo di $d/2$: si ottengono così classi contigue. Nel nostro caso $d=1$

Classe di valori	fr. ass
16 <x≤ 17	1
18 <x≤ 19	8
20 <x≤ 21	5
22 <x≤ 23	3
24 <x≤ 25	3
Tot	20

→

Classe di valori	fr. ass
15,5 < x ≤ 17,5	1
17,5 < x ≤ 19,5	8
19,5 < x ≤ 21,5	5
21,5 < x ≤ 23,5	3
23,5 < x ≤ 25,5	3
Tot	20

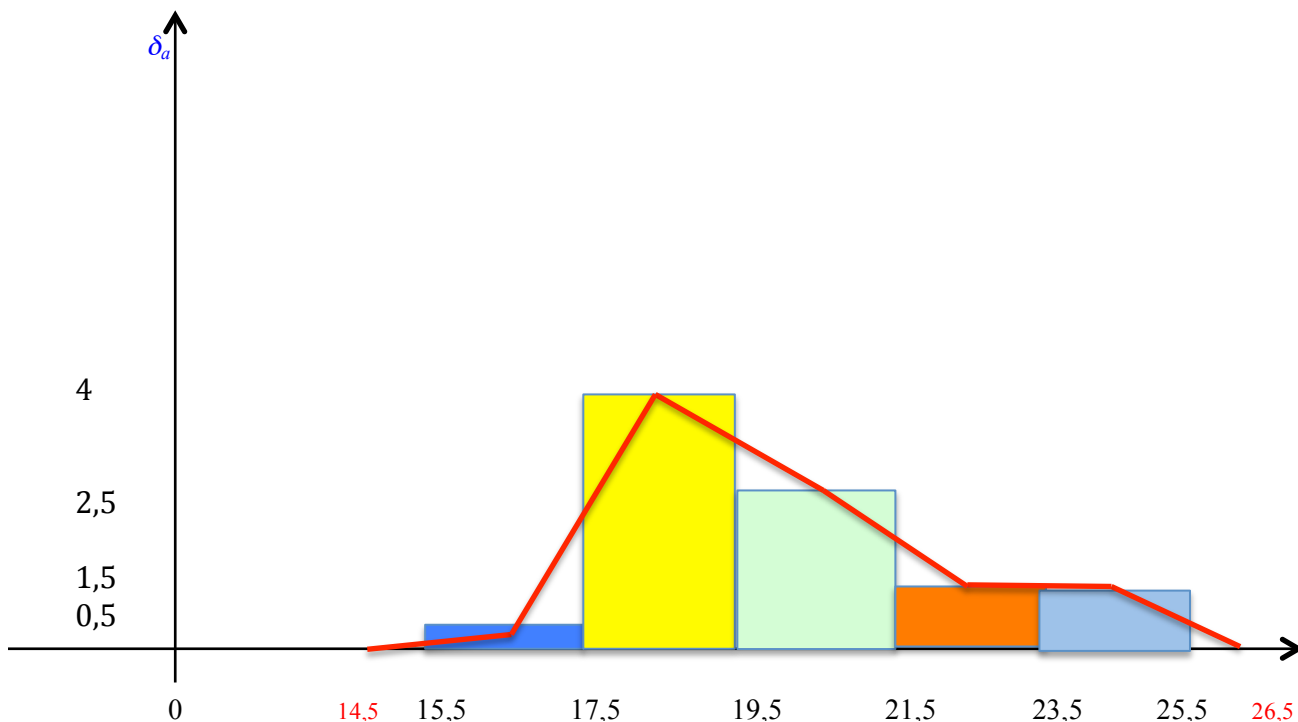
↔

Classe di valori	fr. ass
15,5 --- 17,5	1
17,5 --- 19,5	8
19,5 --- 21,5	5
21,5 --- 23,5	3
23,5 --- 25,5	3
Tot	20

2.2. Costruiamo l'istogramma riportando sull'asse delle ascisse di un riferimento cartesiano le classi dei valori indicate attraverso i limiti reali (le classi coincidono allora con intervalli contigui della retta reale) e innalzando su ogni classe un rettangolo di altezza pari alla densità di frequenza assoluta (freq.ass./ampiezza classe): in questo modo l'area del rettangolo è uguale alla frequenza corrispondente alla classe. calcoliamo quindi la densità assoluta per ogni classe

Classe di valori	ampiezza	fr. ass	$\delta_a = \text{fr. ass} / \text{ampiezza}$
15,5 --- 17,5	2	1	$\frac{1}{2} = 0,5$
17,5 --- 19,5	2	8	$\frac{8}{2} = 4$
19,5 --- 21,5	2	5	$\frac{5}{2} = 2,5$
21,5 --- 23,5	2	3	$\frac{3}{2} = 1,5$
23,5 --- 25,5	2	3	$\frac{3}{2} = 1,5$
		20	20

Disegniamo l'istogramma riportando la densità assoluta sull'asse delle ordinate



Poiché le classi sono tutte di uguale ampiezza (2), per disegnare il poligono delle frequenze (linea rossa) uniamo i punti di mezzo delle basi superiori dei rettangoli ottenendo dei triangolini di base 1; completiamo unendo il punto medio della base superiore del primo rettangolo con un punto dell'asse delle ascisse che precede il rettangolo di 1 (metà ampiezza della base) e il punto medio della base superiore dell'ultimo rettangolo con il punto dell'asse delle ascisse che segue il rettangolo di 1.

2.3

Classe di valori	fr. ass	fr. rel	F
15,5 --- 17,5	1	0,05	0,05
17,5 --- 19,5	8	0,4	0,45
19,5 --- 21,5	5	0,25	0,70
21,5 --- 23,5	3	0,15	0,85
23,5 --- 25,5	3	0,15	1
Tot	20	1	

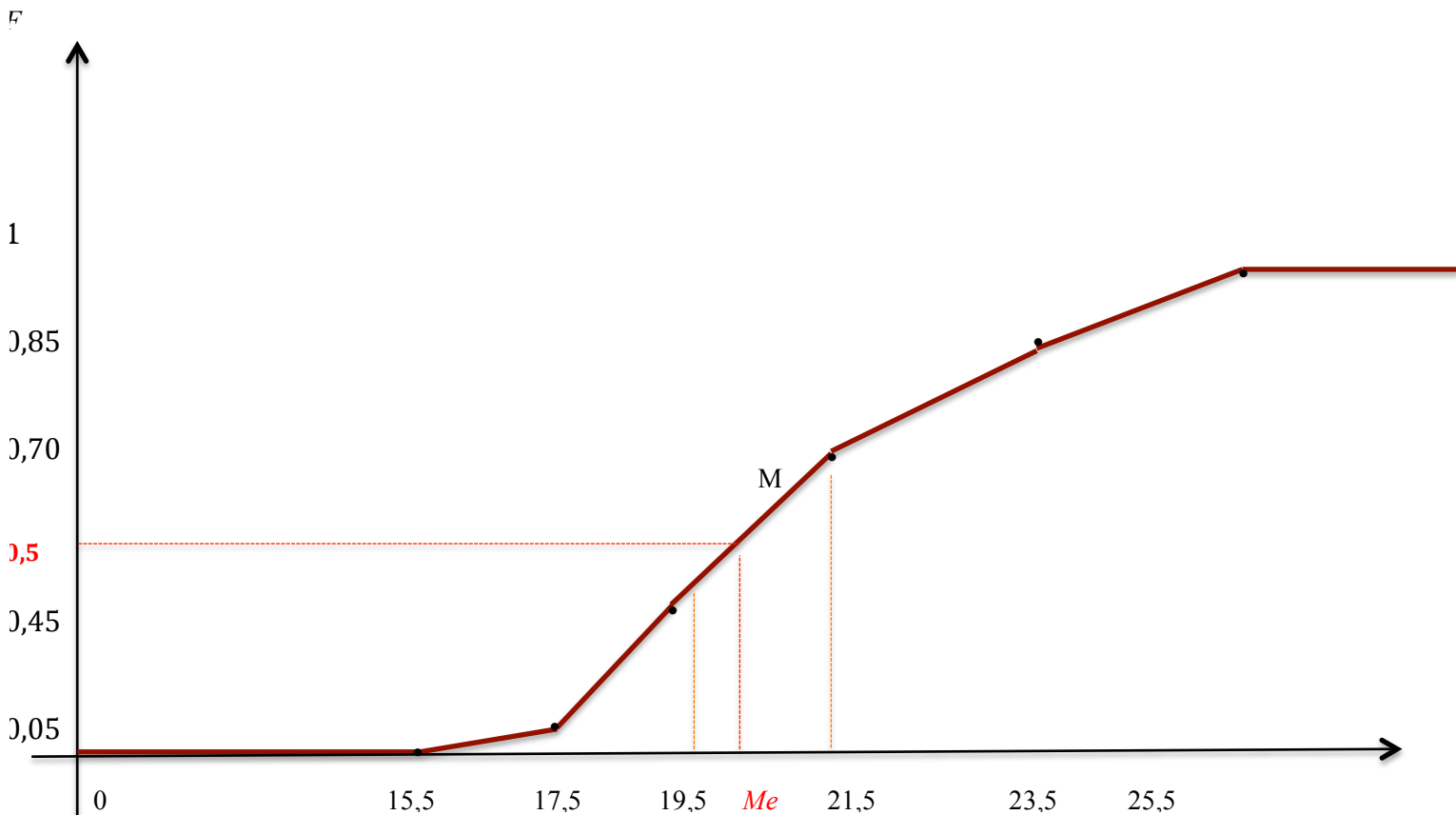
Il valore $F(x)$ della funzione di ripartizione indica la frequenza relativa cumulata a x ; nel caso in cui le classi sono chiuse a destra, i valori nell'ultima colonna indicano in realtà le frequenze relative cumulate agli estremi superiori delle classi (cioè la porzione dei casi in cui il valore non supera l'estremo superiore della classe).

x	$F(x)$	punti grafico
15,5	0	• (15,5; 0)
17,5	0.05	• (17,5; 0.05)
19,5	0.45	• (19,5; 0.45)
21,5	0.70	• (21,5; 0.70)
23,5	0.85	• (23,5; 0.85)
25,5	1	• (25,5; 1)

Non conosciamo i valori di $F(x)$ nei punti interni alle classi; se però ipotizziamo che all'interno di ogni classe la distribuzione delle frequenze assolute sia uniforme, ciò vuol dire che $F(x)$ cresce linearmente nella classe e allora il grafico di F su ogni classe è un segmento. Il grafico completo è l'ogiva che si ottiene coi passi seguenti:

1° Riportiamo nel piano cartesiano i punti $(x, F(x))$ dove gli x indicano gli estremi superiori delle classi e il punto $(15,5, 0)$, dove il primo 15,5 è il primo estremo della prima classe e il secondo 0 è la frequenza relativa cumulata che gli associamo.

2° Congiungiamo i punti consecutivi (cioè con ascisse che sono estremi della stessa classe) con dei segmenti e prolunghiamo il grafico ottenuto a destra e a sinistra con semirette parallele all'asse delle ascisse.



2.4.

classe mediana 19,5--|21 (è la classe cui corrisponde la prima frequenza cumulata relativa che supera 0,5= 1/2 e che quindi comprende il 50% dei casi)

Calcoliamo la mediana applicando la formula con le frequenze cumulate relative

Mediana:

$$Me = 19,5 + \frac{21,5-19,5}{0,70-0,45} (0,50-0,45) = 19,5 + \frac{2}{0,25} (0,05) = 19,5 + 0,4 = 19,9$$

sull'ogiva: la mediana è l'ascissa del punto M di ordinata $F(x) = 0,5$

2.5

Poiché non si conoscono i valori effettivamente presi in ogni classe, non si può calcolare la media aritmetica esatta; si ottiene una stima di essa operando come se tutti i valori presi in una classe coincidessero con il valore centrale della classe (punto medio della classe). Calcoliamo quindi i valori centrali delle classi che indichiamo con c_i .

Classe di valori	fr. ass	fr. rel	F	c_i =val. centr	fr ass $\cdot c_i$	$c_i - \bar{x} = c_i - 20,4$	$(c_i - \bar{x})^2$	$f_i (c_i - \bar{x})^2$
15,5 --- 17,5	1	0,05	0,05	16,5	16,5	16,5-20,4=-3,9	15,21	1•15,21= 15,21
17,5 --- 19,5	8	0,4	0,45	18,5	148	18,5-20,4=-1,9	3,61	8•3,61= 28,88
19,5 --- 21,5	5	0,25	0,70	20,5	102,5	20,5-20,4= 0,1	0,01	5•0,01= 0,05
21,5 --- 23,5	3	0,15	0,85	22,5	67,5	22,5-20,4 =2,1	4,41	3•4,41= 13,23
23,5 --- 25,5	3	0,15	1	24,5	73,5	24,5-20,4 =4,1	16,81	3•16,81= 50,43
Tot	20	1			Tot=408			$\sum_i f_i (c_i - \bar{x})^2 = 107,8$

$$\text{Stima Media arit.: } \bar{x} = \frac{1 \cdot 16,5 + 8 \cdot 18,5 + 5 \cdot 20,5 + 3 \cdot 22,5 + 3 \cdot 24,5}{20} = \frac{16,5 + 148 + 102,5 + 67,5 + 73,5}{20} = \frac{408}{20} = 20,4$$

$$\text{Devianza} = \sum_i f_i (c_i - \bar{x})^2 = 1 \cdot 15,21 + 8 \cdot 3,61 + 5 \cdot 0,01 + 3 \cdot 4,41 + 3 \cdot 16,81 = 107,8$$

$$\text{Varianza } \sigma^2 = \frac{\sum_i f_i (c_i - \bar{x})^2}{\sum_i f_i} = \frac{107,8}{20} = 5,39$$

$$\text{Scarto quadratico medio: } \sigma = \sqrt{\frac{\sum_i f_i (c_i - \bar{x})^2}{\sum_i f_i}} = \sqrt{5,39} \approx 2,32163$$

2.6. per confrontare le variabilità delle due distribuzioni confrontiamo i coefficienti di variazione

$$\text{coeff. Variazione distribuzione dell'esercizio 1: } CV = \frac{\sigma}{\bar{x}} = \frac{1,1149...}{2,47} \approx 0,4514$$

$$\text{coeff. Variazione distribuzione dell'esercizio 2: } CV = \frac{\sigma}{\bar{x}} = \frac{2,32163}{20,4} \approx 0,1138$$

Poiché il primo coefficiente è maggiore del secondo è la prima distribuzione a presentare maggiore variabilità