

# Sintesi della distribuzione di un carattere

## le medie

### Obiettivi dell'unità didattica

- Descrivere e sintetizzare le informazioni disponibili
- Calcolare le principali misure di tendenza centrale
- Interpretare i risultati ottenuti

In questa unità didattica verranno trattati i seguenti argomenti:

- **La moda**
- **La mediana**
- **I percentili**
- **La media aritmetica**
- **La media geometrica**

### Premessa e contenuti

Abbiamo già visto come uno degli obiettivi principali sia quello di sintetizzare le informazioni disponibili. Nella precedente Unità didattica abbiamo imparato come, a partire dalle informazioni ottenute attraverso l'effettuazione di un'indagine, sia possibile operare una prima sintesi costruendo (in modo opportuno) delle tabelle e dei grafici.

Un ulteriore passo in questa direzione è l'utilizzo di un'unica informazione quantitativa che ci permetta di descrivere, con un solo "numero", tutta la distribuzione dei dati che abbiamo a disposizione.

### MODA

è quella modalità del carattere di una distribuzione che si presenta con la frequenza più alta.

### MEDIANA

è la modalità del carattere presentata dall'individuo che occupa la posizione centrale in una distribuzione ordinata di elementi.

### QUARTILI

sono quei valori della distribuzione in grado di dividere la distribuzione in quattro parti uguali.

### PERCENTILI

sono quei valori della distribuzione in grado di dividere la distribuzione in cento parti uguali.

## **MEDIA ARIMETICA**

è data dalla sommatoria delle modalità osservate diviso il numero di unità statistiche considerate:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

se abbiamo a disposizione una distribuzione di frequenze, è necessario tenere in considerazione anche il numero di volte in cui tali modalità si sono presentate (le frequenze, appunto):

$$\bar{X} = \frac{\sum_{i=1}^n (X_i \cdot n_i)}{n}$$

Quando, poi, il carattere sarà suddiviso in classi, utilizzeremo, nel calcolo della media aritmetica, il valore centrale della classe:

$$\bar{X} = \frac{\sum_{i=1}^n (c_i \cdot n_i)}{n}$$

## **MEDIA GEOMETRICA**

di  $n$  valori positivi è data dalla radice  $n$ -esima del prodotto dei singoli valori.

$$\bar{X}_g = \sqrt[n]{X_1 \cdot X_2 \cdot \dots \cdot X_n}$$

# Sintesi della distribuzione di un carattere: le medie di posizione

## Le medie

- Conoscere le caratteristiche principali del fenomeno investigato  
Quale reddito disponibile ha mediamente una certa popolazione?  
Qual è l'età media al pensionamento di un gruppo di lavoratori?  
Qual è la classe di addetti in cui si addensano il maggior numero delle imprese operanti in un territorio?
- Informazione quantitativa che, da sola, ci dà un'idea del problema che stiamo analizzando
- Interpretazione vs. Contenuto

## La Moda

- **Defin.:** E' la modalità del carattere che si presenta con la frequenza più elevata
- **Caratteri qualitativi e quantitativi**

$X_i$	$n_i$
Maschi	170
Femmine	130
	300

$X_i$	$n_i$
Economia	120
Statistica	100
Matematica	80
	300

$X_i$	$n_i$
0	45
1	60
2	60
3	50
4	35
	250

**Moda 2.**

$X_i$	$n_i$	$h_i$
10  — 30	78	3,90
30  — 60	96	3,20
60  — 100	69	1,72
100  — 200	42	0,42
200  — 500	15	0,05
	300	

**12 ; 18; 20; 24 ; 22**

$X_i$	$n_i$
12	1
18	1
20	1
22	1
24	1
	5

**La Mediana**

**15 ; 12; 8; 11 ; 14**

**La MEDIANA è la modalità del carattere che occupa la posizione centrale in una distribuzione ordinata di dati**

**8 ; 11; 12; 14 ; 15**

**Dispari**

→  $\frac{(n+1)}{2}$

**Pari**

→  $\frac{n}{2}$  e  $\frac{n}{2}+1$

**La Mediana 2.**

$X_i$	$n_i$	$N_i$
1	11	11
2	21	32
3	32	64
4	28	92
5	16	108
6	7	115
	115	

$$\frac{(n+1)}{2} = \frac{(115+1)}{2} = \frac{116}{2} = 58$$

1,1,1,1..... 11 volte

2,2,2,2..... 21 volte

3,3,3,3..... 32 volte

4,4,4,4..... 28 volte

**La Mediana 3.**

$X_i$	$n_i$	$N_i$
170  —175	12	12
175  —180	28	40
180  —185	45	85
185  —190	20	105
190  —195	11	116
195  —200	4	120
	120	

$$\frac{n}{2} = \frac{120}{2} = 60$$

$$\frac{n}{2} + 1 = \frac{120}{2} + 1 = 60 + 1 = 61$$

### La mediana 4

$X_i$	$n_i$	$N_i$
1-5	45	45
6-9	35	80
10-19	26	106
20-29	20	126
30-39	15	141
40-49	10	151
	151	

$$\frac{(n+1)}{2} = \frac{(151+1)}{2} = \frac{152}{2} = 76$$

- Quante sono le imprese intervistate?
- Qual è la moda?
- Qual è la mediana?

**N=151**

**Moda = (1-5)**

**Mediana = (6-9)**

### La Mediana 5

$X_j$	$n_j$	$N_j$
0  — 20	94	94
20  — 30	104	198
30  — 50	132	330
50  — 60	121	451
60  — 80	70	521
80  — 100	18	539
	539	

$$\frac{(n+1)}{2} = \frac{(539+1)}{2} = \frac{540}{2} = 270$$

- Quante sono le unità considerate?
- Di che tipo di carattere si tratta?
- Qual è la moda?
- Qual è la mediana?

**N=539**

**Moda = (50-60)**

**Mediana = (30-50)**

## I percentili

**Percentili:** sono quei valori della distribuzione in grado di dividere la distribuzione in 100 parti uguali

**Mediana:** e' il 50-esimo percentile

**Quartili:** dividono la distribuzione in 4 parti uguali

**1° Quartile:** 25% più piccoli, 75% più grandi

**2° Quartile:** 50% più piccoli, 50% più grandi

**3° Quartile:** 75% più piccoli, 25% più grandi

**Mediana:** e' il 2° Quartile

**Calcolo:** come la mediana: si ordina la distribuzione, si individua la posizione, si osserva la modalità presentata

# Sintesi della distribuzione di un carattere: le medie analitiche

- Non individuano una posizione, ma sono il risultato di una operazione algebrica
- Coinvolge tutti gli elementi della distribuzione
- Caratteri quantitativi

## La media aritmetica

10 8 15 7

$$(10 + 8 + 15 + 7)/4 = 40/4 = 10$$

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

## La media aritmetica -Proprietà

1° Proprietà

è sempre compresa tra la più piccola e la più grande delle modalità presenti nella distribuzione

7 8 10 15

2° Proprietà

la somma di tutte le differenze tra ogni  $x_i$  e la media aritmetica è sempre uguale a zero

$$\sum_{i=1}^n (X_i - \bar{X}) = 0$$

$$(10 - 10) + (8 - 10) + (15 - 10) + (7 - 10)$$

$$= 0 - 2 + 5 - 3 = 0$$

### 3° Proprietà

la somma di tutte le differenze tra ogni  $x_i$  e la media aritmetica elevate al quadrato costituisce un “minimo”

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

$$\begin{aligned} & (10 - 10)^2 + (8 - 10)^2 + (15 - 10)^2 + (7 - 10)^2 = \\ & = 0^2 - 2^2 + 5^2 - 3^2 = 0 + 4 + 25 + 9 = 38 \end{aligned}$$

### La media aritmetica – Distribuzioni di frequenze

$$\bar{X} = \frac{\sum_{i=1}^n (x_i \cdot n_i)}{n}$$

$x_i$	$n_i$	$x_i \cdot n_i$
1	11	1 x 11 = 11
2	21	2 x 21 = 42
3	32	3 x 32 = 96
4	28	4 x 28 = 112
5	16	5 x 16 = 80
6	7	6 x 7 = 42
	115	$\Sigma(x_i \cdot n_i) = 383$

$$\bar{X} = \frac{\sum_{i=1}^n (x_i \cdot n_i)}{n} = \frac{383}{115} = 3,3$$

### La media aritmetica – modalità divise in classi

$X_i$	$n_i$	$N_i$	$C_i$	$C_i \cdot n_i$
170   175	12	12	172,5	$172,5 \times 12 = 2.070,0$
175   180	28	40	177,5	$177,5 \times 28 = 4.970,0$
180   185	45	85	182,5	$182,5 \times 45 = 8.212,5$
185   190	20	105	187,5	$187,5 \times 20 = 3.750,0$
190   195	11	116	192,5	$192,5 \times 11 = 2.117,5$
195   200	4	120	197,5	$197,5 \times 4 = 790,0$
	120			21.910,0

**Moda = 180-185      Mediana = 180-185**

$$\bar{X} = \frac{\sum_{i=1}^n (c_i \cdot n_i)}{n} = \frac{21.910}{120} = 182,6$$

### La Media aritmetica – esercizio 1

$X_i$	$n_i$	$N_i$	$C_i$	$C_i \cdot n_i$
1-5	45	45	3,0	$3 \times 45 = 135,0$
6-9	35	80	7,5	$7,5 \times 35 = 262,5$
10-19	26	106	14,5	$14,5 \times 26 = 377,0$
20-29	20	126	24,5	$24,5 \times 20 = 490,0$
30-39	15	141	34,5	$34,5 \times 15 = 517,5$
40-49	10	151	44,5	$44,5 \times 10 = 445,0$
	151			2.227,0

**Moda = 1 - 5      Mediana = 6 - 9**

$$\bar{X} = \frac{\sum_{i=1}^n (c_i \cdot n_i)}{n} = \frac{2.227}{151} = 14,7$$

## La Media aritmetica – esercizio 2

$X_j$	$n_j$	$C_i$	$C_i \cdot n_j$
0   20	94	10	$10 \times 94 = 940$
20   30	104	25	$25 \times 104 = 2.600$
30   50	132	40	$40 \times 132 = 5.280$
50   60	121	55	$55 \times 121 = 6.655$
60   80	70	70	$70 \times 70 = 4.900$
80   100	18	90	$90 \times 18 = 1.620$
	539		21.995

$$\bar{X} = \frac{\sum_{i=1}^n (c_i \cdot n_i)}{n} = \frac{21.995}{539} = 40,8$$

## La media aritmetica e la mediana

5; 8; 9; 11; 12; 15; 320

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{5+8+9+11+12+15+320}{7} = \frac{350}{7} = 50$$

$$Me = 11$$

Generalmente, in caso di outliers la mediana rappresenta meglio i dati

## La media geometrica

Quando per la natura stessa dei valori ha senso il loro prodotto più della loro somma

- capitalizzazione (tassi di interesse)
- prezzi (tassi di inflazione)

$$\bar{X}_g = \sqrt[n]{X_1 \cdot X_2 \cdot \dots \cdot X_n}$$

## La media geometrica -esempio

2,7    3,1    3,4    2,9    3,2

1,027    1,031    1,034    1,029    1,032

$$\bar{X}_g = \sqrt[n]{X_1 \cdot X_2 \cdot \dots \cdot X_n}$$

$$\sqrt[5]{1,027 \cdot 1,031 \cdot 1,034 \cdot 1,029 \cdot 1,032}$$

**3,06%**

## **Le medie –competenze acquisite**

### **Cosa abbiamo imparato?**

- Sintetizzare la distribuzione
- Medie di posizione
- Medie analitiche

### **Quando usare le une o le altre?**

(Dipende....spesso il risultato è molto simile)

- Dal tipo di caratteri
- Dalla distribuzione che incontriamo
- Dagli aspetti concettuali