

Esercitazione 2 del corso di Statistica (parte 1)

Dott.ssa Paola Costantini

11 Ottobre 2011

Dataset Studenti

<i>N</i>	<i>SESSO</i>	<i>ETA'</i>	<i>PESO</i>	<i>ALTEZZA</i>	<i>DIPLOMA</i>	<i>COMPONENTI</i>	<i>OCCHIALI</i>	<i>FUMO</i>
1	0	20,6	65	180	Ist.Tecnico	6	0	1
2	0	20,2	75	180	Liceo	4	0	0
3	0	20,3	60	173	Ist.Tecnico	4	1	0
4	0	23,9	93	193	Liceo	8	0	1
5	0	21,4	66	164	Ist.Tecnico	5	0	0
6	0	25	84	188	Ist.Tecnico	4	0	0
7	0	20,8	67	175	Altro dipl.	4	0	1
8	0	20,6	89	170	Liceo	3	1	0
9	0	27,1	71	180	Liceo	1	0	1
10	0	23,3	63	170	Liceo	4	0	0
11	1	20,5	51	161	Ist.Tecnico	4	0	1
12	1	19,1	58	167	Ist.Tecnico	5	1	1
13	1	22,1	67	165	Altro dipl.	5	1	1
14	1	21,8	51	153	Ist.Tecnico	4	0	0
15	1	19,2	60	170	Ist.Tecnico	5	1	1
16	1	20,8	55	165	Liceo	4	1	1
17	1	21	55	155	Liceo	5	1	0
18	1	20,9	58	170	Liceo	5	1	1
19	1	22,7	76	170	Liceo	6	1	0
20	1	21	55	165	Liceo	7	0	0

Esercizio 1 A partire dalla distribuzione di frequenza del carattere Altezza, calcolare la media aritmetica, la media troncata al 20% e la moda.

x_i	n_i
153	1
155	1
161	1
164	1
165	3
167	1
170	5
173	1
175	1
180	3
188	1
193	1
tot	20

Media aritmetica

$$\mu = \frac{\sum_{i=1}^n x_i \times n_i}{n} = \frac{153 \times 1 + 155 \times 1 + 161 \times 1 + \dots + 193 \times 1}{20} = 170,7$$

Media troncata

Nel nostro esercizio, la media troncata sarà ottenuta escludendo i due valori più grandi e i due più piccoli.

$$Mt_{20\%} = \frac{\sum_{i=(0,1n)+1}^{n-(0,1n)} X_i}{n - (0,2n)}$$

$$Mt_{20\%} = \frac{\sum_{i=3}^{18} X_i}{16} = \frac{2725}{16} = 170,3$$

La media troncata elimina l'influenza di valori anomali.

Moda

$$Mo_1 = 170$$

Esercizio n. 2

Calcolare media e moda della variabile Altezza ripartita in 3 classi equifrequenti.

$$N/3 = 6,67$$

Per una distribuzione in classi di frequenza, la media si calcola:

C_i	n_i	f_i	N_i	F_i	\hat{C}_i	a_i	d_i
$C_1 = [153; 165]$	7	0,35	7	0,35	159	12	0,029
$C_2 =] 165; 173]$	7	0,35	14	0,70	169	8	0,043
$C_3 =] 173; 193]$	6	0,30	20	1	183	20	0,015
Totali	20	1,00					

Media

$$\mu_x = \frac{\sum_{i=1}^k \hat{x}_i \cdot n_i}{n}$$

dove

$$\hat{x}_i = \frac{x_{i-1} + x_i}{2}$$

$$\mu_x = \frac{\sum_{i=1}^k (159 \times 7 + 169 \times 7 + 183 \times 6)}{20} = 169,7$$

Classe Modale = $C_2 =] 165; 173]$ = la classe modale è quella cui corrisponde la densità di frequenza più elevata.

Esercizio 3

Verificare la proprietà associativa della media aritmetica

- 1) Verificare la proprietà associativa della media aritmetica del carattere ALTEZZA rispetto al carattere GENERE.

Secondo il raggruppamento indotto dalle modalità del carattere GENERE, verifichiamo che la media generale dell'ALTEZZA è uguale alla media delle medie di gruppo.

MASCHI	ALTEZZA
0	180
0	180
0	173
0	193
0	164
0	188
0	175
0	170
0	180
0	170

FEMMINE	ALTEZZA
1	161
1	167
1	165
1	153
1	170
1	165
1	156
1	170
1	170
1	165

$$\mu_M = \frac{\sum_{i=1}^n x_i \times n_i}{n} = \frac{164 \times 1 + 170 \times 2 + 173 \times 1 + \dots + 187 \times 1}{20} = 177,3$$

$$\mu_F = \frac{\sum_{i=1}^n x_i \times n_i}{n} = \frac{153 \times 1 + 156 \times 1 + 161 \times 1 + \dots + 170 \times 3}{20} = 164,2$$

Media uomini = 177,3

Media donne 164,2

Media totale 170,7

La proprietà associativa è verificata dalla seguente uguaglianza:

$$\frac{\sum_{i=1}^3 \mu_i n_i}{n} = \frac{177,3 \times 10 + 164,2 \times 10}{20} = 170,7$$

Esercizio n 4

Calcolare la mediana e i quartili del carattere ALTEZZA.

Mediana

Consideriamo la successione ordinata di valori del carattere:

153 156 161 164 165 165 165 167 170 170 170 170 173 175 180 180 180 188 193

Essendo N pari la mediana è ottenuta come:

$$Me = \frac{x_{\left(\frac{N}{2}\right)} + x_{\left(\frac{N}{2}+1\right)}}{2} = \frac{10+11}{2} = 10,5(\text{posizione})$$

per cui la mediana sarà pari a $Me_1 = \frac{170+170}{2} = 170$

Quartili

Per il calcolo del primo quartile consideriamo la mediana delle prime 20 osservazioni:

Calcolo del primo quartile:

153 156 161 164 165 165 165 167 170 170

$$Q_1 = \frac{x_{\left(\frac{N}{2}\right)} + x_{\left(\frac{N}{2}+1\right)}}{2} = \frac{5+6}{2} = 5,5(\text{posizione})$$

per cui il primo quartile sarà pari a $Q_1 = \frac{165+165}{2} = 165$

In ogni caso Q_1 si individua scorrendo la successione ordinata della X e corrisponde al primo valore x^* tale che $F(x^*) \geq 0,25$

Per il calcolo del terzo quartile consideriamo la mediana delle prime 20 osservazioni:

Calcolo del terzo quartile:

170 170 170 173 175 180 180 180 188 193

$$Q_3 = \frac{x_{\left(\frac{N}{2}\right)} + x_{\left(\frac{N}{2}+1\right)}}{2} = \frac{5+6}{2} = 5,5(\text{posizione})$$

per cui il terzo quartile sarà pari a $Q_3 = \frac{175+180}{2} = 177,5$

In ogni caso Q_3 si individua scorrendo la successione ordinata della X e corrisponde al primo valore x^* tale che $F(x^*) \geq 0,75$

x_i	n_i	f_i	F_i	
153	1	0,05	0,05	
155	1	0,05	0,1	
161	1	0,05	0,15	
164	1	0,05	0,2	
165	3	0,15	0,35	1°quartile
167	1	0,05	0,4	
170	5	0,25	0,65	mediana
173	1	0,05	0,7	
175	1	0,05	0,75	} 3°quartile
180	3	0,15	0,9	
188	1	0,05	0,95	
193	1	0,05	1	
tot	20	1		

Esercizio 5

Calcolare la media interquartile utilizzando i valori di cui all'esercizio 4.

$$\mu_1 = \frac{Q_3 - Q_1}{2} = \frac{177,5 - 165}{2} = 171,25$$

Confrontando la media interquartile con la mediana, la distribuzione si dice obliqua a destra se la mediana risulta più piccola rispetto alla media interquartile, e obliqua a sinistra se invece è la media interquartile a precedere la mediana.

Nel nostro caso la mediana è pari a 170, la media interquartile è pari a 171,25, per cui concludiamo che la distribuzione del carattere ALTEZZA è obliqua a destra.

Esercizio 6

Data la seguente tabella:

ESAME	VOTO	CREDITI
ESAME 1	18	6
ESAME 2	20	4
ESAME 3	25	12
		TOTALE CREDITI: 22

Calcolare la media ponderata.

Soluzione

Per media ponderata si intende la sommatoria del prodotto tra il voto di ciascun esame e il credito ad esso associato.

$$\frac{(18 \times 6) + (20 \times 4) + (25 \times 12)}{22} = \frac{108 + 80 + 300}{22} = \frac{498}{22} = 22,18$$