

# Esercitazione 1

## ESERCIZIO 1

Si consideri il carattere titolo di studio. Di seguito è riportata la distribuzione di frequenze. Si individuino mediana, quartili e moda.

Titolo di studio	$n_i$
nessuno	2293
licenza elementare	4240
licenza media inferiore	5671
diploma media superiore	5738
diploma universitario	146
laurea	1421
specializzazione post-laurea	42
Totale	19551

## SOLUZIONE

Titolo di studio	$n_i$	$C_i$
nessuno	2293	2293
licenza elementare	4240	6533
licenza media inferiore	5671	12204
diploma media superiore	5738	17942
diploma universitario	146	18088
laurea	1421	19509
specializzazione post-laurea	42	19551
Totale	19551	

La mediana è la modalità corrispondente all'osservazione  $\frac{n+1}{2} = \frac{19552}{2} = 9776$ , dobbiamo quindi calcolare le frequenze cumulate e guardare in quale modalità cade tale osservazione. La mediana risulta quindi essere la modalità “licenza media inferiore”.

Il calcolo dei quartili procede in maniera analoga, i quartili vengono ottenuti come

$$\begin{aligned}q_1 &= x_{\left(\frac{n+1}{4}\right)} = x_{(4888)} \\q_2 &= x_{\left(2 \times \frac{n+1}{4}\right)} = x_{(9776)} \\q_3 &= x_{\left(3 \times \frac{n+1}{4}\right)} = x_{(14664)}\end{aligned}$$

Il primo quartile quindi cade nella modalità “licenza elementare”, il secondo quartile è la mediana, il terzo quartile cade nella modalità “diploma media superiore”.

La moda è data dalla modalità con frequenza più alta. Essendo il carattere in analisi di natura qualitativa, possiamo considerare le frequenze assolute. La moda corrisponde alla modalità “licenza media superiore”. Riguardo alla sua rappresentatività, la moda è un indice rappresentativo se la modalità corrispondente ha frequenza relativa maggiore di 0.5, in questo caso è inferiore.

## ESERCIZIO 2

Si considerino i seguenti 10 individui, per essi si osserva l'età. Si individuino mediana, quartili, moda, media aritmetica, media armonica, media quadratica e media geometrica.

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$
58	36	32	86	52	60	56	19	37	51

## SOLUZIONE

Ordiniamo la tabella in ordine crescente

	$x_{(1)}$	$x_{(2)}$	$x_{(3)}$	$x_{(4)}$	$x_{(5)}$	$x_{(6)}$	$x_{(7)}$	$x_{(8)}$	$x_{(9)}$	$x_{(10)}$
$x_i$	19	32	36	37	51	52	56	58	60	86
$\ln(x_i)$	2,944	3,466	3,584	3,611	3,932	3,951	4,025	4,060	4,094	4,454

La mediana corrisponde a

$$Me = x_{(\frac{n+1}{2})} = x_{(5,5)} = x_{(5)} + 0,5 \times (x_{(6)} - x_{(5)}) = 51,5$$

I quartili sono calcolati come

$$q_1 = x_{(\frac{n+1}{4})} = x_{(2,75)} = x_{(2)} + 0,75 \times (x_{(3)} - x_{(2)}) = 35$$

$$q_2 = Me$$

$$q_3 = x_{(8,25)} = x_{(8)} + 0,25 \times (x_{(9)} - x_{(8)}) = 58,5$$

La media aritmetica è calcolata come

$$M_1 = \frac{1}{n} (x_1 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{10} (487) = 48,7$$

La media armonica è calcolata come

$$M_{-1} = \frac{1}{\frac{1}{n} \left( \frac{1}{x_1} + \dots + \frac{1}{x_n} \right)} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}} = \frac{10}{0,241} = 41,494$$

La media quadratica è calcolata come

$$M_2 = \sqrt{\frac{1}{n} (x_1^2 + \dots + x_n^2)} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} = \sqrt{\frac{1}{10} 26851} = 51,818$$

La media geometrica è calcolata come

$$M_0 = \sqrt[n]{x_1 \times \dots \times x_n} = \exp \left[ \frac{1}{n} \sum_{i=1}^n \log(x_i) \right] = \exp \left[ \frac{1}{10} 38,122 \right] = 45,251$$

## ESERCIZIO 3

La seguente tabella riporta la distribuzione di frequenze delle età degli individui di età compresa tra i 31 anni e i 40 anni. Si individuino mediana, quartili, decili, moda, media aritmetica, media armonica, media quadratica e media geometrica.

$x_i$	31	32	33	34	35	36	37	38	39	40	TOTALE
$n_i$	222	227	227	215	201	294	259	282	271	276	2474

## SOLUZIONE

i	1	2	3	4	5	6	7	8	9	10
$x_i$	31	32	33	34	35	36	37	38	39	40
$n_i$	222	227	227	215	201	294	259	282	271	276
$C_i$	222	449	676	891	1092	1386	1645	1927	2198	2474
$\frac{1}{x_i}$	0,032	0,031	0,030	0,029	0,029	0,028	0,027	0,026	0,026	0,025
$\ln(x_i)$	3,434	3,466	3,497	3,526	3,555	3,584	3,611	3,638	3,664	3,689

La mediana corrisponde a

$$Me = x_{(\frac{n+1}{2})} = x_{(1237,5)} = 36$$

I quartili sono calcolati come

$$\begin{aligned} q_1 &= x_{(\frac{n+1}{4})} = x_{(618,75)} = 33 \\ q_2 &= Me \\ q_3 &= x_{(3 \times \frac{n+1}{4})} = x_{(1856,25)} = 38 \end{aligned}$$

I decili sono calcolati in maniera analoga

$$\begin{aligned} d_1 &= x_{(\frac{n+1}{10})} = x_{(247,5)} = 32 \\ d_2 &= x_{(2 \times \frac{n+1}{10})} = x_{(495)} = 33 \\ &\dots \end{aligned}$$

La moda è la modalità con frequenza maggiore, quindi “36”. Non è rappresentativa.

La media aritmetica è calcolata come

$$M_1 = \frac{1}{n} (n_1 x_1 + \dots + n_k x_k) = \frac{1}{n} \sum_{i=1}^k n_i x_i = \frac{1}{2474} (88474) = 35,762$$

La media armonica è calcolata come

$$M_{-1} = \frac{1}{\frac{1}{n} \left( n_1 \frac{1}{x_1} + \dots + n_k \frac{1}{x_k} \right)} = \frac{n}{\sum_{i=1}^k n_i \frac{1}{x_i}} = \frac{2474}{69,637} = 35,527$$

La media quadratica è calcolata come

$$M_2 = \sqrt{\frac{1}{n} (n_1 x_1^2 + \dots + n_k x_k^2)} = \sqrt{\frac{1}{n} \sum_{i=1}^k n_i x_i^2} = \sqrt{\frac{1}{2474} 3184352} = 35,877$$

La media geometrica è calcolata come

$$M_0 = \sqrt[n]{x_1^{n_1} \times \dots \times x_k^{n_k}} = \exp \left[ \frac{1}{n} \sum_{i=1}^k n_i \log(x_i) \right] = \exp \left[ \frac{1}{2474} 8841,104 \right] = 35,645$$

## ESERCIZIO 4

Si considera ora l'intero campione e di esso si analizza ancora l'età degli individui. Poiché le modalità sono numerose queste sono state raggruppate in classi. Si individuino mediana, quartili, moda, media aritmetica, media armonica, media quadratica e media geometrica.

i	$x_l - x_u$	$n_i$
1	18-30	5782
2	31-40	2474
3	41-50	2973
4	51-65	4191
5	66-80	4131

## SOLUZIONE

i	$x_l - x_u$	$n_i$	$C_i$	$a_i$	$c_i$	$\frac{1}{c_i}$	$n_i \times \frac{1}{c_i}$	$c_i^2$	$c_i^2 \times n_i$	$\ln(c_i)$	$\ln(c_i) \times n_i$
1	18-30	5782	5782	13	24	0,0417	240,917	576	3330432	3,178	18375,507
2	31-40	2474	8256	10	35,5	0,0282	69,690	1260,25	3117858,5	3,570	8831,024
3	41-50	2973	11229	10	45,5	0,0220	65,341	2070,25	6154853,25	3,818	11350,059
4	51-65	4191	15420	15	58	0,0172	72,259	3364	14098524	4,060	17017,317
5	66-80	4131	19551	15	73	0,0137	56,589	5329	22014099	4,290	17723,888

Il valore di interesse cade nell'osservazione  $\frac{n+1}{2} = \frac{19552}{2} = 9776$ . La mediana cade quindi nella terza classe, ed è pari a

$$Me = x_l + S \left( \left( \frac{n+1}{2} - C_{i-1} \right) \frac{a_i}{n_i} \right) - 1 = 41 + S \left( (9776 - 8256) \frac{10}{2973} \right) - 1 = 46$$

dove  $S(x)$  indica il primo numero intero più grande di  $x$ .

I quartili si ottengono come

$$q_1 = x_{\left(\frac{n+1}{4}\right)} = x_{(4888)} = 18 + S \left( (4888 - 0) \frac{13}{5782} \right) - 1 = 28$$

$$q_2 = Me$$

$$q_3 = x_{(14664)} = 41 + S \left( (14664 - 11229) \frac{15}{4191} \right) - 1 = 53$$

La moda è la modalità con frequenza maggiore, quindi la classe "18-30". Non è rappresentativa.

La media aritmetica è calcolata come funzione dei valori centrali delle classi

$$M_1 = \frac{1}{n} (n_1 c_1 + \dots + n_k c_k) = \frac{1}{n} \sum_{i=1}^k n_i c_i = \frac{1}{19551} (906507,5) = 46,366$$

La media armonica è calcolata come funzione dei valori centrali delle classi

$$M_{-1} = \frac{1}{\frac{1}{n} \left( n_1 \frac{1}{c_1} + \dots + n_k \frac{1}{c_k} \right)} = \frac{n}{\sum_{i=1}^k n_i \frac{1}{c_i}} = \frac{19551}{504,795} = 38,731$$

La media quadratica è calcolata come funzione dei valori centrali delle classi

$$M_2 = \sqrt{\frac{1}{n} (n_1 c_1^2 + \dots + n_k c_k^2)} = \sqrt{\frac{1}{n} \sum_{i=1}^k n_i c_i^2} = \sqrt{\frac{1}{19551} 48715766,75} = 49,917$$

La media geometrica è calcolata come funzione dei valori centrali delle classi

$$M_0 = \sqrt[n]{c_1^{n_1} \times \dots \times c_k^{n_k}} = \exp \left[ \frac{1}{n} \sum_{i=1}^k n_i \log(c_i) \right] = \exp \left[ \frac{1}{19551} 73297,794 \right] = 42,481$$

## ESERCIZIO 5

Si considerano ora le 7768 famiglie. Per ognuna di esse si osserva il reddito disponibile netto in migliaia di Euro. I dati sono stati raggruppati in classi. Si individuino mediana, quartili, e media aritmetica. Si verifichi la prima proprietà della media aritmetica.

i	$x_l - x_u$	$n_i$
1	0-50	5782
2	50-100	2474
3	100-150	2973
4	150-350	4191
5	350-850	4131

## SOLUZIONE

i	$x_l -  x_u$	$n_i$	$c_i$	$a_i$	$C_i$	$c_i - M_1$	$(c_i - M_1) \times n_i$
1	0- 50	5782	25	50	5782	-191,259	-1105858,144
2	50- 100	2474	75	50	8256	-141,259	-349474,170
3	100- 150	2973	125	50	11229	-91,259	-271312,291
4	150- 350	4191	250	200	15420	33,741	141409,540
5	350- 850	4131	600	500	19551	383,741	1585235,065

Il valore di interesse cade nell'osservazione  $\frac{n+1}{2} = \frac{19552}{2} = 9776$ . La mediana cade quindi nella terza classe, ed è pari a

$$Me = x_l + \left(\frac{n}{2} - C_{i-1}\right) \times \frac{a_i}{n_i} = 100 + (9775,5 - 8256) \times \frac{50}{2973} = 125,555$$

I quartili si ottengono come

$$q_1 = x_{\left(\frac{n+1}{4}\right)} = x_{(4888)} = x_l + \left(\frac{n+1}{4} - C_{i-1} - \frac{1}{2}\right) \times \frac{a_i}{n_i} = 0 + \left(4888 - 0 - \frac{1}{2}\right) \times \frac{50}{5782} = 42,265$$

$$q_2 = Me$$

$$q_3 = x_{(14664)} = 150 + \left(14664 - 11229 - \frac{1}{2}\right) \times \frac{200}{4191} = 313,899$$

La media aritmetica è calcolata come funzione dei valori centrali delle classi

$$M_1 = \frac{1}{n} (n_1 c_1 + \dots + n_k c_k) = \frac{1}{n} \sum_{i=1}^k n_i c_i = \frac{1}{19551} (4228075) = 216,259$$

La proprima proprietà della media aritmetica, applicata in questo contesto, dice che

$$\sum_{i=1}^k n_i (c_i - M_1) = 0$$

la somma degli scarti della media è uguale a zero, è valida (NB è valida facendo i conti su qualsiasi software e senza portare avanti errori di approssimazione, se fatto con la calcolatrice e approssimando di volta in volta verrà un risultato simile a 0).