

Statistica: principi e metodi



Capitolo 3

Rappresentazioni grafiche

Rappresentazioni grafiche

Le rappresentazioni grafiche hanno lo scopo di **illustrare** le distribuzioni di frequenze o di quantità.

Presentano diversi vantaggi:

- ▣ Consentono di **visualizzare** immediatamente le caratteristiche delle distribuzioni
- ▣ Rendono possibile il **confronto** tra più distribuzioni in spazi ristretti
- ▣ Agevolano l'investigazione dei fenomeni, **mettendo in rilievo** dati anomali, andamenti, relazioni
- ▣ Sono un **efficace strumento** per la divulgazione dei dati.

Rappresentazioni grafiche

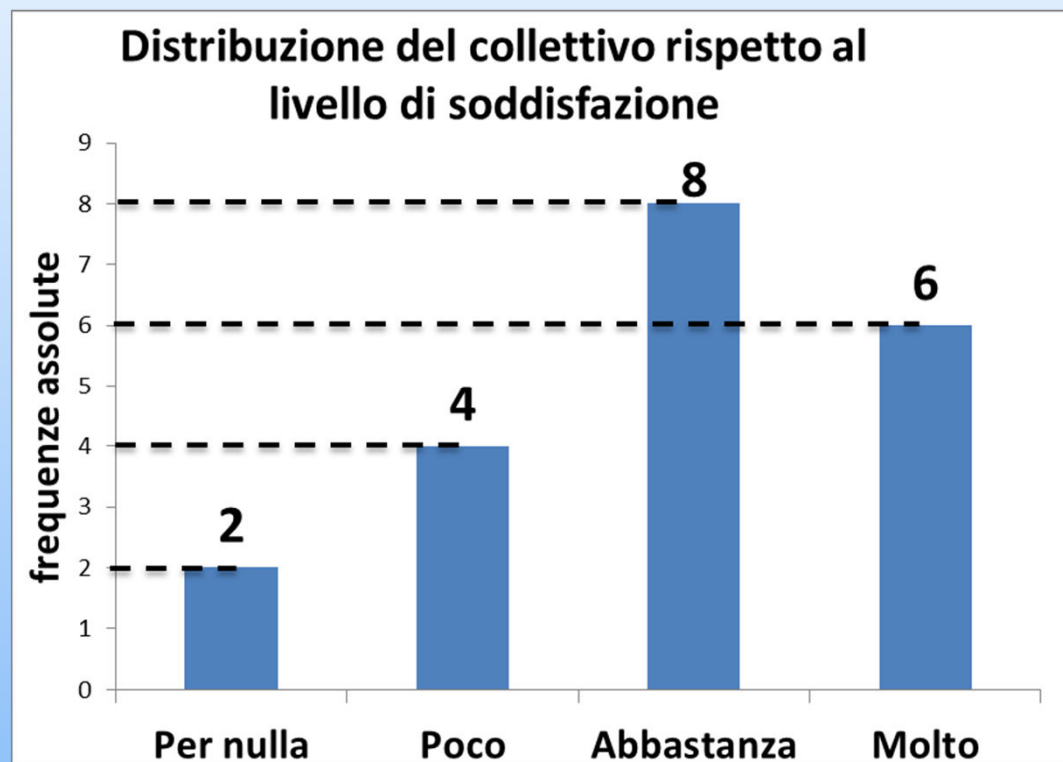
IMPORTANTE

Per ogni distribuzione statistica, è possibile individuare la rappresentazione grafica più adatta, la soluzione, cioè, che concilia **correttezza metodologica** ed **efficacia informativa**

Serie sconnesse

Le serie sconnesse vengono generalmente rappresentate con **grafici di tipo areale**, in cui alle modalità del carattere si fanno corrispondere figure geometriche (rettangoli, quadrati, settori circolari ecc.) con **aree proporzionali** alle grandezze da rappresentare (possono essere frequenze o quantità). Le figure geometriche più spesso utilizzate sono i rettangoli (**nastri**).

Esempio Grafico a nastri verticali



Livello di soddisfazione sul corso di laurea

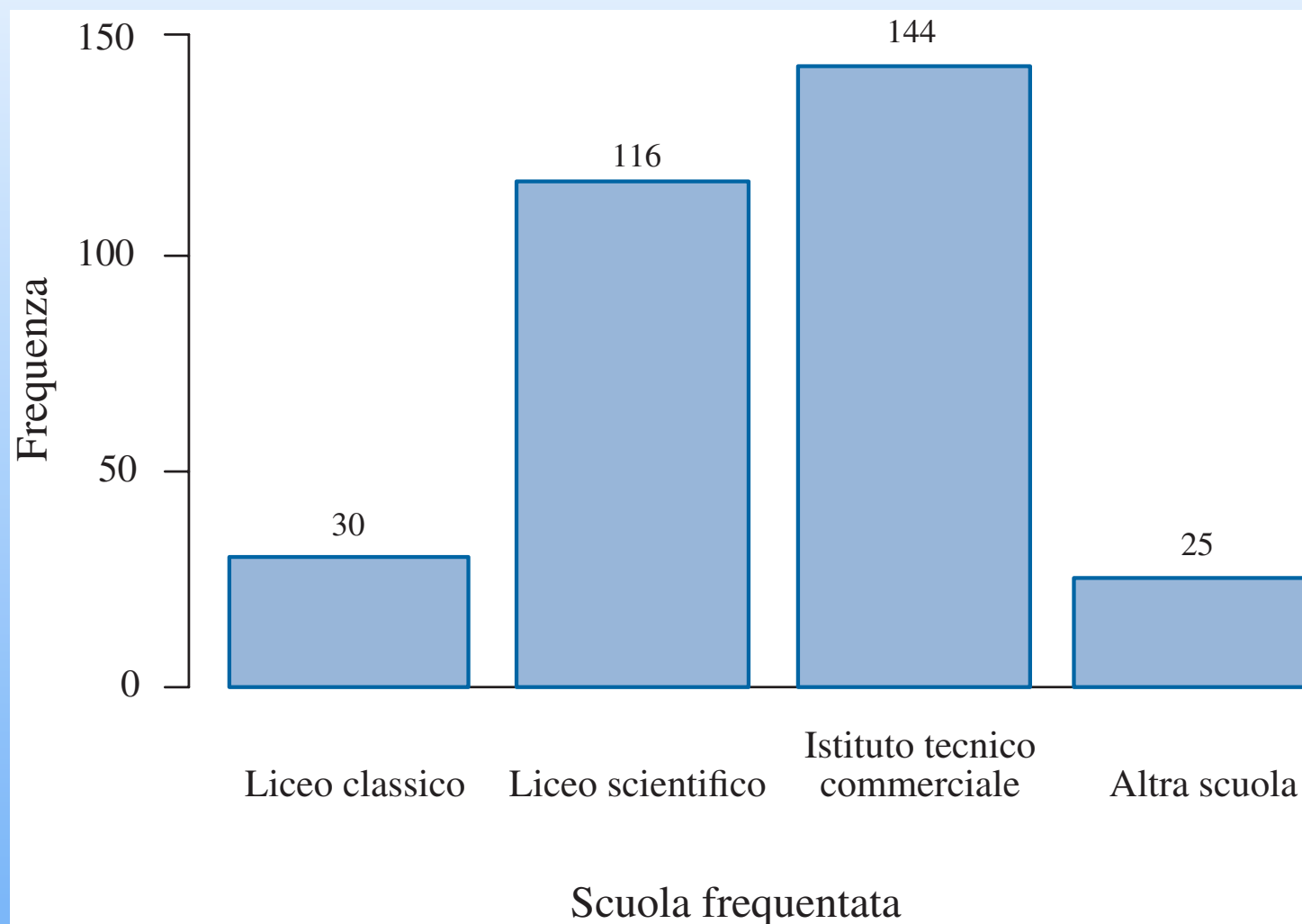
x_i	n_i
Per nulla	2
Poco	4
Abbastanza	8
Molto	6
Totale	20

Naturalmente, il grafico a nastri può essere realizzato in modi diversi.

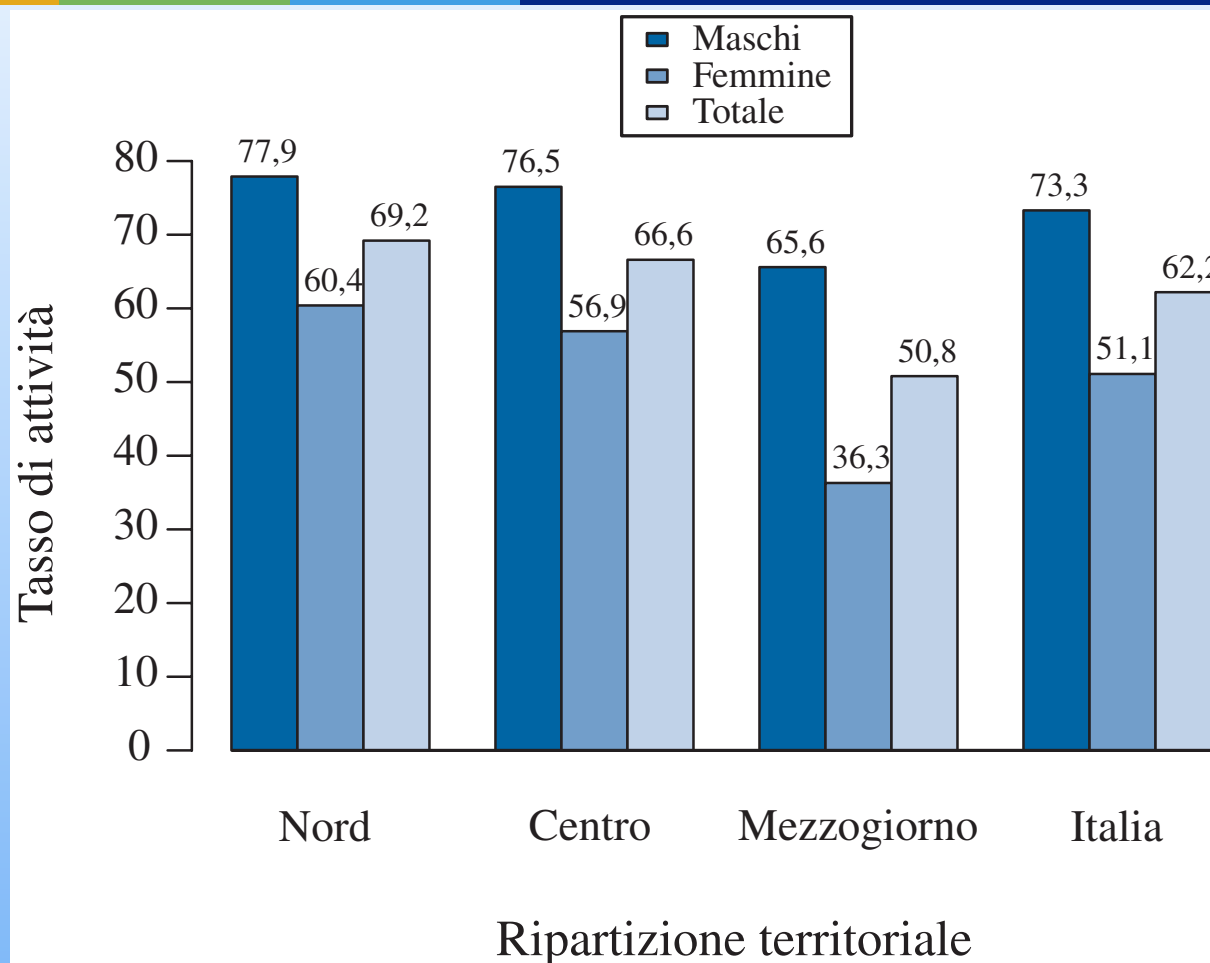
Ad esempio, disponendo i nastri in orizzontale

L'aspetto visivo del grafico rimane immutato se si utilizzano le frequenze relative o percentuali

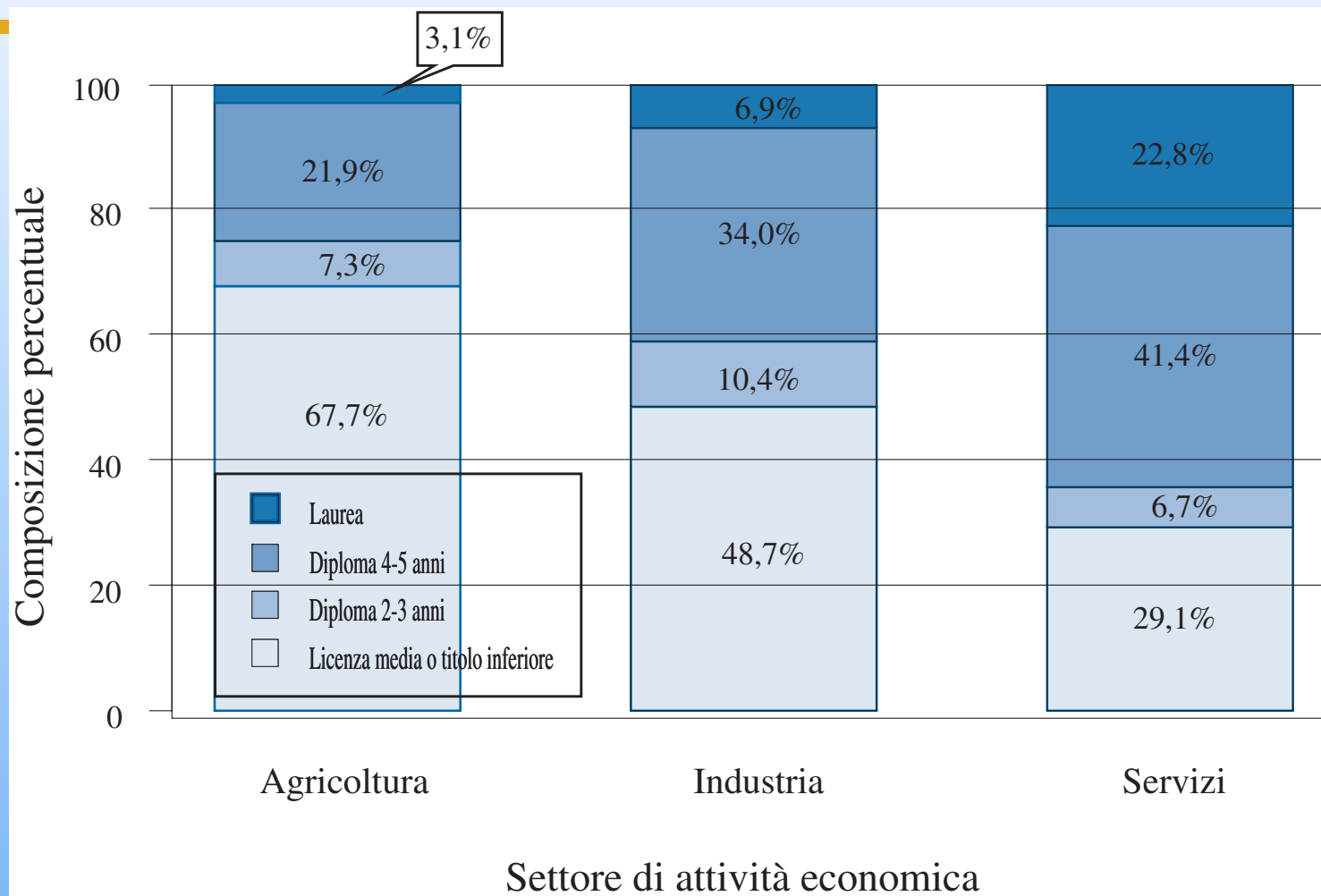
Grafico a nastri per la serie sconnessa dei laureati per scuola superiore



Tassi di attività per ripartizione territoriale e sesso - Anno 2010



Occupati per grado di istruzione e settore di attività



Il grafico a nastri si presta a essere utilizzato anche per i rapporti di composizione.

Settori circolari

Una rappresentazione grafica alternativa, utilizzabile quando il numero delle modalità non è elevato, è il grafico a settori circolari, in cui le frequenze o le quantità associate alle varie modalità del carattere vengono rappresentate con le **aree dei settori circolari** in cui è suddiviso un cerchio

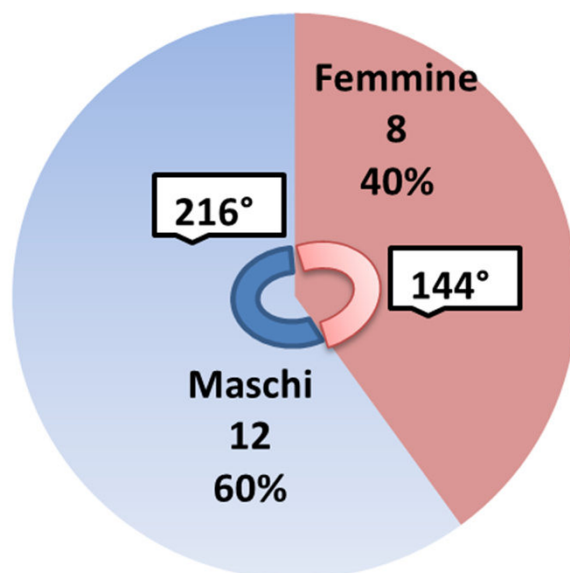
Esempio Settori circolari

<i>Sesso</i>	<i>frequenza assoluta</i>	<i>frequenza relativa</i>	<i>angolo</i>
x_i	n_i	f_i	α_i
Femmine	8	0.4	144
Maschi	12	0.6	216
Totale	20	1	360

$$n_i : N = \alpha_i : 360^\circ$$

$$\begin{aligned}\alpha_i &= \frac{n_i}{N} \cdot 360^\circ \\ &= f_i \cdot 360^\circ\end{aligned}$$

Distribuzione del collettivo di studenti
rispetto al sesso



Settori circolari

distribuzione dei laureati per scuola frequentata

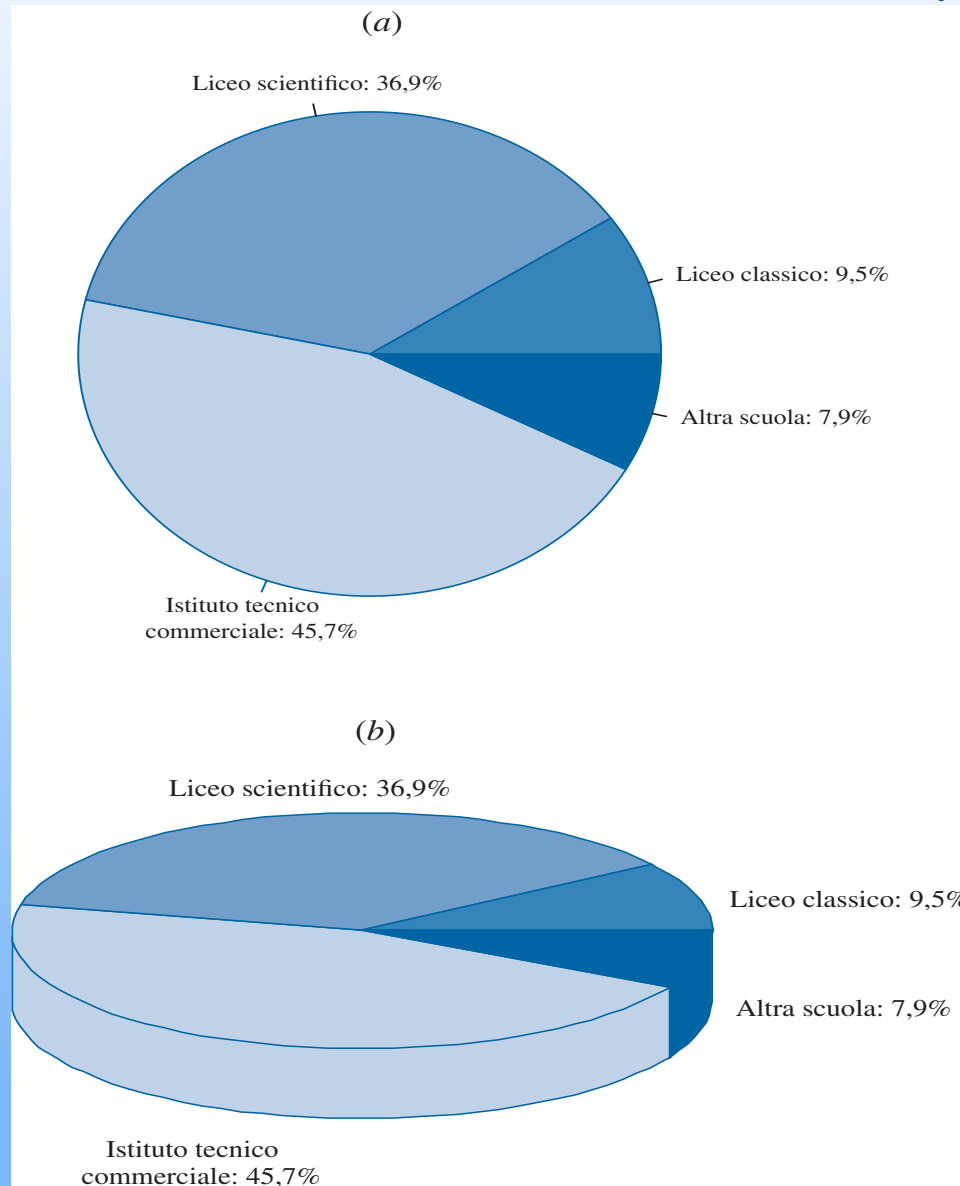


grafico a settori circolari
versione bidimensionale

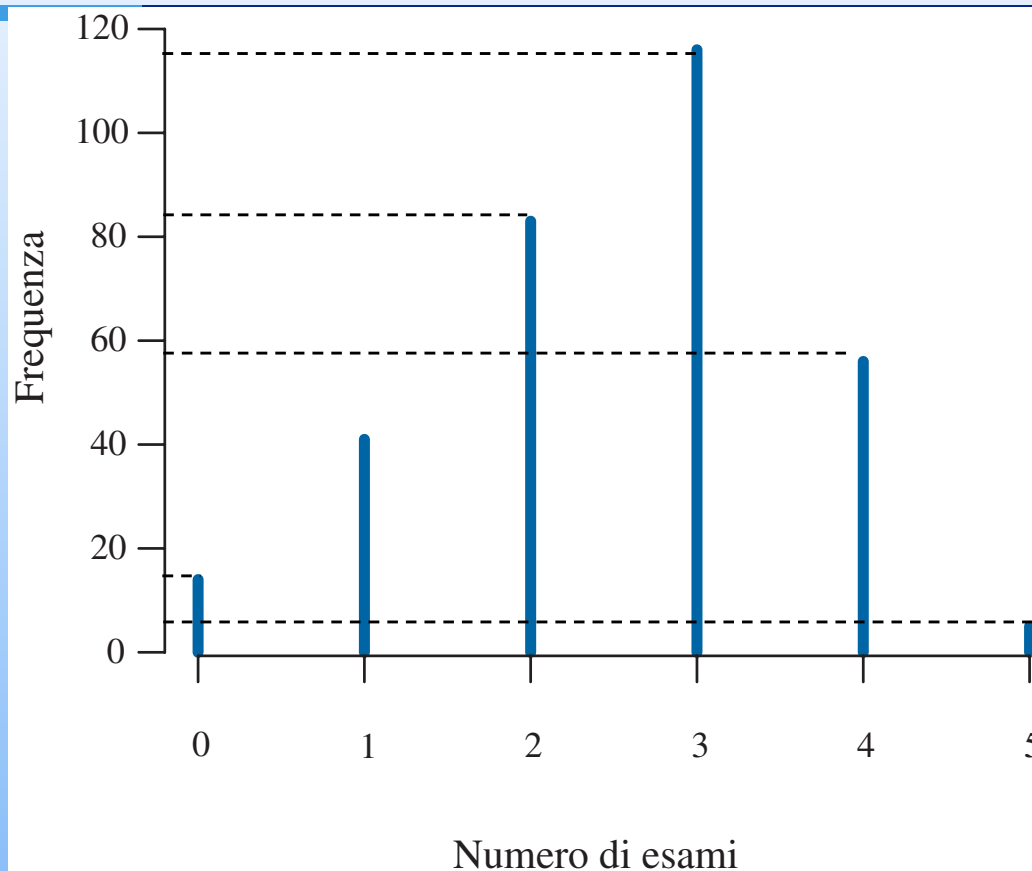
grafico a settori circolari
versione tridimensionale

Caratteri quantitativi discreti

- La rappresentazione grafica più idonea per una distribuzione di frequenze secondo un carattere discreto è quella cartesiana
- Sull'asse delle ascisse vengono poste le modalità x_1, x_2, \dots, x_k , sull'asse delle ordinate le frequenze corrispondenti n_1, n_2, \dots, n_k
- La rappresentazione grafica va sotto il nome di **diagramma ad aste**

Esempio: laureati per numero di esami sostenuti nel primo anno di corso

N. esami	Frequenza
0	14
1	41
2	83
3	116
4	56
5	5
Totale	315



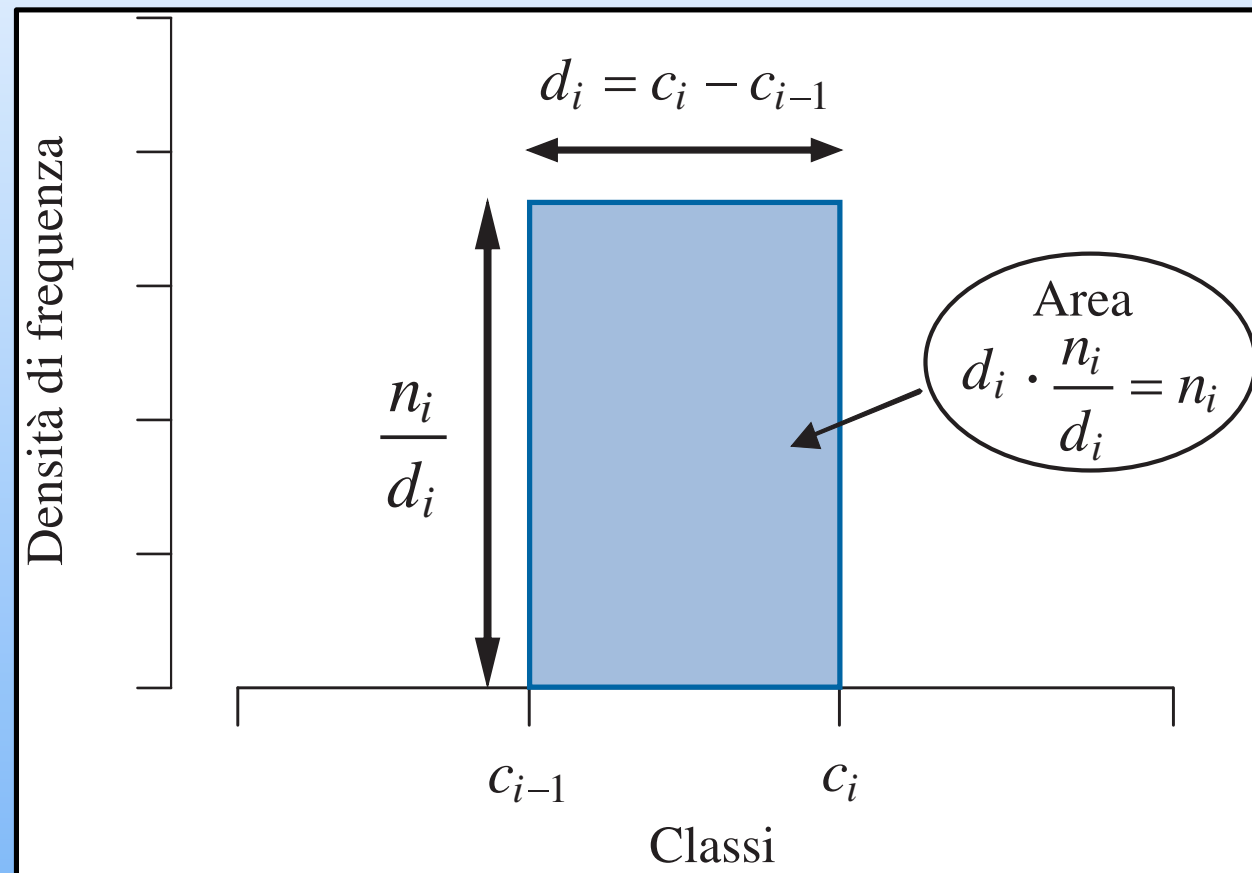
L'uso delle **frequenze relative (o percentuali)** non altera l'immagine della distribuzione fornita dal grafico.

Caratteri quantitativi continui divisi in intervalli

La rappresentazione grafica più appropriata è l'**istogramma** in cui l'area del rettangolo rappresenta la frequenza assoluta o relativa della corrispondente classe.

L'istogramma si ottiene ponendo sull'asse delle ascisse gli estremi di classe c_0, c_1, \dots, c_k e disegnando per ogni classe (c_{i-1}, c_i) , $i = 1, 2, \dots, k$, un **rettangolo** avente per base il segmento dell'asse delle ascisse di estremi c_{i-1} e c_i e per altezza la **densità di frequenza** n_i/d_i , dove d_i è l'ampiezza di classe

Istogramma

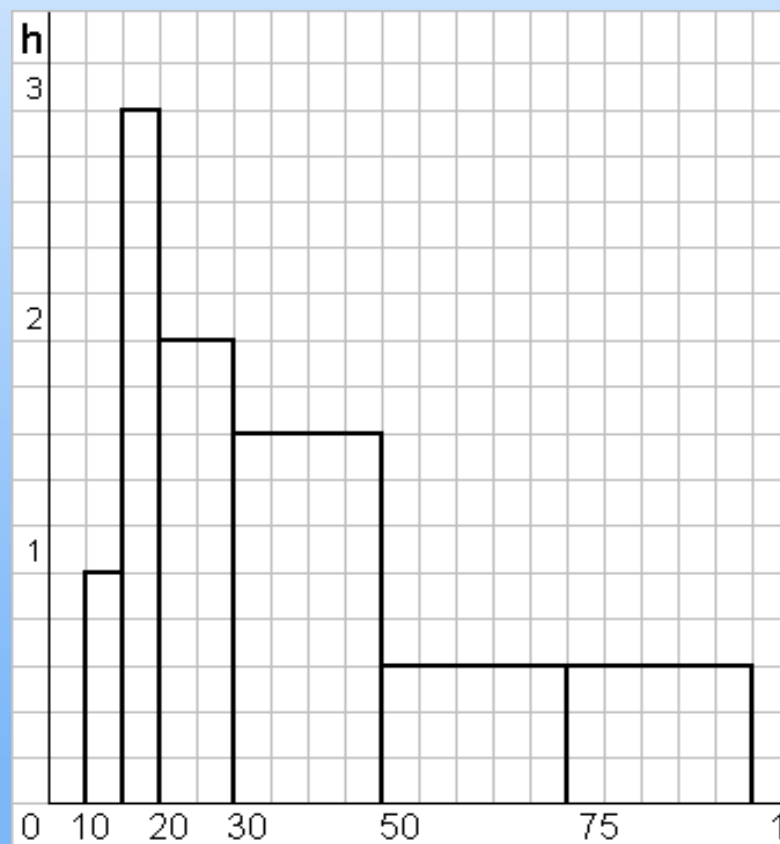


Istogramma: esempio

Classi di peso (in Kg)	Frequenza assoluta	Ampiezza di classe	Densità di frequenza
$c_{i-1} - c_i$	n_i	d_i	h_i
10 -- 15	5	5	1
15 -- 20	15	5	3
20 -- 30	20	10	2
30 -- 50	30	20	1,5
50 -- 75	15	25	0,6
75 -- 100	15	25	0,6
Totale	100		

$$d_i = c_{i-1} - c_i$$

$$h_i = \frac{n_i}{d_i}$$



Nell'istogramma si costruiscono dei rettangoli la cui area è proporzionale alla frequenza della classe.

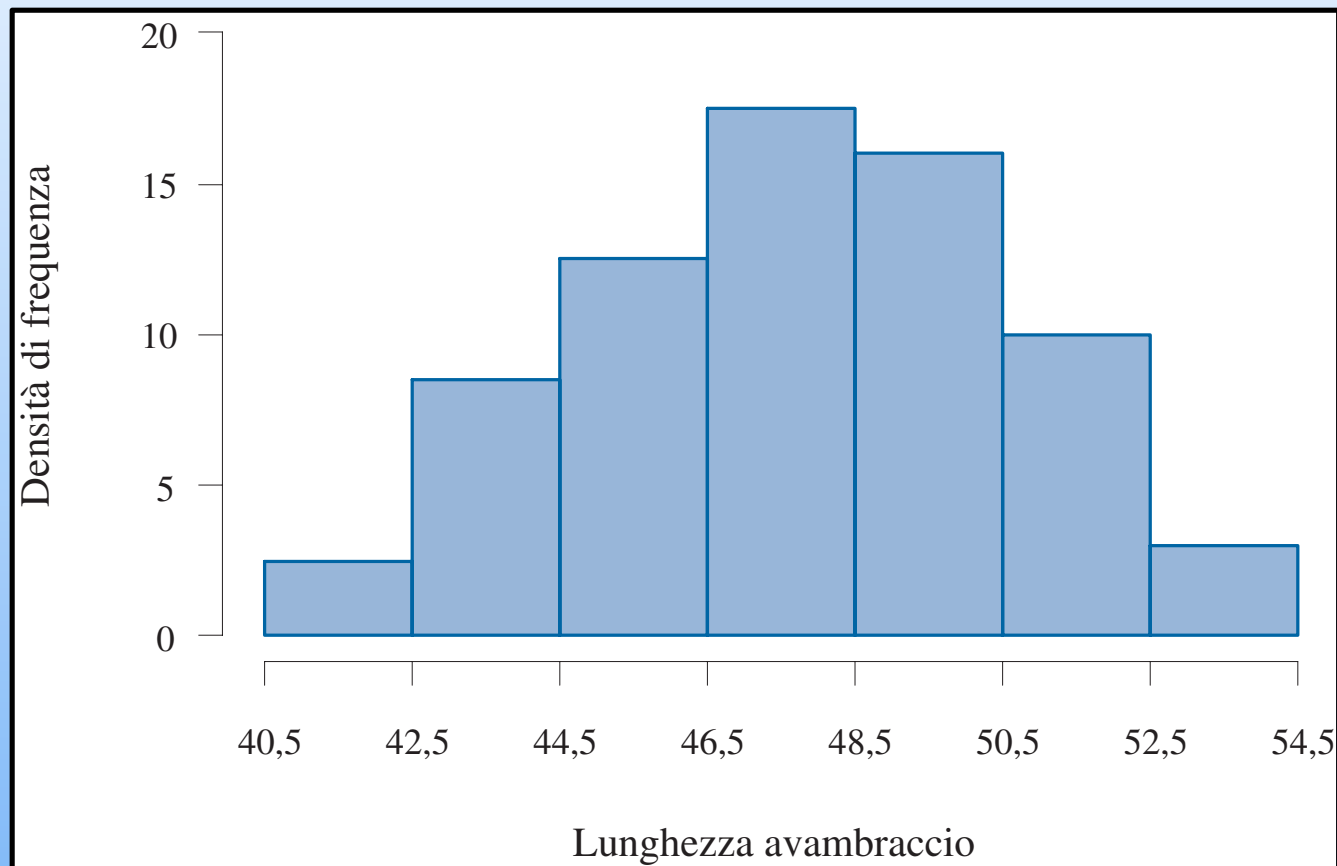
La base è data dall'ampiezza di classe, l'altezza dalla densità di frequenza

Istogramma: esempio

Distribuzione di frequenze della lunghezza dell'avambraccio

Lunghezza avambraccio (cm)	Frequenza	Ampiezza della classe	Densità di frequenza
40.5-42.5	5	2	2.5
42.5-44.5	17	2	8.5
44.5-46.5	25	2	12.5
46.5-48.5	35	2	17.5
48.5-50.5	32	2	16.0
50.5-52.5	20	2	10.0
52.5-54.5	6	2	3.0
Totale	140		

Istogramma per la distribuzione delle lunghezze dell'avambraccio



Serie storiche

- Per la rappresentazione grafica delle serie storiche riferite a **fenomeni di stato**, si ricorre, generalmente, ai **diagrammi cartesiani**. Si pongono sull'asse delle ascisse i tempi e su quello delle ordinate le intensità associate: i conseguenti punti del piano cartesiano vengono poi uniti con **segmenti di retta** per facilitare la percezione visiva dell'andamento del fenomeno.

Serie storiche

Quando la serie storica riguarda un **fenomeno di movimento**, la rappresentazione grafica più appropriata è quella a **nastri**. I periodi vengono indicati lungo una linea orizzontale, mentre su un asse verticale viene riportata l'appropriata scala; sulla base di questa, si tracciano dei rettangoli, con la stessa base e altezze pari alle **intensità** da rappresentare.

Serie storiche

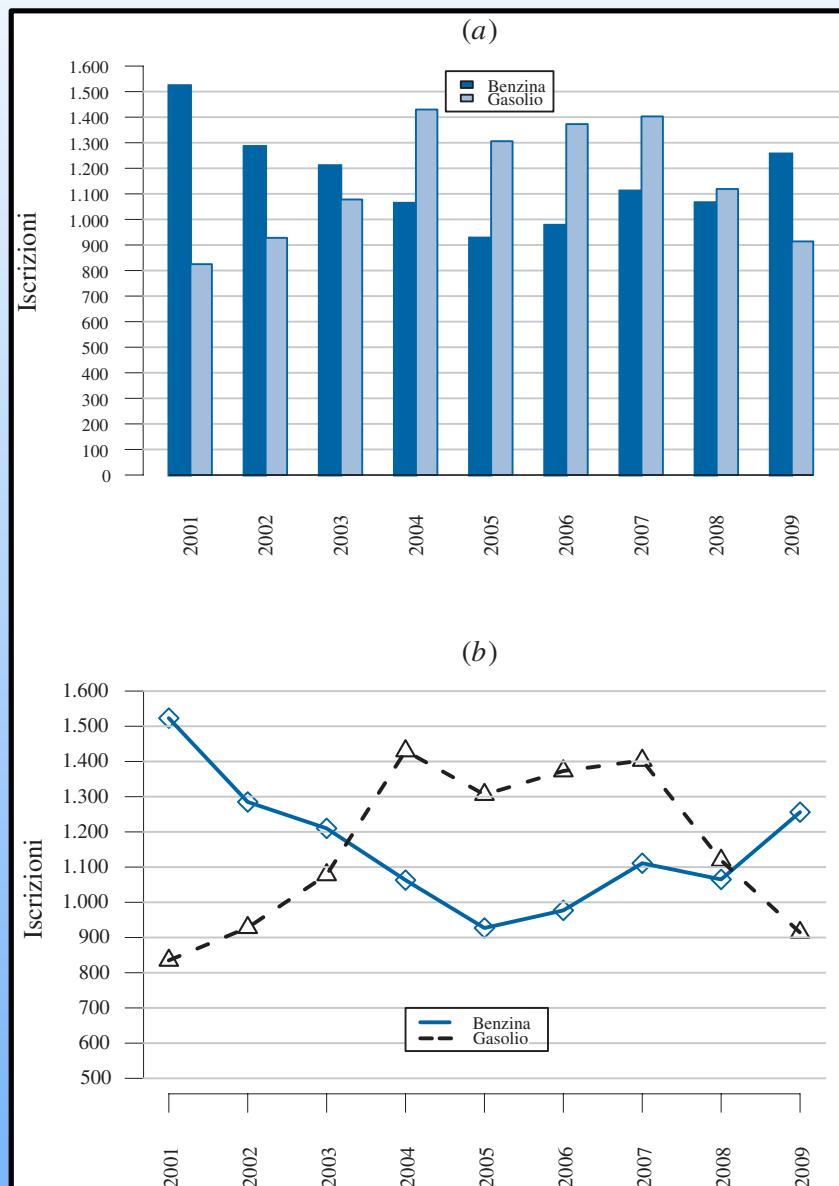


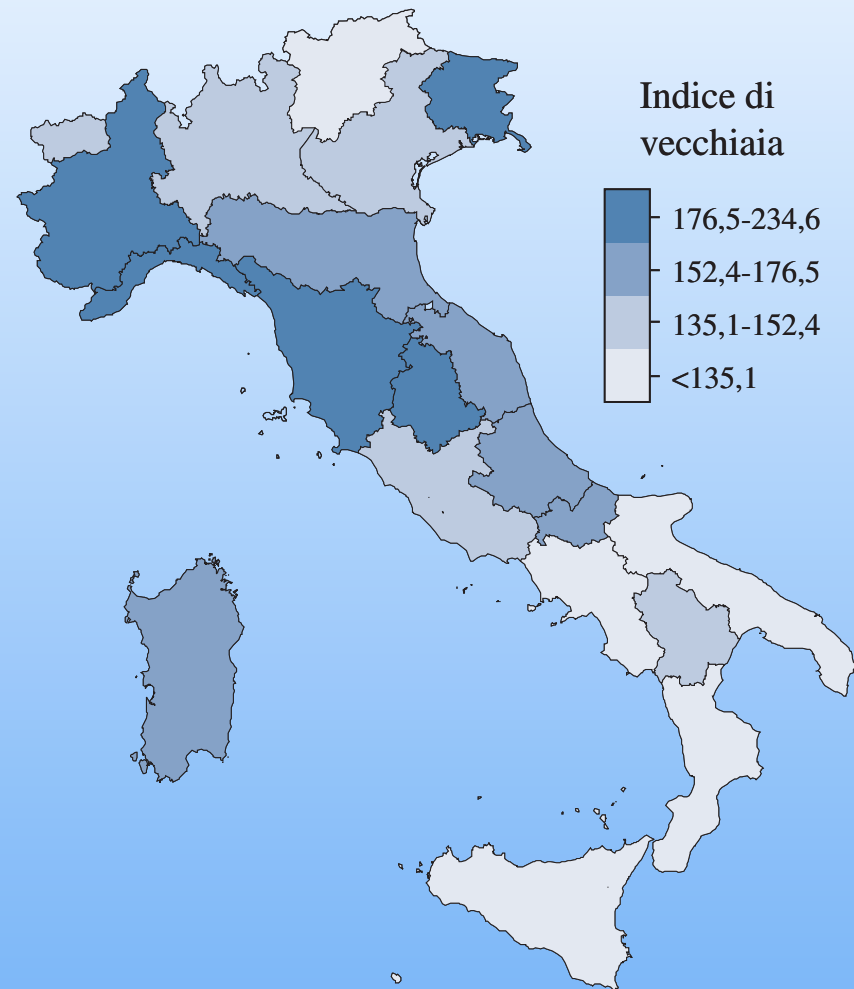
Grafico più appropriato
fenomeno movimento

Grafico più efficace mostra
l'andamento del fenomeno

Serie territoriali

- si rappresentano spesso mediante i **cartogrammi**: le ripartizioni territoriali sono individuate sull'appropriata **cartina geografica**; le intensità corrispondenti (frequenze o quantità) vengono rappresentate tramite **colori** o **tratteggi diversi**, il cui significato è specificato in apposite **legende**
- una variante è il **cartodiagramma**: a ogni ripartizione territoriale viene associato un grafico relativo a una distribuzione o ai livelli di uno o più fenomeni

cartogramma relativo all'indice di
vecchiaia per regione



Funzione di ripartizione per caratteri quantitativi discreti

Consideriamo il sottoinsieme così definito

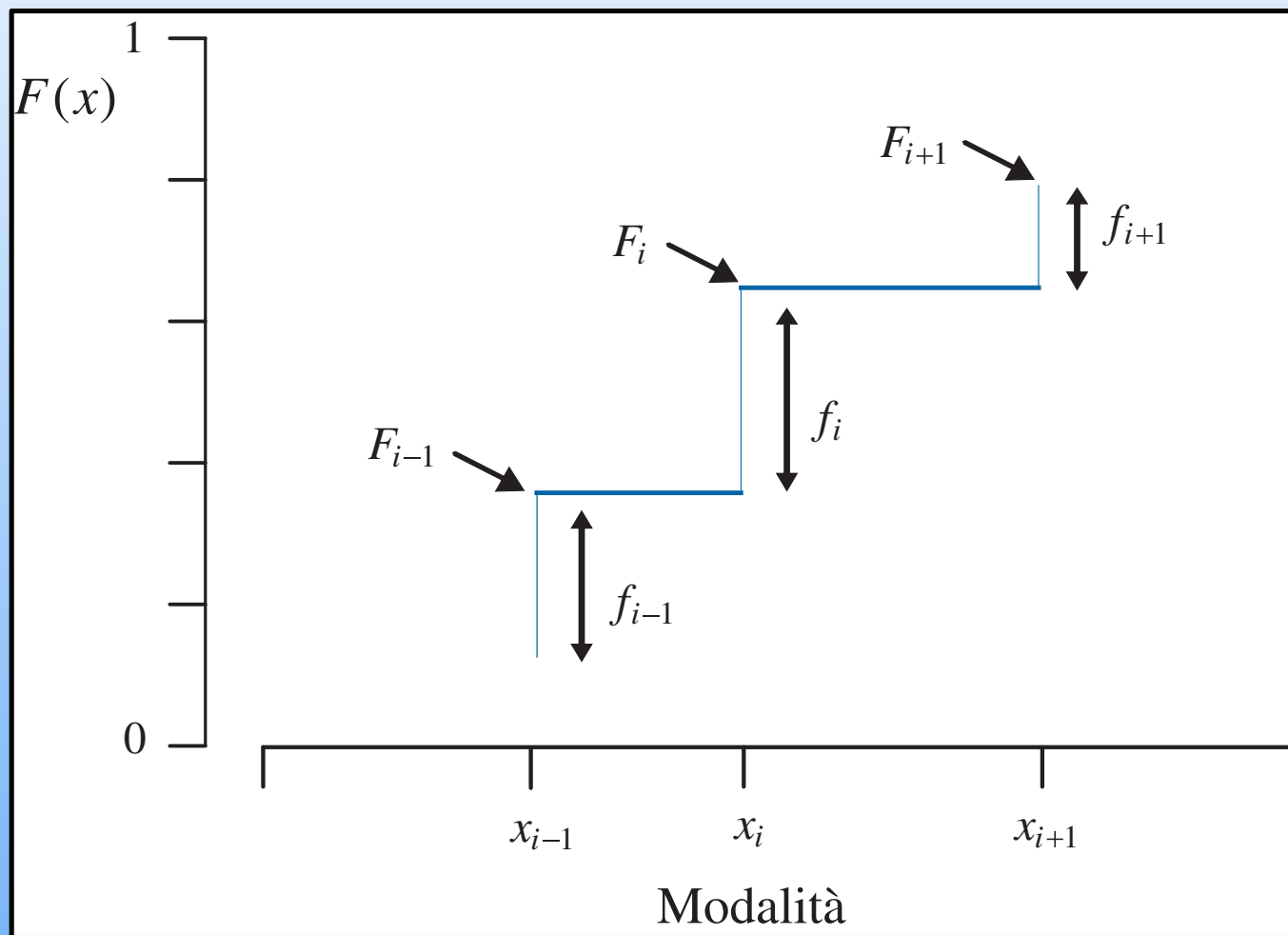
$$C_x = \{u : X(u) \leq x\},$$

costituito dalle unità del collettivo in cui il carattere assume un valore **minore o uguale** a un livello assegnato x .

Si chiama funzione di ripartizione, $F(x)$, il **rapporto** tra la **numerosità** di C_x e il **totale delle unità** N .

Si tratta della **frequenza relativa** delle unità del collettivo nelle quali il carattere assume un valore non superiore alla quantità x .

Funzione di ripartizione nel caso discreto

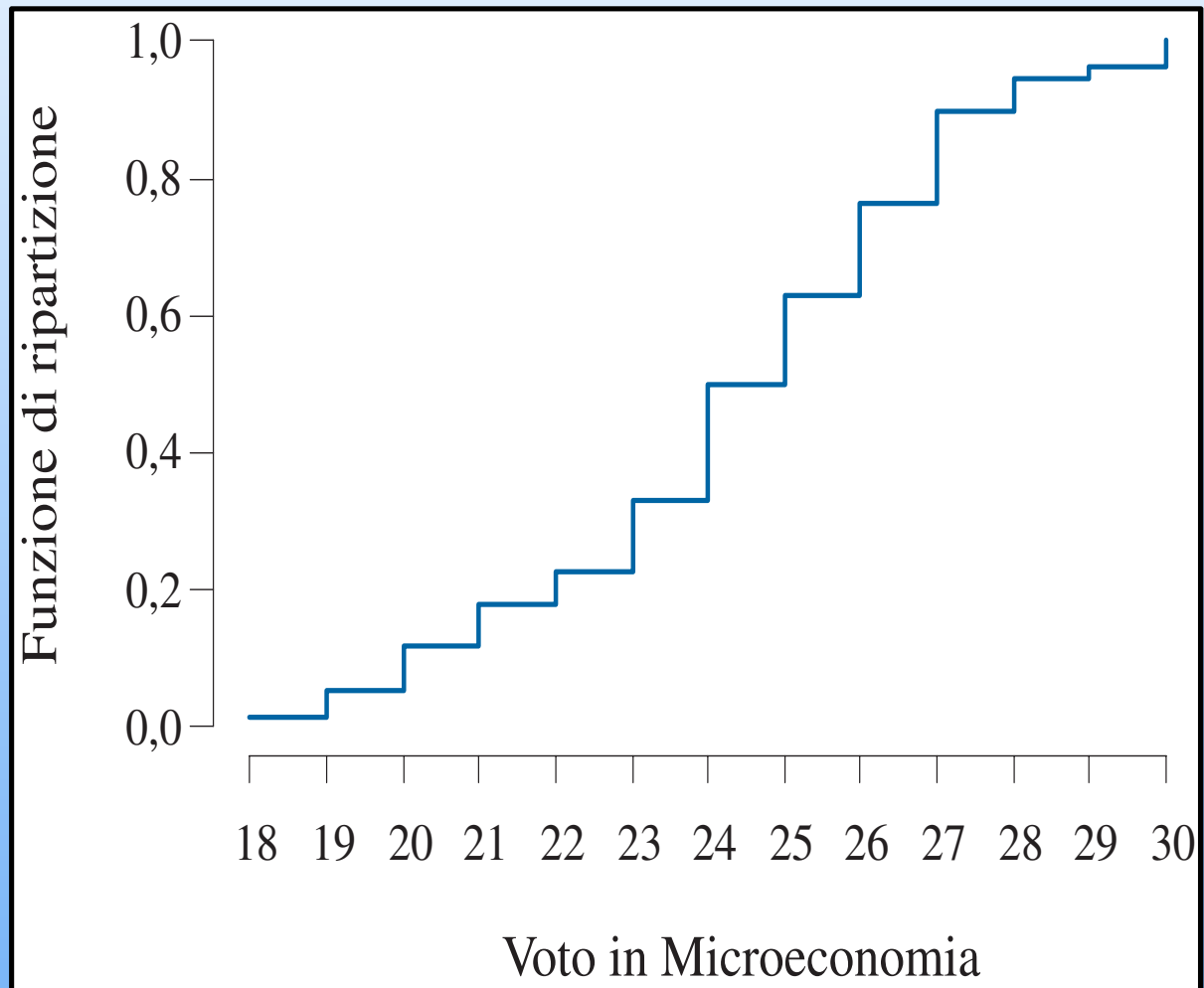


Proprietà della funzione di ripartizione per caratteri quantitativi discreti

- È definita per qualsiasi valore di x
- È pari a F_{i-1} quando $x_{i-1} \leq x < x_i$ (F_{i-1} : frequenza relativa cumulata)
- Assume il valore 0 quando x è minore di x_1 (perché nel collettivo non vi sono unità con modalità più piccole di x_1)
- Assume il valore 1 quando x è maggiore o uguale a x_k (perché nel collettivo non vi sono unità con modalità maggiori di x_k)
- È non decrescente

Funzione di ripartizione per il voto in microeconomia

Modalità	Frequenza relativa cumulata
18	0.01
19	0.05
20	0.12
21	0.18
22	0.23
23	0.33
24	0.50
25	0.63
26	0.77
27	0.90
28	0.95
29	0.96
30	1.00



Funzione di ripartizione per caratteri quantitativi continui

Consideriamo il sottoinsieme così definito

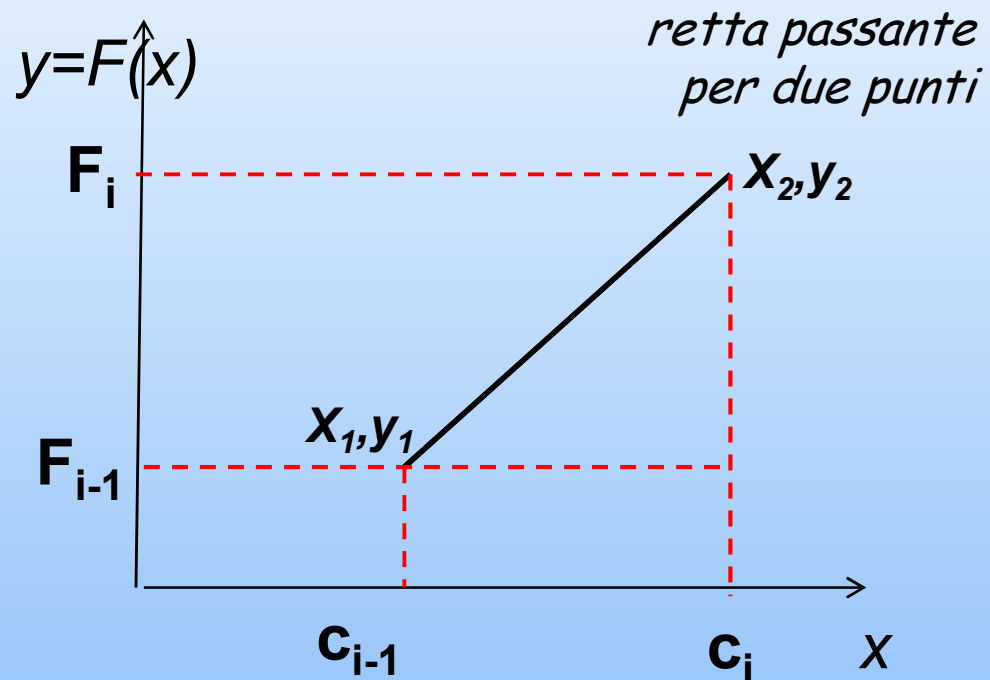
$$C_x = \{u : X(u) \leq x\},$$

nel caso di un carattere continuo diviso in intervalli, la funzione di ripartizione è la frequenza relativa delle unità del collettivo in cui il carattere X non supera un fissato livello x .

Ipotizzando che le unità siano uniformemente distribuite all'interno delle classi la funzione di ripartizione è espressa da

$$F_i(x) = F_{i-1} + \frac{1}{N} \times \frac{n_i}{d_i} (x - c_{i-1}) = F_{i-1} + \frac{f_i}{d_i} (x - c_{i-1})$$

Spiegazione della formula della funzione di ripartizione caratteri quantitativi continui



$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}$$

$$(x_1, y_1) = (c_{i-1}, F_{i-1});$$

$$(x_2, y_2) = (c_i, F_i)$$

$$\frac{y - F_{i-1}}{F_i - F_{i-1}} = \frac{x - c_{i-1}}{c_i - c_{i-1}}$$

$$y = F_{i-1} + \frac{F_i - F_{i-1}}{c_i - c_{i-1}}(x - c_{i-1})$$

$$F_i(x) = F_{i-1} + \frac{f_i}{d_i}(x - c_{i-1})$$

Proprietà della funzione di ripartizione per caratteri quantitativi continui

$F(x)$:

- È definita per qualsiasi valore di x
- È pari a 0 quando x è minore o uguale a c_0
- È pari a 1 quando x è maggiore o uguale a c_k
- È crescente all'interno di ogni classe assumendo l'andamento di una retta

Funzione di ripartizione per la distribuzione delle lunghezze dell'avambraccio

Lunghezza avambraccio (cm)	Frequenza	Frequenza relativa cumulata
40.5-42.5	5	0.04
42.5-44.5	17	0.16
44.5-46.5	25	0.34
46.5-48.5	35	0.59
48.5-50.5	32	0.81
50.5-52.5	20	0.96
52.5-54.5	6	1.00
Totale	140	

