

Università degli Studi di Napoli «Federico II»
Facoltà di Architettura

Upta

Corso di laurea in Urbanistica e Scienze della Pianificazione Territoriale e Ambientale
AA 2014-2015

Corso integrato di Matematica e statistica

Docente: Livia D'Apuzzo.

Statistica Descrittiva Univariata

3. - 1ª parte- Distribuzioni di frequenza e di intensità: rappresentazione grafica nel caso di caratteri quantitativi

RAPPRESENTAZIONE GRAFICA DELLE DISTRIBUZIONI DI FREQUENZA o INTENSITÀ PER CARATTERI QUANTITATIVI

Un carattere discreto quantitativo o variabile può essere discreto (e prende cioè un numero finito o numerabile di valori) o continuo (l'insieme X dei possibili valori è un intervallo della retta dei numeri reali). Per entrambi i casi abbiamo considerato le distribuzioni di frequenza e/o intensità assoluta, relativa o percentuale e abbiamo dato di tali distribuzioni una rappresentazione mediante tabelle (cfr. fascicolo 2). In questo fascicolo diamo la rappresentazione grafica di tali distribuzioni distinguendo il caso in cui la variabile è discreta con un numero finito e gestibile di valori distinti, $x_1, x_2, \dots, x_k$, dal caso in cui la variabile è continua o discreta ma trattata come continua perché con grande quantità di valori dispersi: in ogni caso la rappresentazione è operata in un piano con riferimento cartesiano ortogonale e i valori o le classi dei valori sono riportati sull'asse delle ascisse.

1. RAPPRESENTAZIONE GRAFICA DELLE DISTRIBUZIONI DI FREQUENZE E INTENSITÀ

PER VARIABILI DISCRETE CON UN NUMERO FINITO DI VALORI DISTINTI $x_1, x_2, \dots, x_k$,

Una distribuzione di frequenza o intensità assoluta o relativa o percentuale riferita ad un carattere discreto con valori $x_1, x_2, \dots, x_k$, è una funzione definita nell'insieme $X = \{x_1, x_2, \dots, x_k\}$ dei valori rilevati e che associa ad ogni x_i un numero reale $f(x_i)$ che, a seconda della distribuzione considerata, ha il significato di frequenza o intensità

$$f: x_i \in X \rightarrow f(x_i) \in R \quad (1)$$

R insieme dei numeri reali

$x_i = \text{valore}$	$f(x_i)$	<i>rappresentazione tabellare di f</i>
x_1	$f(x_1)$	
x_2	$f(x_2)$	
...		
x_k	$f(x_k)$	

Come rappresentare graficamente f ?

Poiché per ogni $i=1,2,\dots,k$, x_i e $f(x_i)$ sono numeri reali, la funzione f è una **funzione reale di una variabile reale**, la cui naturale rappresentazione nel piano, in cui è fissato un riferimento cartesiano ortogonale, è il **diagramma cartesiano ortogonale** che può essere completato nel così detto **diagramma cartesiano a canne d'organo**.

- il **diagramma cartesiano ortogonale**: insieme dei punti del piano di coordinate $(x_i, f(x_i))$, costruito portando i valori x_i del carattere sull'asse delle ascisse di un riferimento cartesiano ortogonale e le frequenze $f(x_i)$ di f sull'asse delle ordinate.
- il **diagramma cartesiano a canne d'organo detto anche a segmenti, ad aste, a pettine**: insieme dei segmenti, paralleli all'asse delle ordinate, che congiungono i punti P_i del diagramma cartesiano all'asse delle ascisse

Sia data la seguente

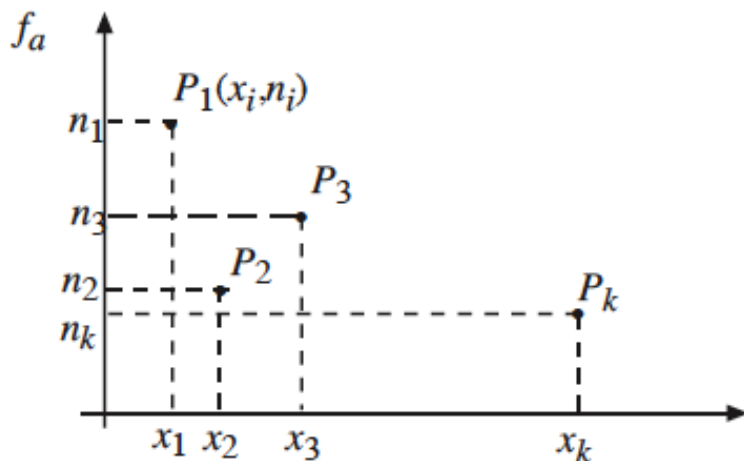
Tabella valori/ frequenze

Valori	Freq. ass. f_a $(x_i)=n_i$	Fre . rel. $f_r (x_i)=n_i/n$	Freq. perc. $f_{\%} (x_i)= (n_i/n)100$
x_1	n_1	n_1/n	$(n_1/n)100$
x_2	n_2	n_2/n	$(n_2/n)100$
...	...	...	...
x_k	n_k	n_k/n	$(n_k/n)100$
	n	1	100

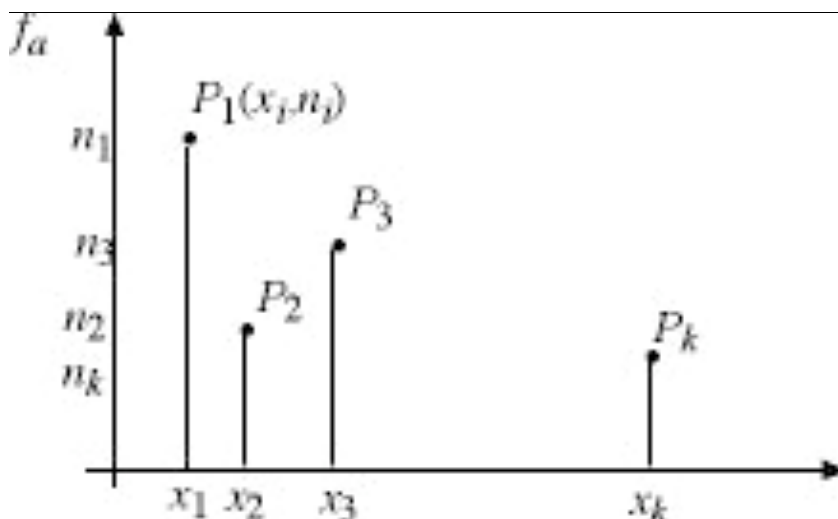
e sia f la seguente distribuzione di frequenze assolute

$$f_a: x_i \in \{x_1, x_2, \dots, x_k\} \rightarrow f_a(x_i) = n_i \quad \text{distribuzione semplice di frequenza assoluta}$$

Le due tipologie di rappresentazioni grafiche sono date dai seguenti diagrammi.



**diagramma cartesiano
ortogonale**



**diagramma cartesiano a
canne d'organo o a
pettine**

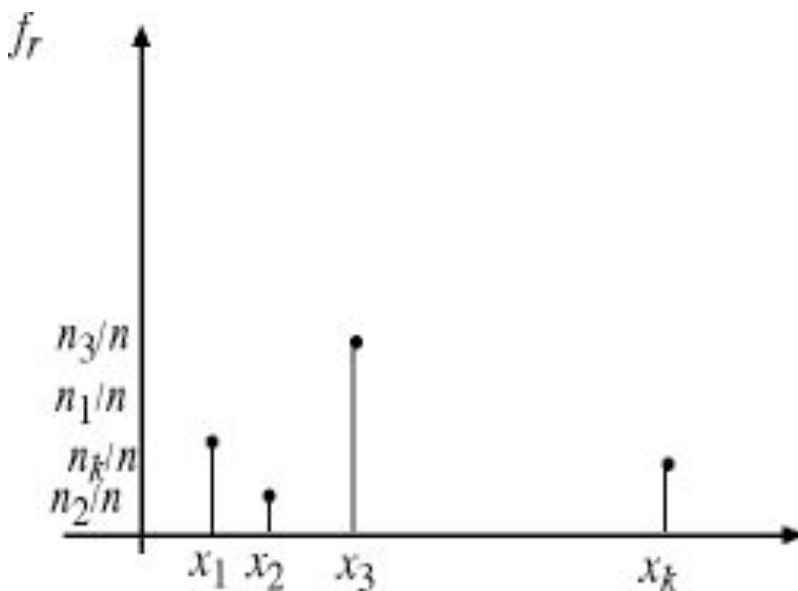
L'altezza del bastoncino è pari alla frequenza assoluta.

Le stesse rappresentazioni grafiche sono utilizzate per le distribuzioni di frequenze relative o percentuali o per le distribuzioni di intensità relative o percentuali

Consideriamo ad esempio una distribuzione di frequenze relative

$$f_r: x_i \in \{x_1, x_2, \dots, x_k\} \rightarrow f_r(x_i) = f_a(x_i) / n = n_i / n \quad \text{distribuzione semplice di frequenza relativa}$$

Portando i valori della variabile sull'asse delle ascisse e le frequenze relative sull'asse delle ordinate otteniamo i punti $(x_i, n_i/n)$ del grafico cartesiano e collegando questi punti con l'asse delle ascisse con segmenti ad esso perpendicolari otteniamo un diagramma come il seguente



**Diagramma cartesiano
a canne d'organo**

L'altezza del bastoncino
è pari alla frequenza relativa

Analoghe rappresentazioni grafiche si utilizzano per le distribuzioni di intensità, per le quali ad ogni valore del carattere è associata una quantità di interesse per quel valore

Esempio. 1

Indagine sull'età di dieci individui

Età	frequenze
22	2
23	1
24	3
25	2
27	2
totale	10

Grafico cartesiano

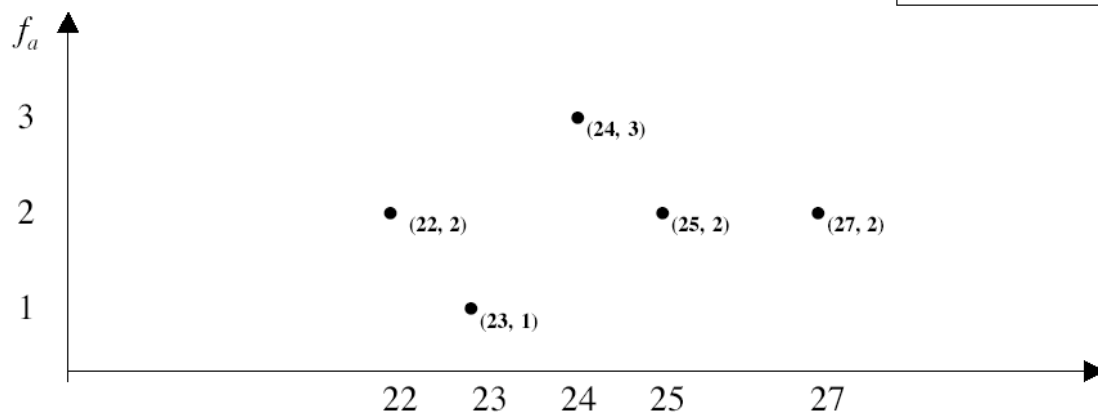
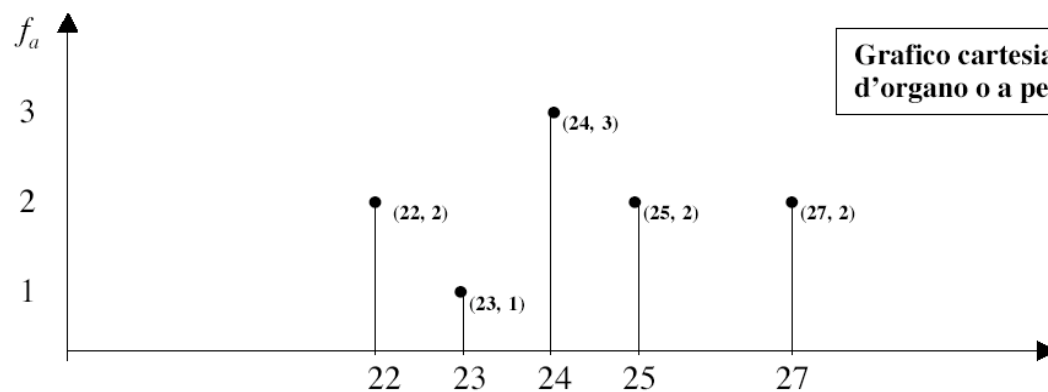


Grafico cartesiano a canne
d'organo o a pettine

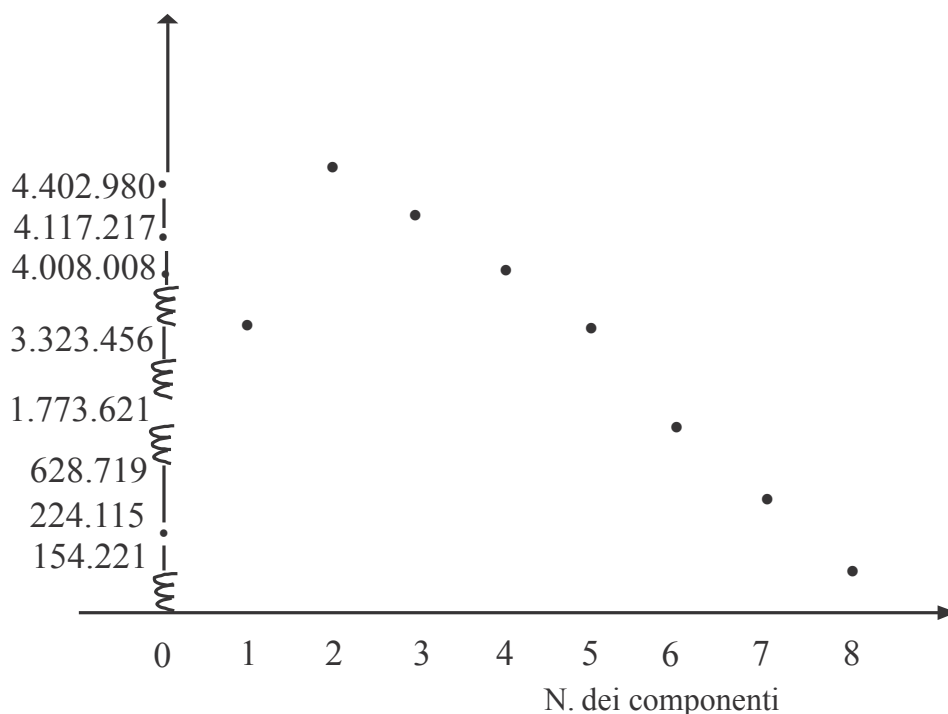


Esempio. 2 Indagine sulle famiglie residenti in Italia nel 1981 secondo il carattere “ numero componenti”

Censimento 1981 – Valori assoluti e percentuali

x_i = n. componenti	n_i = n. famiglie	$(n_i/n) \times 100$
1 componente	3.323.456	17,9
2 componenti	4.402.980	23,6
3 componenti	4.177.217	22,1
4 componenti	4.008.008	21,5
5 componenti	1.773.621	9,5
6 componenti	628.719	3,4
7 componenti	224.115	1,2
8 componenti	154.22	0,8

Insieme dei valori del carattere: $X=\{1, 2, 3, 4, 5, 6, 7,8\}$, $f_a(x_i)= n_i$

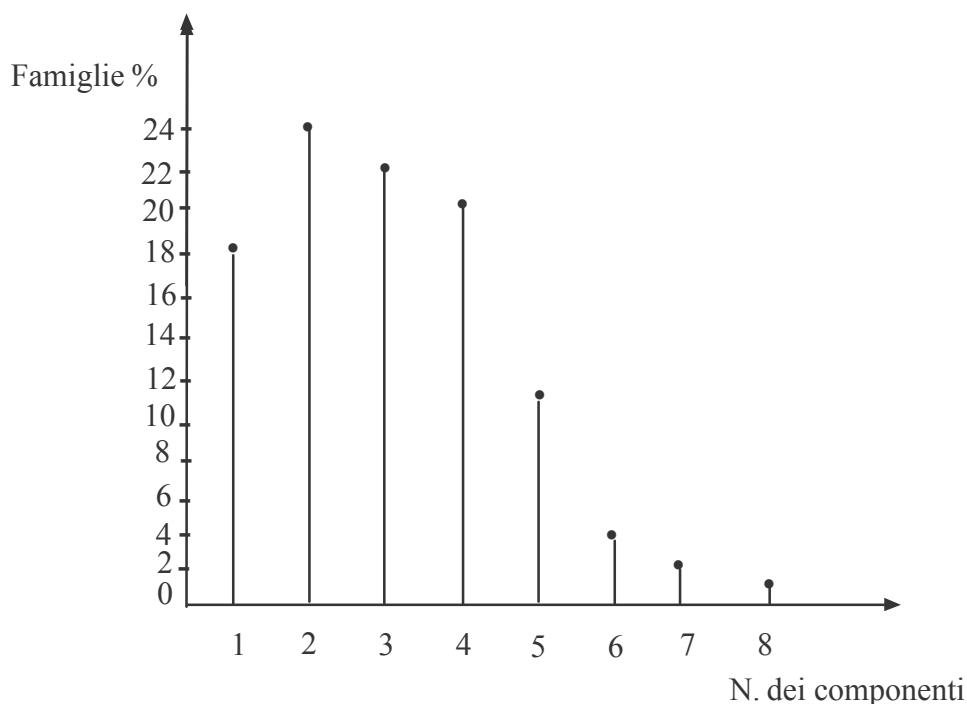


Per la disparità di grandezza tra i valori del carattere riportati sull'asse delle ascisse e le frequenze riportate sull'asse delle ordinate è impossibile rappresentare tutto il grafico della distribuzione di frequenza assoluta su un foglio il grafico: si disegna allora l'asse delle ordinate nel modo su indicato per dare l'idea che in alcuni tratti l'asse è stato ripiegato al fine di accorciare le distanze e far leggere le frequenze

Rappresentiamo ora la distribuzione delle frequenze percentuali con un diagramma a canne d'organo; per rappresentare in modo compatto tale diagramma è sufficiente scegliere su i due assi due diverse unità di misura

Censimento 1981 - Valori assoluti e percentuali

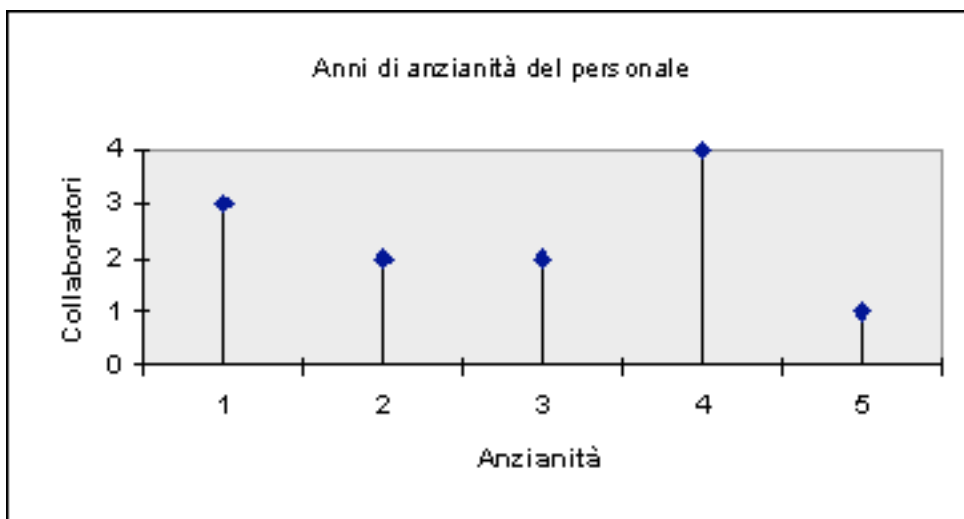
n. di componenti	<i>ni</i>	$(n_i/n) \times 100$ %
1 componente	3.323.456	17,9
2 componenti	4.402.980	23,6
3 componenti	4.177.217	22,1
4 componenti	4.008.008	21,5
5 componenti	1.773.621	9,5
6 componenti	628.719	3,4
7 componenti	224.115	1,2
8 e più	154.221	0,8



Esempio. 3. Indagine sugli anni di anzianità di lavoro del personale di una ditta di 12 dipendenti

Anzianità (x)	dipendenti $f_a(x)$
1	3
2	2
3	2
4	4
5	1
	12

Con Excell si ottiene il seguente diagramma ad aste



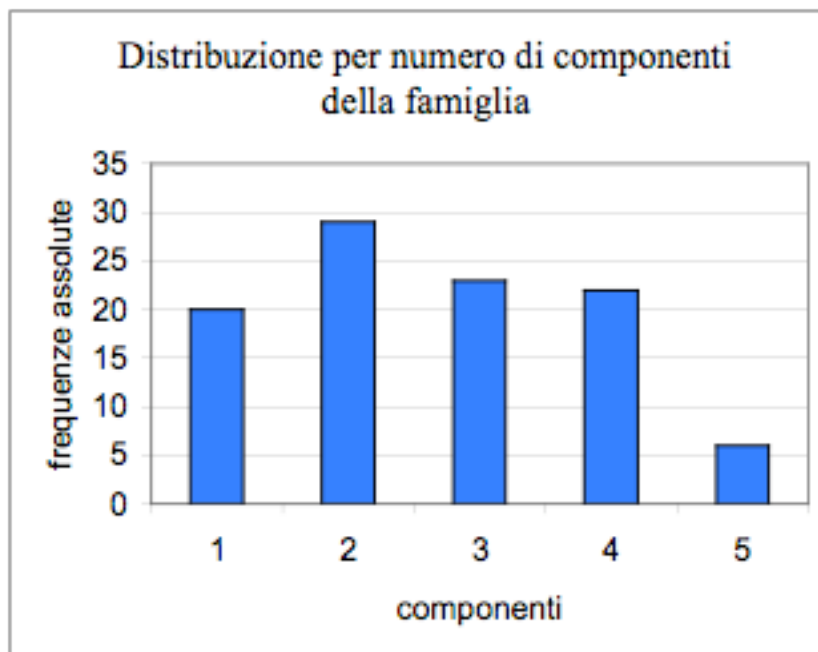
1.1 Una diversa rappresentazione per il carattere quantitativo discreto con valori interi

Alle tipologie illustrate di rappresentazioni grafica si aggiunge, talora e solo per variabili con valori interi, un tipo di rappresentazione grafica, il **diagramma a barre**, mutuato dalle rappresentazioni grafiche per caratteri qualitativi: a ogni valore del carattere si fa corrispondere un rettangolo la cui area è pari alla frequenza e i rettangoli, corrispondenti a singoli valori, sono poggiati su una stessa base; un regolo verticale ne misura le altezze. Per confrontare facilmente le aree e quindi le frequenze, i rettangoli associati ai diversi valori del carattere sono costruiti con basi uguali; in tal modo le misure delle altezze sono proporzionali alle aree ed anzi ad esse uguali se la basi si scelgono di misura 1.

La figura che si ottiene, a rettangoli separati, è detta **ortogramma a barre**.

Esempio. 4

Numero componenti	f_a
1	20
2	28
3	23
4	22
5	7
Tot.	100



Ortogramma a barre

2. RAPPRESENTAZIONE GRAFICA DELLE DISTRIBUZIONI DI FREQUENZE E PER CARATTERI CON VALORI RAGGRUPPATI IN CLASSI

Consideriamo in questa sezione il caso di una variabile continua o trattata come tale; per tale variabile i valori vengono raggruppati in classi - intervalli, $X_1, X_2, \dots, X_k$, con le indicazioni date nel fascicolo 2, ed ad ogni classe viene associata la sua frequenza o intensità assoluta o relativa o percentuale.

classi di valori / frequenze assolute, relative e percentuali

Classi di valori	$f_a(X_i) = n_i$	$f_r(X_i) = n_i/n$	$f\%(X_i) = (n_i/n) \times 100$
X_1	n_1	n_1/n	$(n_1/n) \times 100$
X_2	n_2	n_2/n	$(n_2/n) \times 100$
...	...	...	...
X_k	n_k	n_k/n	$(n_k/n) \times 100$
	n	1	100

n = numero totale delle unità statistiche

classi di valori / quantità assolute, relative e percentuali

Classi di valori	$q_a(X_i) = q_i$	$q_r(X_i) = q_i/Q$	$q\%(X_i) = (q_i/Q) \times 100$
X_1	q_1	q_1/Q	$(q_1/Q) \times 100$
X_2	q_2	q_2/Q	$(q_2/Q) \times 100$
...	...	...	...
X_k	q_k	q_k/Q	$(q_k/Q) \times 100$
	Q	1	100

Q = ammontare totale delle quantità

Una distribuzione di frequenza associa ad ogni classe di valori la sua frequenza

$$f: X_i \in \mathcal{X} \rightarrow f(X_i) \in \mathbb{R} \quad (2)$$

ed è una funzione reale definita nell'insieme $\mathcal{X} = \{X_1, X_2, \dots, X_k\}$ delle classi.

Non essendo la (2) una funzione reale di una variabile reale non possiamo ricorrere al diagramma cartesiano per rappresentarla graficamente; per la sua rappresentazione grafica è il concetto di *densità di frequenza (o intensità)* e di *funzione densità* che illustriamo nel prossimo paragrafo riferendoci alle distribuzioni di frequenza.

2.1 La densità di frequenza (e il caso di classi di valori di diversa ampiezza)

La ripartizione in classi di uguale ampiezza dei valori di una variabile continua o trattata come tale non è sempre praticabile (cfr. esempio 20 del fascicolo 2): infatti, molti fenomeni tendono a diradare la propria frequenza nelle classi estreme e a concentrarsi molto in quelle centrali; adottando classi di uguale ampiezza si corre il rischio di perdita di informazioni nelle classi estreme. Si conviene allora adottare una soluzione di compromesso: scegliere classi di uguale ampiezza laddove le osservazioni sono più addensate e classi di ampiezza maggiore laddove le osservazioni sono più diradate. La scelta di classi di diversa ampiezza comporta però perdita di significatività nel confronto tra le frequenze delle varie classi: l'informazione fornita dalla frequenza sull'addensarsi dei valori in una classe va integrata con l'informazione sull'ampiezza della classe stessa. Per eliminare l'effetto della dimensione della classe si calcola la *densità assoluta* di frequenza qui di seguito definita

Densità di frequenza assoluta riferita a una classe di valori limitata è il **rapporto tra la frequenza assoluta corrispondente alla classe e l'ampiezza della classe**

Esempio 5

Classi di valori	Ampiezza classe	f_a	$\delta_a = \text{dens. di freq. ass.}$
1 -- 5	4	8	8/4=2
5-- 15	10	10	10/10=1
15-- 30	15	6	6/15=0,4
30-- 50	20	4	4/20=0,2
50 -- 80	30	7	7/30=0,233
		35	

La classe 5--|15 è
← ampia ma non
densamente
popolata

Leggendo le frequenze assolute abbiamo la sensazione che la prima classe sia meno popolata della seconda, ma *le densità di frequenza ci dicono che la prima classe è più densamente popolata della seconda* in quanto, pensando a valori uniformemente distribuiti su sotto intervalli di ampiezza 1, la prima classe ha 2 unità per intervallino unitario mentre la seconda ha una sola unità.

Ovviamente *nel caso di ampiezze uguali le densità sono proporzionali alle frequenze assolute e le informazioni sul confronto tra le classi, relativamente a quanto esse siano “popolate”, ottenute a partire dalle frequenze, sono le stesse di quelle ottenute a partire dalle densità.*

Esempio 6

Classi di valori	f_a	$\delta = \text{dens. di freq.}$
10— 20	7	$7/10=0,7$
20— 30	4	$4/10=0,4$
30— 40	5	$5/10=0,5$
40— 50	2	$2/10=0,2$
50— 60	2	$2/10=0,2$
	20	

Ogni densità è un decimo della frequenza assoluta corrispondente

$$(7/10, 4/10, 5/10, 2/10, 2/10) = \frac{1}{10}(7, 4, 5, 2, 2)$$

Nel caso considerato il rapporto tra frequenze di classi distinte è uguale al rapporto delle corrispondenti densità e dà le stesse informazioni sulla distribuzione delle unità statistiche sulle varie classi di valori

Accanto al concetto di densità di frequenza assoluta sono dati anche i concetti di densità di frequenza relativa e percentuale

Densità di frequenza relativa riferita a una classe di valori(limitata): *rapporto tra la frequenza relativa corrispondente alla classe e l'ampiezza della classe.*

Densità di frequenza percentuale riferita ad una classe di valori(limitata): *rapporto tra la frequenza relativa corrispondente alla classe e l'ampiezza della classe.*

Esempio 7

Classi di valori	<i>Ampiezza classe</i>	f_a	f_r	$\delta_a = \text{dens. di freq. ass.}$	$\delta_r = \text{dens. di freq. rel.}$
1 -- 5	4	8	8/35	8/4=2	8/(35×4)=2/35
5-- 15	10	10	10/35	10/10=1	10/(35×10)=1/35
15-- 30	15	6	6/35	6/15=0,4	6/(35×15)=0,4/35=4/350
30-- 50	20	4	4/35	4/20=0,2	4/(35×20)=0,2/35=2/350
50 -- 80	30	7	7/35	7/30=0,233	7/(35×30)=0,233/35=233/35000
		35			

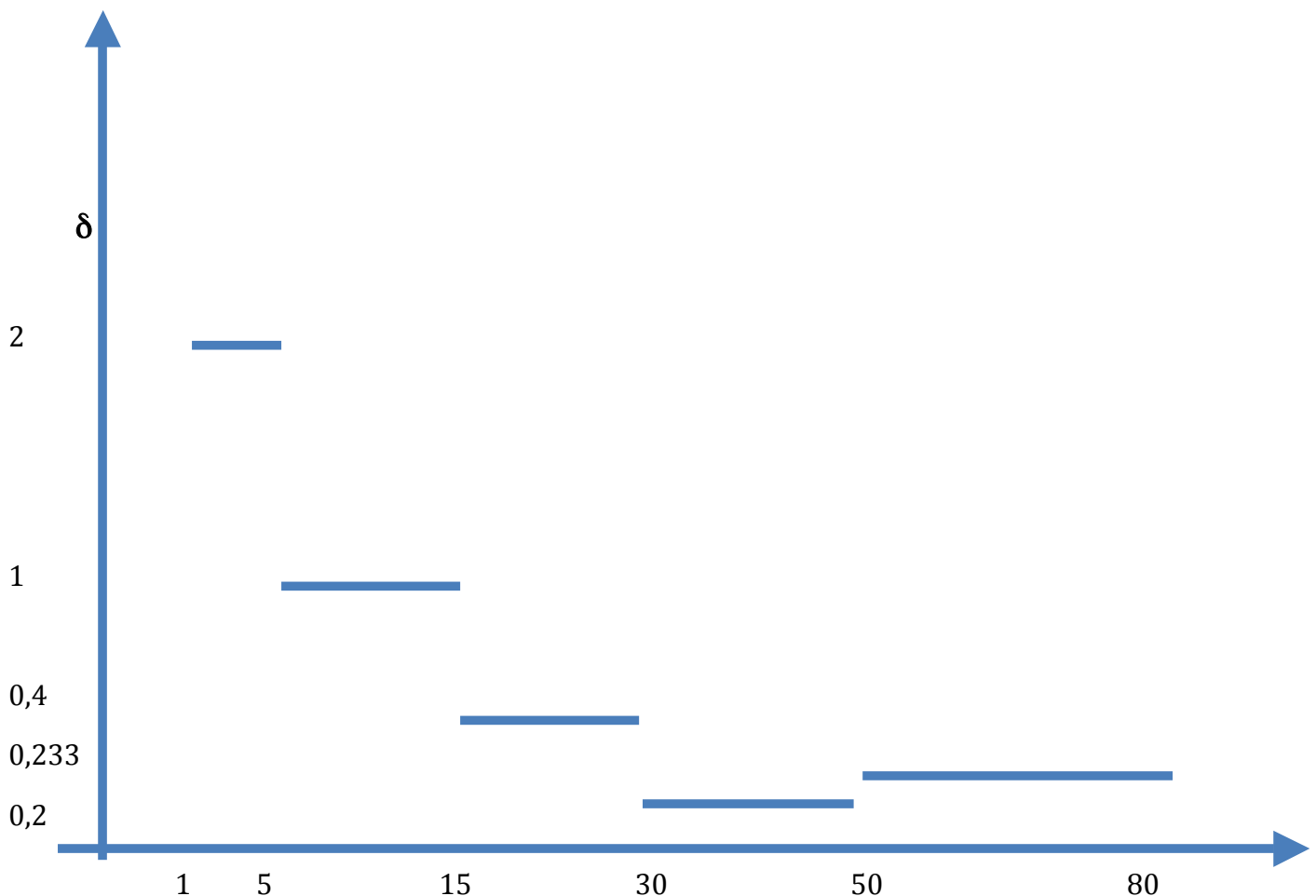
Esercizio: calcolare le densità di frequenze percentuali

2.2 Funzione densità

Nel caso di un carattere con valori raggruppati in classi, è detta **funzione densità assoluta (relativa o percentuale)** la funzione che a ogni valore del carattere associa la densità di frequenza assoluta (relativa o percentuale) corrispondente alla classe cui appartiene il valore. La funzione densità è una funzione costante a tratti e il suo grafico è un grafico a scalino

Esempio 8. La funzione densità assoluta corrispondente alla distribuzione di frequenze assolute di cui all'esempio 7 è la seguente:

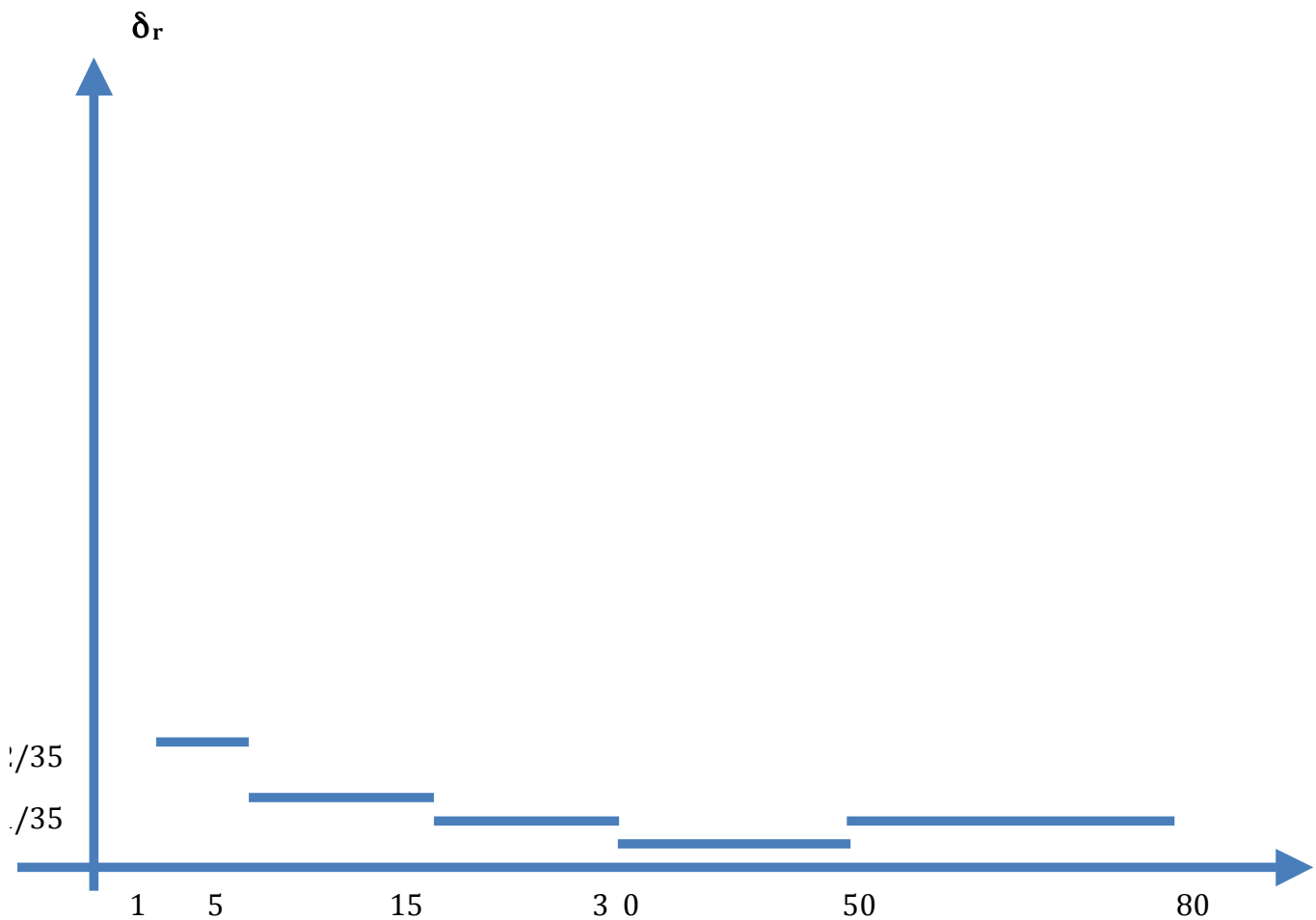
$$\delta: x \in [1, 80] \rightarrow \delta(x) = \begin{cases} 2 & \text{se } x \in [1, 5] \\ 1 & \text{se } x \in]5, 15] \\ 0,4 & \text{se } x \in]15, 30] \\ 0,2 & \text{se } x \in]30, 50] \\ 0,233 & \text{se } x \in]50, 80] \end{cases}$$



La funzione densità relativa corrispondente alla distribuzione di frequenze relative di cui all'esempio 10, distribuzione corrispondente a quella delle frequenze assolute di cui all'esempio 8, è la seguente:

$$\delta_r : x \in [1, 80] \rightarrow \delta(x) = \begin{cases} 2/35 & \text{se } x \in [1, 5] \\ 1/35 & \text{se } x \in]5, 15] \\ 4/350 & \text{se } x \in]15, 30] \\ 2/350 & \text{se } x \in]30, 50] \\ 233/35000 & \text{se } x \in]50, 80] \end{cases}$$

e il grafico della funzione è a scalino come il grafico della funzione densità assoluta della funzione

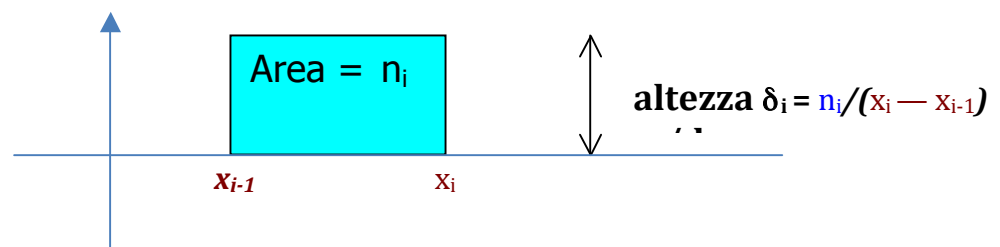


Esercizio: determinare la funzione di densità percentuale

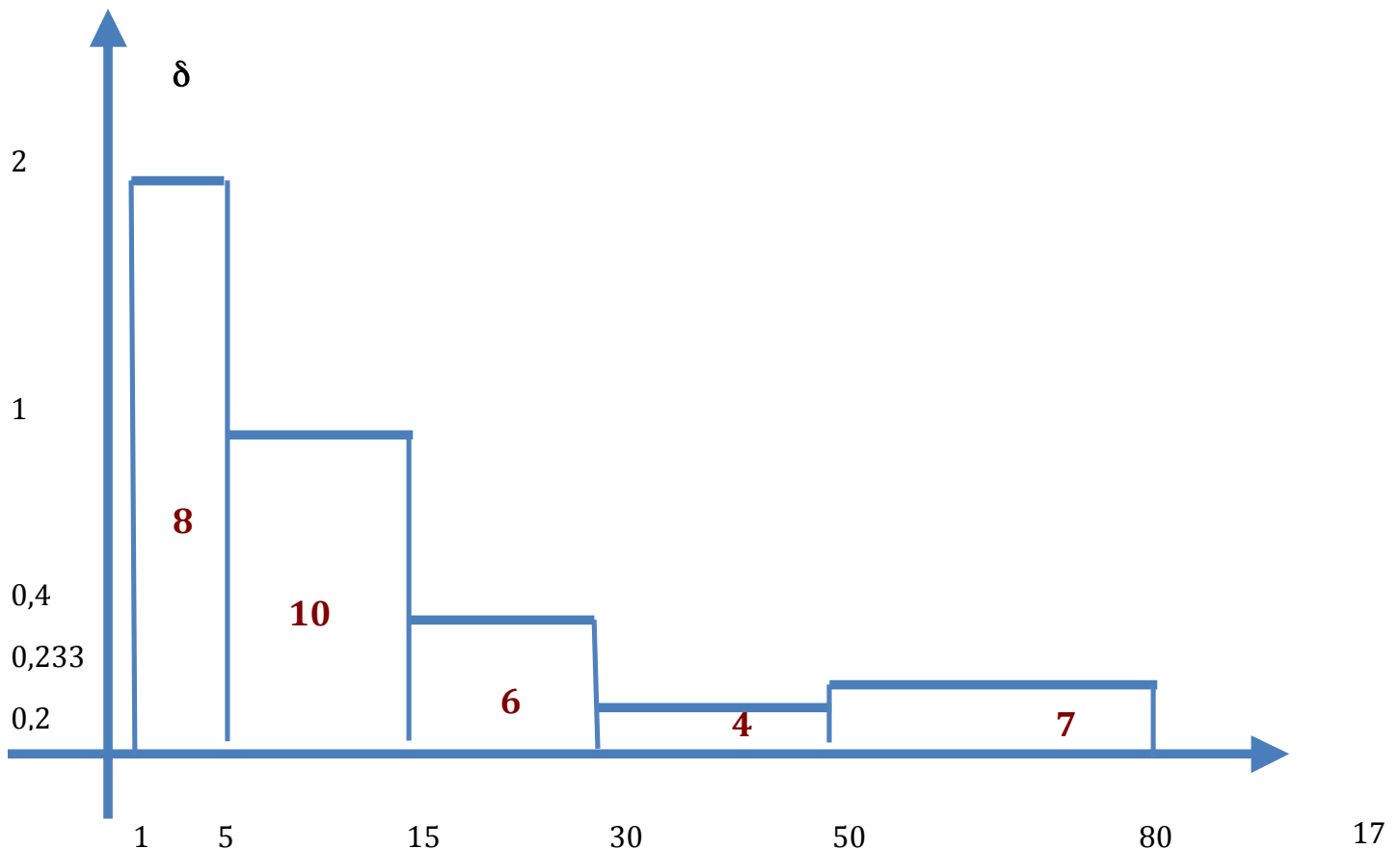
2.3 . Rappresentazione grafiche delle distribuzioni di frequenza per caratteri quantitativi con valori raggruppati in classi: l'istogramma e il poligono delle frequenze

Dalla funzione densità all'istogramma delle frequenze Per giungere alla rappresentazione grafica di una distribuzione di frequenze assolute, (risp. relative o percentuali) utilizziamo la funzione densità assoluta (risp. relativa o percentuale) e il suo **rettangoloide**, cioè la zona di piano compresa tra il grafico della funzione e l'asse delle ascisse. Il rettangoloide è unione di rettangoli contigui ognuno dei quali ha per base un intervallo classe, $x_{i-1} - x_i$, e altezza pari alla densità spettante alla classe quindi la sua area corrisponde alla frequenza associata alla classe; infatti con riferimento alle frequenze assolute si ha:

$$\text{area rettangolo} = \text{base} \times \text{altezza} = \text{ampiezza classe} \times \text{densità} = (x_i - x_{i-1}) \times n_i / (x_i - x_{i-1}) = n_i.$$



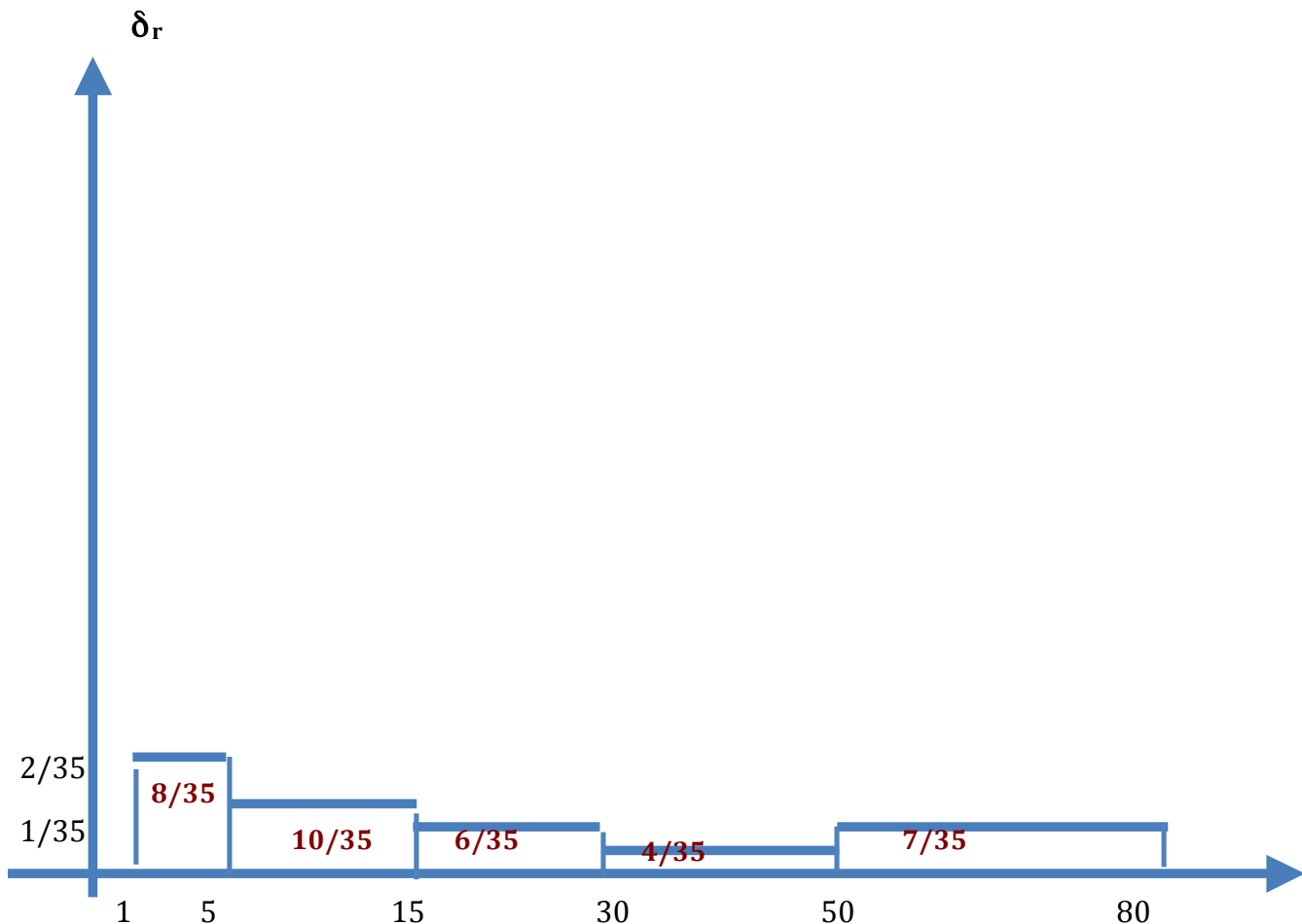
Consideriamo la distribuzione di frequenze assolute di cui all'esempio 7 e la funzione di densità assoluta costruita nell'esempio 8.. Il rettangoloide della funzione densità assoluta, di seguito disegnato, è unione di rettangoli e ogni rettangolo ha area pari alla frequenza corrispondente all'intervallo classe su cui è costruito



Consideriamo la funzione di densità relativa

$$\delta_r : x \in [1, 80] \rightarrow \delta(x) = \begin{cases} 2/35 & \text{se } x \in [1, 5] \\ 1/35 & \text{se } x \in]5, 15] \\ 4/350 & \text{se } x \in]15, 30] \\ 2/350 & \text{se } x \in]30, 50] \\ 233/35000 & \text{se } x \in]50, 80] \end{cases}$$

Il rettangoloide della funzione densità relativa, di seguito disegnato, è unione di rettangoli contigui ognuno dei quali ha area pari alla frequenza relativa della classe che è indicata alla base del rettangolo



Concludendo:

- *le frequenze assolute associate alle varie classi di valori sono rappresentate dalle aree dei rettangoli costituenti il rettangoloide della funzione densità assoluta: la frequenza assoluta di una classe è data dall'area del rettangolo che ha per base la classe;*
- *le frequenze relative associate alle varie classi di valori sono rappresentate dalle aree dei rettangoli costituenti il rettangoloide della funzione densità relativa: la frequenza relativa di una classe è data dall'area del rettangolo che ha per base la classe;*
- *le frequenze percentuali associate alle varie classi di valori sono rappresentate dalle aree dei rettangoli costituenti il rettangoloide della funzione densità percentuale: la frequenza percentuale di una classe è data dall'area del rettangolo che ha per base la classe.*

Il rettangoloide della funzione densità assoluta, unione di rettangoli adiacenti, è detto istogramma delle frequenze assolute: esso è scelto quale **rappresentazione grafica** della distribuzione delle frequenze assolute *perché permette di leggere la distribuzione di frequenze attraverso le aree dei rettangoli che la costituiscono*; l'area dell'istogramma, pari alla somma delle aree di tutti i rettangoli, è uguale alla frequenza totale.

Il rettangoloide della funzione di densità relativa, detto istogramma delle frequenze relative: *l'area di ogni rettangolo è pari alla frequenza relativa della classe base e l'area dell'istogramma, pari alla somma delle aree di tutti i rettangoli, è uguale a 1.*

Il rettangoloide della funzione di densità percentuale detto istogramma delle frequenze percentuali: *l'area di ogni rettangolo è pari alla frequenza percentuale della classe base e l'area dell'istogramma, pari alla somma delle aree di tutti i rettangoli, è uguale a 100.*

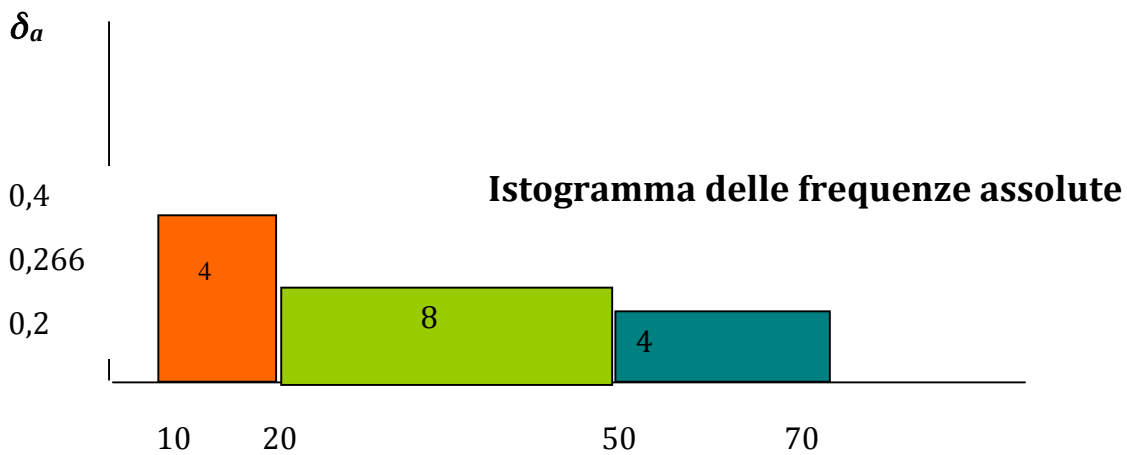
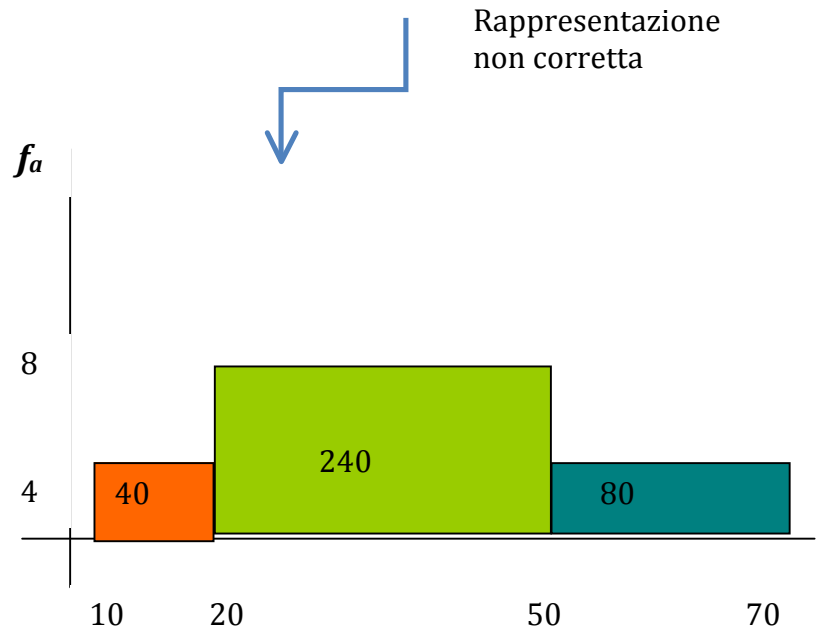
Di seguito alcuni esempi di rappresentazione di distribuzione delle frequenze su classi di valori attraverso l'istogramma.

Esempio 9

Classi di valori	f_a	$\delta_a = \text{dens. di freq. ass.}$
10 — 20	4	$4/10 = 0,4$
20 — 50	8	$8/30 = 0,2666$
50 — 70	4	$4/20 = 0,2$

16

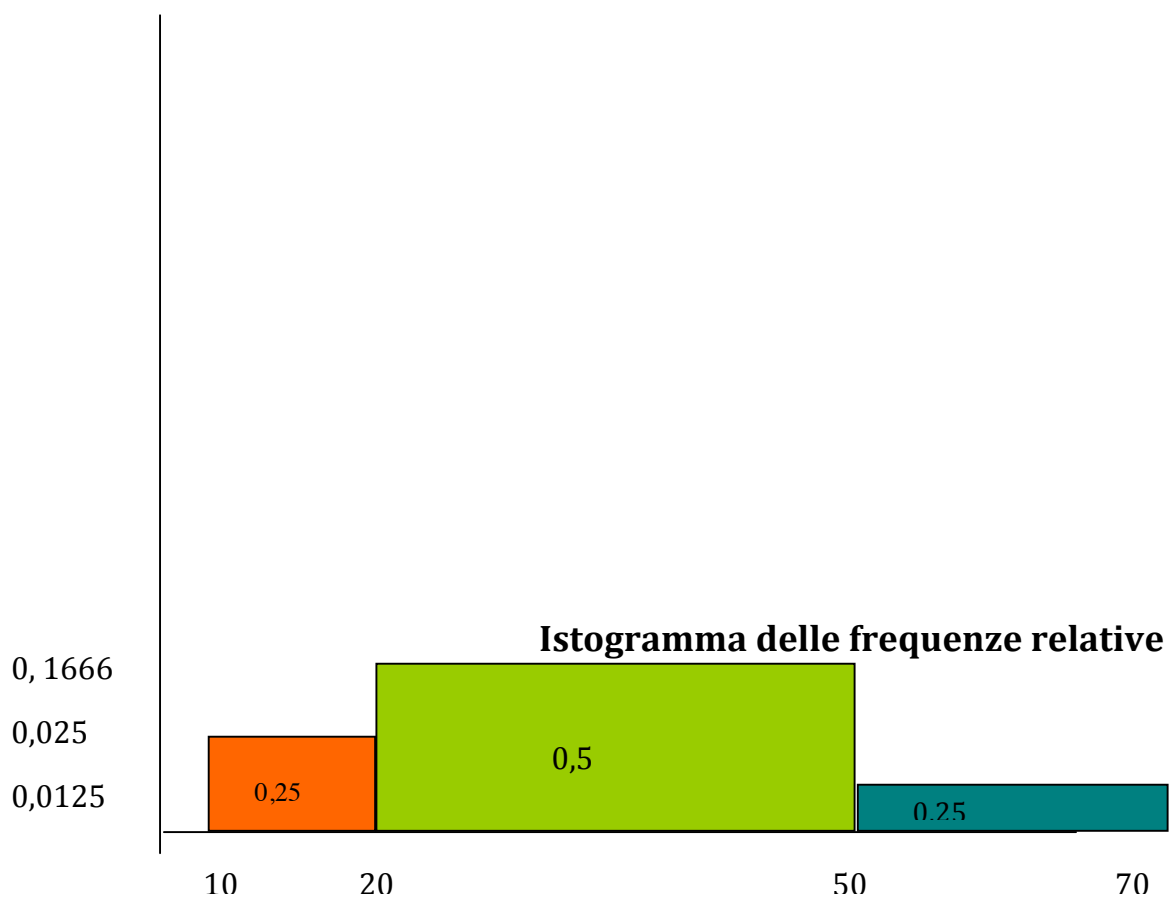
Rappresentazione
corretta



Se sull'asse delle ordinate portiamo i valori di densità relativa, ogni rettangolo avrà area uguale alla frequenza relativa corrispondente alla classe rappresentata dall'intervallo alla base.

Classi di valori	f_a	δ_a =dens. di freq. ass.	f_r	δ_r =dens. di freq. rel
10 — 20	4	$4/10=0,4$	$4/16= 1/4 = 0,25$	$4/160=0,025$
20 — 50	8	$8/30$	$8/16=1/2 = 0,5$	$1/60= 0,0166666..$
50 — 70	4	$4/20$	$4/16= 1/4 = 0,25$	$1/80= 0,0125$

16

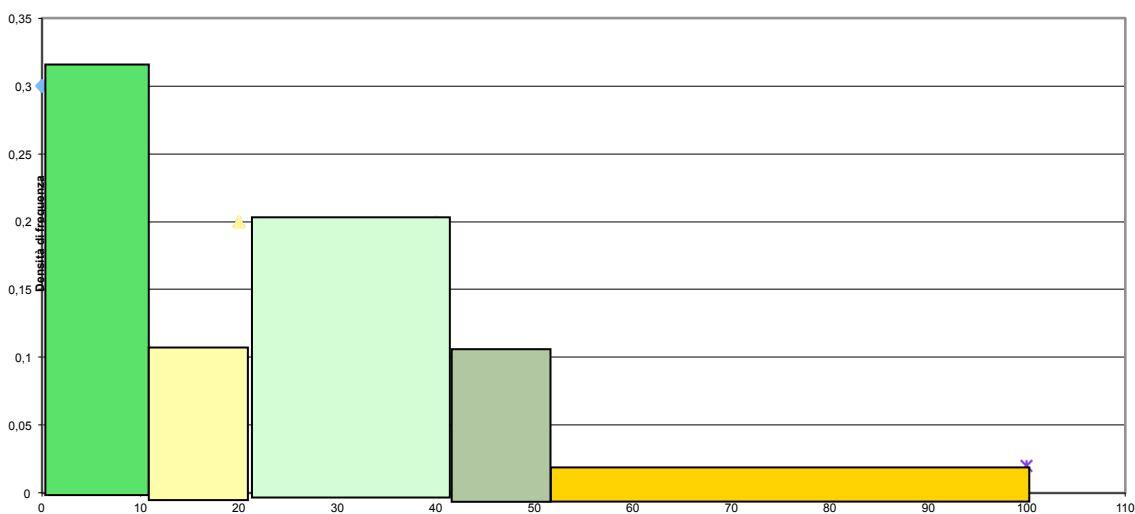


Esempio 10

Classi di Reddito	$n_i = f_a(M_i)$	$n_i/n = p_i = f_r(M_i)$	$p_i * 100 = f_{per}(M_i)$
$M_1 = (0,10]$	3	0,3	30
$M_2 = (10,20]$	1	0,1	10
$M_3 = (20,40]$	4	0,4	40
$M_4 = (40,50]$	1	0,1	10
$M_5 = ((50,100]$	1	0,1	10
TOT	10	1	100

Per costruire l'istogramma delle frequenze assolute, occorre determinare le densità di frequenza assoluta. L'istogramma si ottiene riportando sull'asse delle ascisse le classi contigue di valori e sull'asse delle ordinate le densità di frequenza assoluta

Reddito	n_i	d_i	$\delta_i = n_i/d_i$
(0,10]	3	10	0.3
(10,20]	1	10	0.1
(20,40]	4	20	0.2
(40,50]	1	10	0.1
(50,100]	1	50	0.02

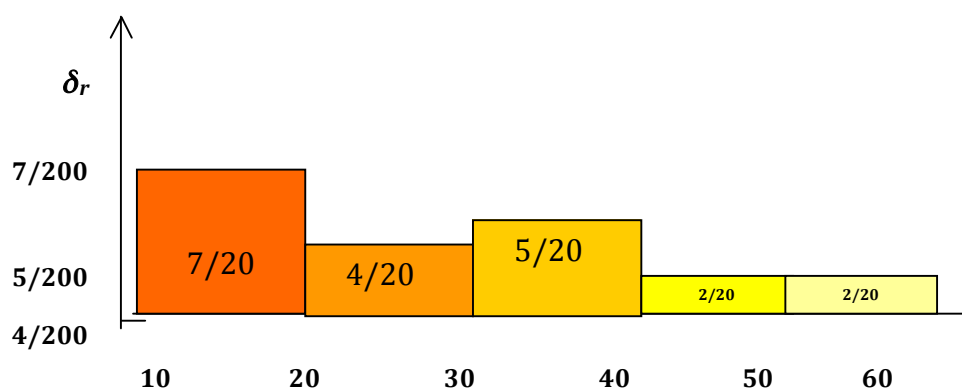


da

Excell

Esempio 11.

Classi di valori	f_a	f_r	$\delta_a = \text{dens. di frequenza ass.}$	$\delta_r = \text{dens. di frequenza rel.}$
10 — 20	7	$7/20=0,35$	$7/10=0,7$	$7/200=0,035$
20 — 30	4	$4/20=0,2$	$4/10=0,4$	$4/200=0,02$
30 — 40	5	$5/20=0,25$	$5/10=0,5$	$5/200=0,025$
40 — 50	2	$2/20=0,1$	$2/10=0,2$	$2/200=0,01$
50 — 60	2	$2/20=0,1$	$2/10=0,2$	$2/200=0,01$
	20	1		



Cosa rappresenta l'istogramma in figura?

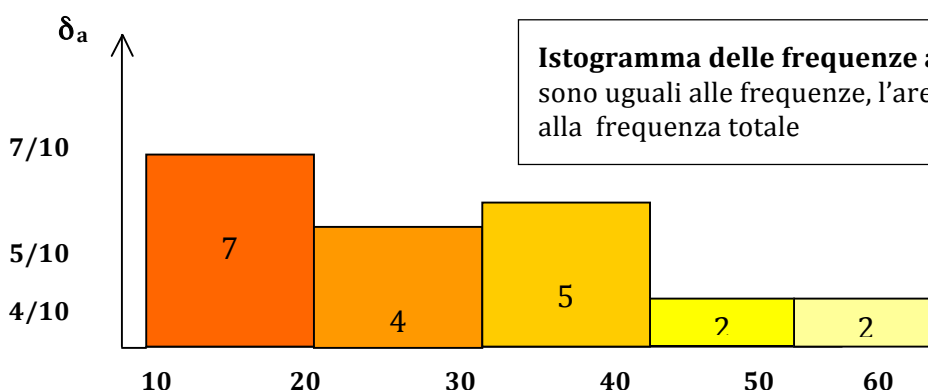
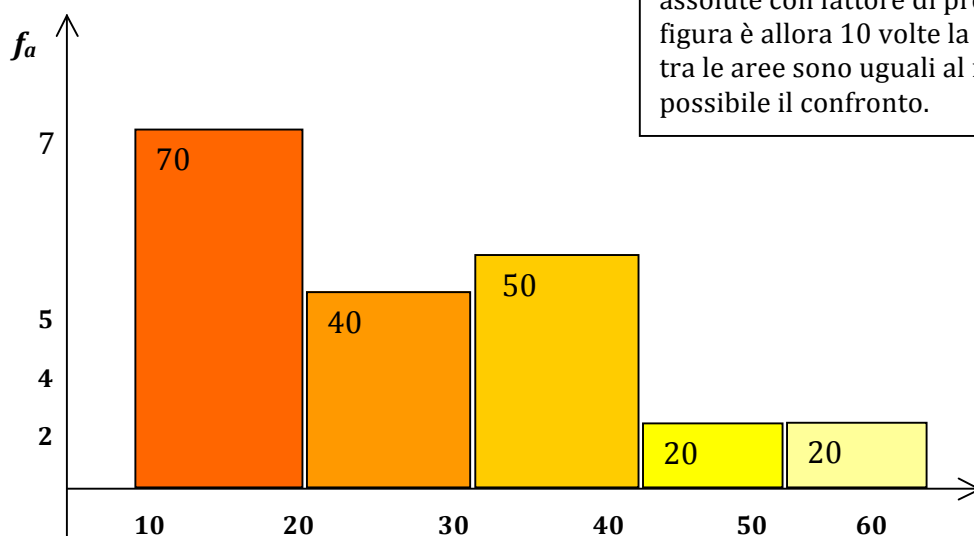
NOTA sul caso di un carattere continuo con valori raggruppati in classi di uguale ampiezza

Riportando le frequenze assolute sull'asse delle ordinate e associando ogni frequenza ai punti della classe a cui la frequenza corrisponde, si ottengono rettangoli le cui aree sono proporzionali alle frequenze con fattore di proporzionalità dato dall'ampiezza comune delle classi: i rapporti tra le aree sono uguali al rapporto tra le frequenze; allora tenendo conto del fattore di proporzionalità si può leggere dalla figura la distribuzione delle frequenze classe per classe

Esempio 12

Classi di valori	f_a	δ
10— 20	7	7/10
20— 30	4	4/10
30— 40	5	5/10
40— 50	2	2/10
50— 60	2	2/10
	20	20/10

Le aree dei rettangoli sono proporzionali alle frequenze assolute con fattore di proporzionalità 10; l'area della figura è allora 10 volte la frequenza totale. I rapporti tra le aree sono uguali al rapporto tra le frequenze: è possibile il confronto.



Istogramma delle frequenze assolute: le aree dei rettangoli sono uguali alle frequenze, l'area dell'istogramma è uguale alla frequenza totale

2. 4. SULLA COSTRUZIONE DELL'ISTORIGRAMMA QUANDO LE CLASSI SONO DATE ATTRAVERSO LIMITI FORMALI

1. nel caso in cui le classi di valori sono riportate attraverso limiti formali, l'istogramma va costruito facendo riferimento ai limiti reali;
2. nel caso in cui ci sono classi aperte, per costruire l'istogramma è necessario assegnare un limite reale inferiore alla classe aperta non limitata inferiormente ed un limite reale inferiore alla classe aperta non limitata superiormente; tali limiti sono scelti a discrezione del valutatore.

Esempio 13

Consideriamo la seguente tabella in cui le classi sono date attraverso limiti formali

<i>altezza ↓</i>	<i>Freq. ass</i>	<i>Freq.rel</i>	<i>Freq. Perc.</i>
140 - 149	2		
150 - 159	6		
160 - 169	20		
170 - 179	10		
180 - su	12		
Totale	50	1	100

Per una giusta rappresentazione della distribuzione delle frequenze e per una corretta rappresentare grafica della stessa attraverso un istogramma, occorre che le classi sia contigue; calcoliamo allora i limiti reali

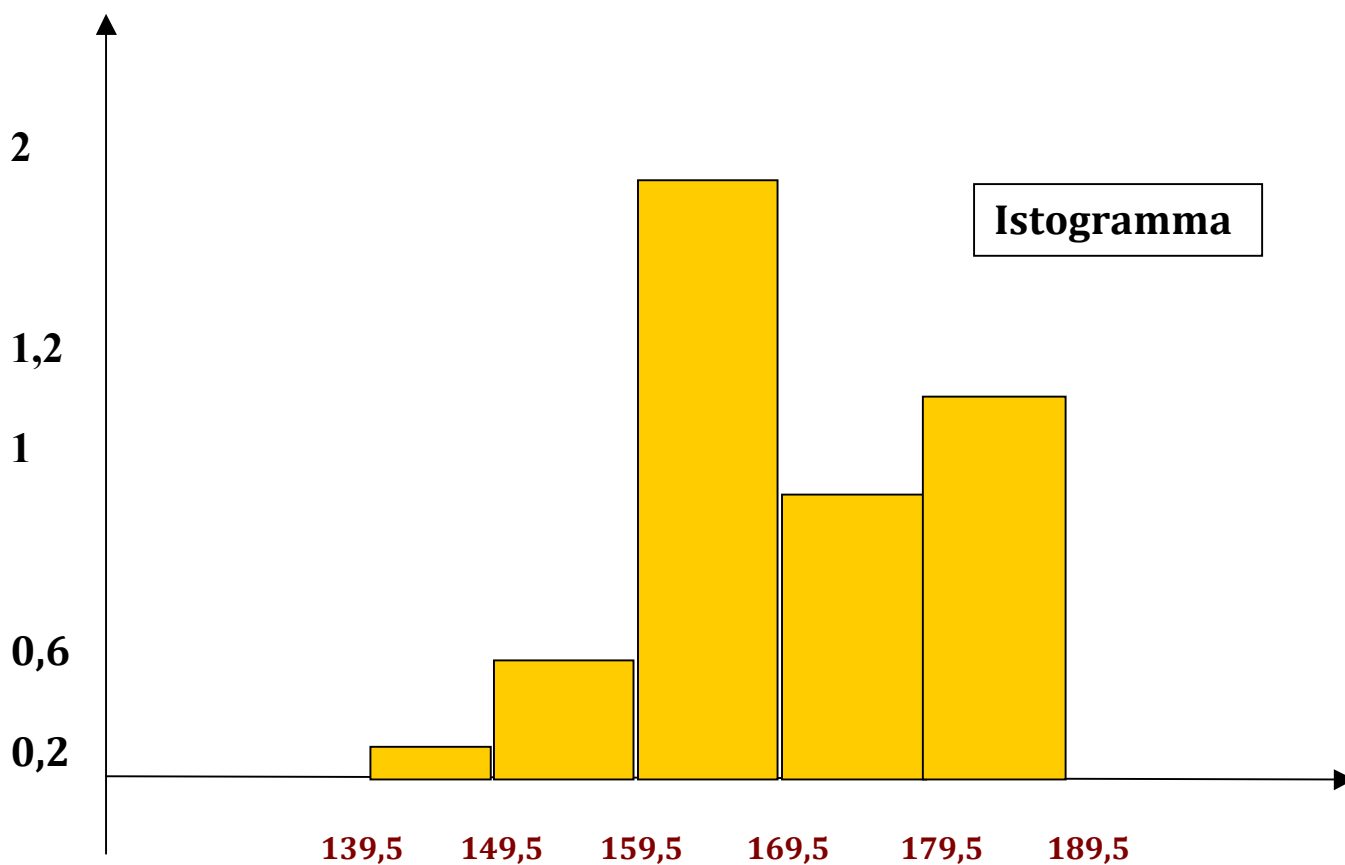
139,5←140 _____ 149→**149,5**←150 _____ 159→**159,5** _____ **169,5** ←170 _____ 179→**179,5**←180 _____ **189,5**

Utilizzando i limiti reali e supponendo che non siano riscontrate altezze superiori a 189,5 in modo da poter assegnare un limite superiore anche all'ultima classe (cioè che ci aiuta nella rappresentazione grafica), la tabella va così cambiata

Distribuzione semplice di frequenze secondo l'altezza di 50 individui - seriazioni di frequenze

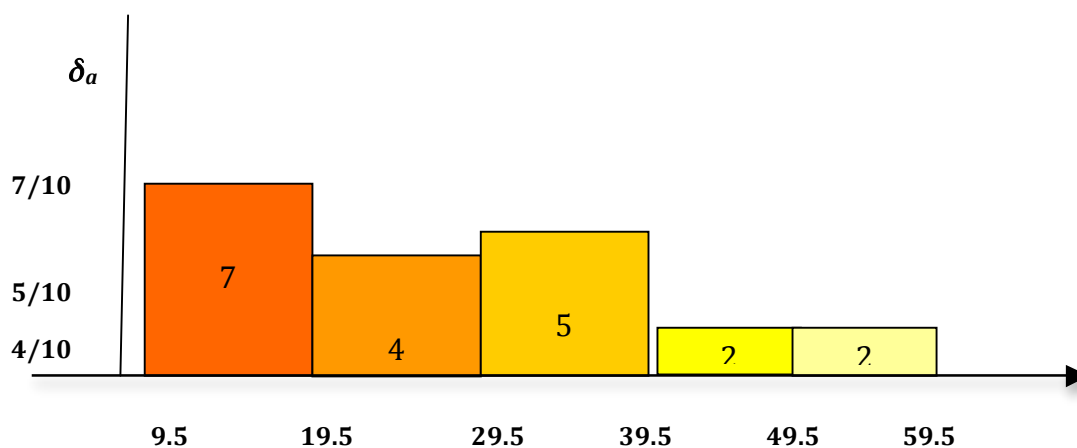
<i>altezza</i> ↓	<i>Freq. ass</i>	δ_a = dens. di frequenza ass.	<i>Freq.rel</i>	<i>Freq. Perc.</i>
139,5 – 149,5	2	2/10 = 0,2		
149,5 – 159,5	6	6/10 = 0,6		
159,5 – 169,5	20	20/10 = 2		
169,5 – 179,5	10	10/10 = 1		
179,5 – 189,5	12	12/10 = 1,2		
Totale	50		1	100

La distribuzione di frequenze assolute è rappresentata dall'istogramma



Esempio 14 (relativo al caso in cui le classi di valori siano date con limiti formali)

Classi di valori	f_a		
10— 19	7		
20— 29	4		
30— 39	5		
40— 49	2		
50— 59	2		
	20		
Limiti reali	f_a	$\delta = \text{dens. di freq. ass}$	
9,5--- 19,5	7	7/10=0,7	
19,5---29,5	4	4/10=0,4	
29,5---39,5	5	5/10=0,5	
39,5---49,5	2	2/10=0,2	
49,5--- 59,5	2	2/20=0,2	



2.4.1. Poligono delle frequenze.

Il poligono delle frequenze assolute (relative o percentuali) è una spezzata che sostituisce il grafico della funzione densità di frequenza assoluta (relativa o percentuale)

Il caso di classi di uguale ampiezza. Il poligono delle frequenze è la spezzata che si ottiene unendo i punti medi delle basi superiori dei rettangoli costituenti l'istogramma, unendo inoltre il punto medio della base superiore del primo rettangolo con un punto dell'asse delle ascisse che precede il rettangolo (di metà ampiezza della base) e il punto medio della base superiore dell'ultimo rettangolo con un punto dell'asse delle ascisse che segue il rettangolo (di metà ampiezza della base): in tal modo la linea spezzata è unita all'asse orizzontale sia all'inizio che alla fine e l'area sotto la spezzata coincide con l'area dell'istogramma.

Classi di valori	c_i valore centrale	f_a	dens. di freq.
10 — 20	15	7	$7/10=0,7$
20 — 30	25	4	$4/10=0,4$
30 — 40	35	5	$5/10=0,5$
40 — 50	45	2	$2/10=0,2$
50 — 60	55	2	$2/10=0,2$
		20	

$P_0(5, 0)$

$P_1(15, 0,7)$

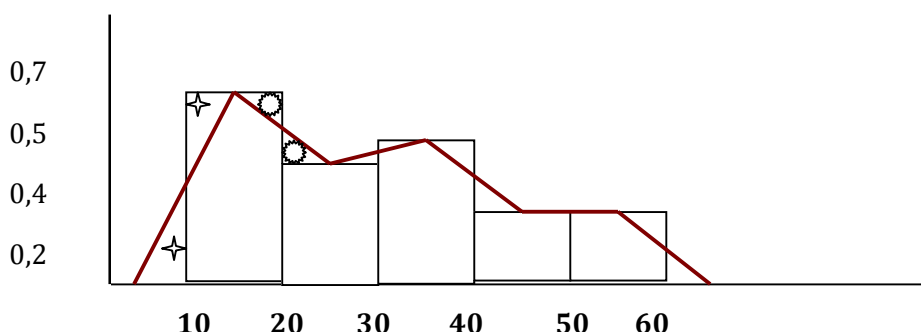
$P_2(25, 0,4)$

$P_3(35, 0,5)$

$P_4(45, 0,2)$

$P_5(55, 0,2)$

$P_6(65, 0)$



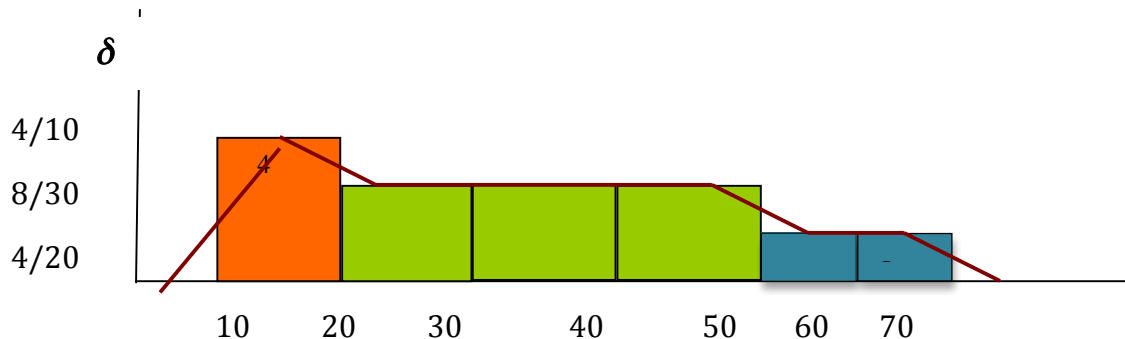
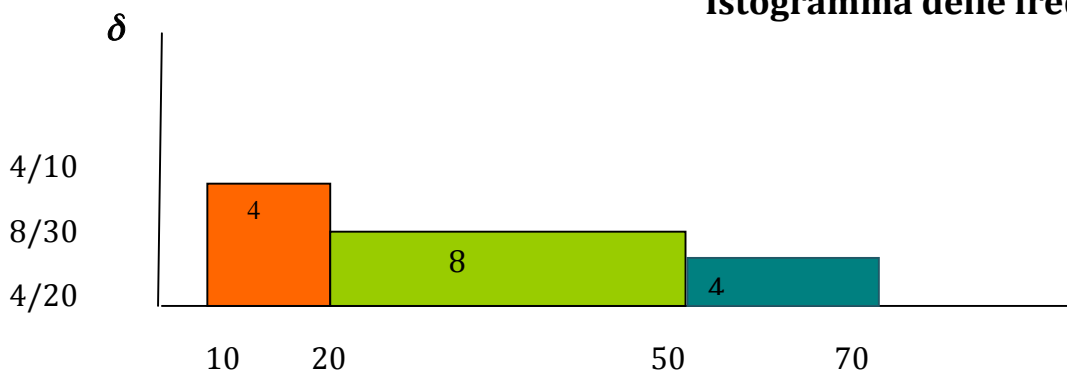
La densità di frequenza δ_i corrispondente alla classe M_i è allora associata al valore centrale x_i della classe e il poligono di frequenza è la spezzata che unisce i punti consecutivi $P_i(x_i, \delta_i)$ e $P_{i+1}(x_{i+1}, \delta_{i+1})$ e il primo e l'ultimo di questi punti con i due punti dell'asse delle ascisse esterni all'istogramma, individuati come detto sopra. Il poligono delle frequenze è detto **poligono delle frequenze assolute** se ottenuto dall'istogramma delle frequenze assolute, è detto **poligono di frequenze relative (risp. poligono di frequenze percentuali)** se ottenuto dall'istogramma delle frequenze relative (risp. percentuali)

Il caso di classi di diversa ampiezza. In tal caso si determina un sottomultiplo comune dei rettangoli e i rettangoli si scompongono nei sottomultipli; il poligono delle frequenze è allora la spezzata che si ottiene unendo i punti medi delle basi superiori dei nuovi rettangoli costituenti l'istogramma, unendo inoltre il punto medio della base superiore del primo rettangolo base con un punto dell'asse delle ascisse che precede il rettangolo (di metà ampiezza della base) e il punto medio della base superiore dell'ultimo rettangolo con un punto dell'asse delle ascisse che segue il rettangolo (di metà ampiezza della base): l'area sotto la spezzata che si viene a determinare coincide con l'area dell'istogramma.

Classi di valori	f_a	$\delta_a = \text{dens}$
10 — 20	4	4/10
20 — 50	8	8/30
50 — 70	4	4/20

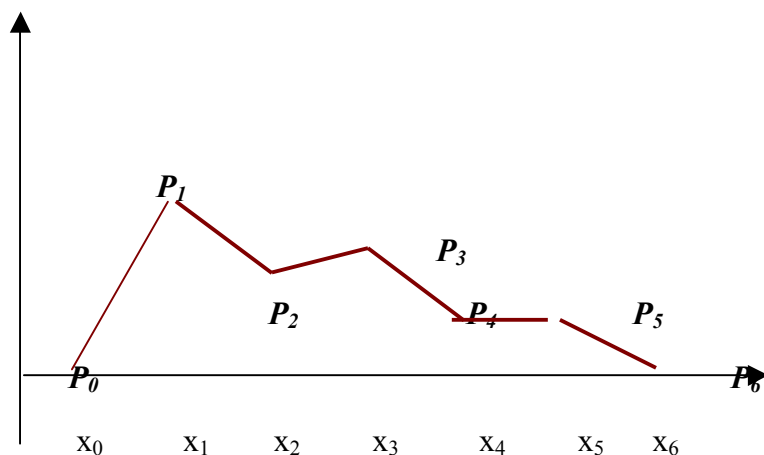
16

Istogramma delle frequenze assolute



Dal poligono delle frequenze alla curva di frequenza

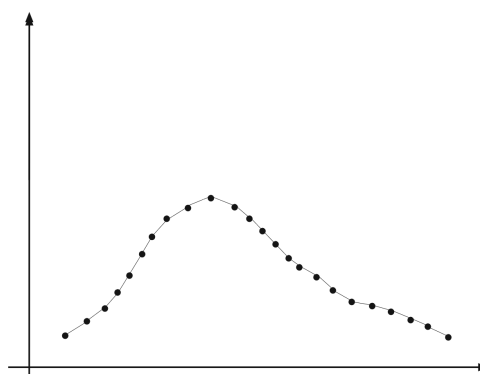
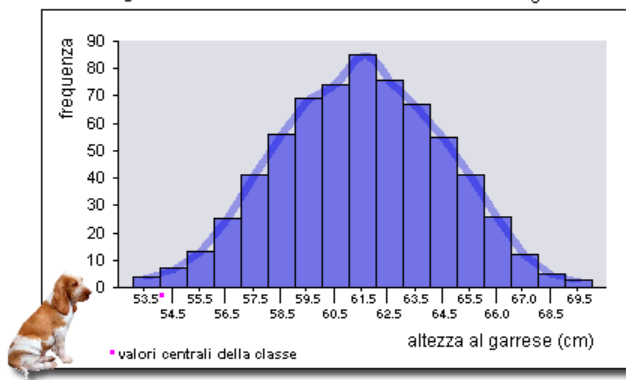
I caratteri quantitativi continui possono teoricamente assumere tutti i valori compresi in un intervallo della retta reale; e l'ideale sarebbe rappresentare una distribuzione di frequenze attraverso un grafico continuo e il corrispondente rettangoloide, in modo tale che *l'area sottostante il rettangoloide e compresa tra i valori x_i e x_j presi dal carattere ci dia la frequenza corrispondente a tutta la classe di valori compresa tra x_i e x_j* , cosa che approssimativamente si tenta di fare con il poligono delle frequenze costituito da una linea spezzata ma continua.



Si può giungere ad una rappresentazione ideale della distribuzione di frequenze con un passaggio al limite, scegliendo un numero sempre più elevato di classi in cui raggruppare i valori e quindi facendo tendere a 0 l'ampiezza delle classi. Più sono piccole le classi di valori su cui si costruisce un istogramma, più i rettangoli sono stretti (e assimilabili a segmenti), più numerosi e vicini sono i punti $P_i(x_i, d_i)$ uniti dai lati del poligono delle frequenze e tanto più il poligono delle frequenze tende ad approssimarsi ad una linea continua non spezzata. La linea teorica a cui tende il poligono delle frequenze, quando si fanno tendere a 0 le dimensioni delle classi, è detta **curva di frequenza** ed è talvolta esprimibile mediante una legge matematica. L'area sotto la curva deve rappresentare la frequenza totale e la porzione di rettangoloide corrispondente ad un intervallo base $[a, b]$ deve misurare in buona approssimazione la frequenza dei valori che cadono tra a e b .

curve di (densità di) frequenza

Altezza al garrese di 659 cani di razza "Bracco italiano". Istogramma.



L'utilità dei poligoni di frequenza o delle curve di frequenza consiste tra l'altro nel fatto che attraverso essi è più semplice confrontare distribuzioni diverse sullo stesso intervallo di valori.