STATISTICA (I modulo - Statistica Descrittiva)
Soluzione esercitazione 3

Esercizio A.

1. Automobile

$$\mu = \frac{23 + 32 + \ldots + 31}{12} = 32{,}167$$
$$\mu_a = \frac{12}{1/23 + 1/32 + \ldots + 1/31} = 30{,}724$$
$$\mu_g = \sqrt[12]{23 \times 32 \times \ldots \times 31} = \exp\left(\frac{\log(23) + \log(32) + \ldots + \log(31)}{12}\right) = 31{,}442$$

Metropolitana

$$\mu = \frac{22 + 24 + \ldots + 24}{12} = 27{,}833$$
$$\mu_a = \frac{12}{1/22 + 1/24 + \ldots + 1/24} = 27{,}077$$
$$\mu_g = \sqrt[12]{22 \times 24 \times \ldots \times 24} = \exp\left(\frac{\log(22) + \log(24) + \ldots + \log(24)}{12}\right) = 27{,}448$$

Come si può vedere le medie calcolate soddisfano la nota relazione $\mu_a \leq \mu_g \leq \mu$.

2. Le distribuzioni disaggregate ordinate in senso non decrescente e la corrispondente distribuzione delle frequenze relative cumulate sono le seguenti:

Unità (i)	1	2	3	4	5	6	7	8	9	10	11	12
Automobile	21	23	28	29	30	31	32	33	34	36	44	45
Metropolitana	22	22	24	24	24	26	28	31	31	32	33	37
F_i	0,0833	0,1667	0,2500	0,3333	0,4167	0,5000	0,5833	0,6667	0,7500	0,8333	0,9167	1,0000

Automobile

mediana: l'unità per cui vale $(i-1)/N \leq 1/2 < i/N$ è l'unità $i = 7$. Essendo $(i-1)/N = 0{,}5$, $m = q_2 = (y_6 + y_7)/2 = (31 + 32)/2 = 31{,}5$.

primo quartile: l'unità per cui vale $(i-1)/N \leq 1/4 < i/N$ è l'unità $i = 4$. Essendo $(i-1)/N = 0{,}25$, $q_1 = (y_3 + y_4)/2 = (28 + 29)/2 = 28{,}5$.

terzo quartile: l'unità per cui vale $(i-1)/N \leq 3/4 < i/N$ è l'unità $i = 10$. Essendo $(i-1)/N = 0{,}75$, $q_3 = (y_9 + y_{10})/2 = (34 + 36)/2 = 35$.

Metropolitana

mediana: l'unità per cui vale $(i-1)/N \leq 1/2 < i/N$ è l'unità $i = 7$. Essendo $(i-1)/N = 0{,}5$, $m = q_2 = (y_6 + y_7)/2 = (26 + 28)/2 = 27$.

primo quartile: l'unità per cui vale $(i-1)/N \leq 1/4 < i/N$ è l'unità $i = 4$. Essendo $(i-1)/N = 0{,}25$, $q_1 = (y_3 + y_4)/2 = (24 + 24)/2 = 24$.

terzo quartile: l'unità per cui vale $(i-1)/N \leq 3/4 < i/N$ è l'unità $i = 10$. Essendo $(i-1)/N = 0{,}75$, $q_3 = (y_9 + y_{10})/2 = (31 + 32)/2 = 31{,}5$.

3. Sia dalle medie analitiche (μ, μ_a, μ_g) sia da quelle di posizione (m, q_1, q_3) emerge che il tempo necessario per recarsi al lavoro è inferiore se viene utilizzata la metropolitana.

Esercizio B. Indichiamo con x_i lo stipendio dell'i-esimo dipendente. La media aritmetica degli stipendi è data da $\mu = \dfrac{\sum_{i=1}^{10} x_i}{10}$.

1. Se lo stipendio di un dipendente, diciamo il primo, viene aumentato di 50 €, quindi $x_1 \to x_1 + 50$, allora la media aritmetica diventa
$$\frac{(x_1+50)+x_2+\ldots+x_{10}}{10} = \frac{50}{10} + \frac{x_1+x_2+\ldots+x_{10}}{10} = 5+\mu$$
Se, invece, tutti gli stipendi sono aumentati di 50 €, quindi $x_i \to x_i + 50$ per ogni $i = 1,\ldots,10$, allora la media aritmetica diventa
$$\frac{(x_1+50)+(x_2+50)+\ldots+(x_{10}+50)}{10} = 50+\mu$$
2. Nel caso della media armonica non è possibile derivare una formula esplicita come nel caso della media aritmetica. L'unica cosa che possiamo dire è che la nuova media sarà più elevata del precedente valore.
3. Se lo stipendio di tutti i dipendenti è aumentato dell'1%, quindi $x_i \to 1{,}01x_i$ per ogni $i = 1,\ldots,10$, allora la media aritmetica diventa
$$\frac{1{,}01x_1+1{,}01x_2+\ldots+1{,}01x_{10}}{10} = 1{,}01\mu$$
mentre per la media armonica si ha
$$\frac{10}{1/(1{,}01x_1)+1/(1{,}01x_2)+\ldots+1/(1{,}01x_{10})} = \frac{10}{1/1{,}01(1/x_1+1/x_2+\ldots+1/x_{10})} = 1{,}01\mu_a$$

Esercizio C.

Reddito (x_i)	n_i	$x_i-\mu$	$(x_i-\mu)n_i$	$(x_i-\mu)^2n_i$	$(x_i-2)^2n_i$
1,5	3	-0,80	-2,40	1,92	0,75
2,0	4	-0,30	-1,20	0,36	0,00
3,5	3	1,20	3,60	4,32	6,75
Totale	10		0,00	6,60	7,50

1. La media aritmetica è pari a $\mu = (1{,}5\times3+2\times4+3{,}5\times3)/10 = 2{,}3$. Come si vede dalla tabella precedente la somma degli scarti da μ è pari a zero.
2. La somma dei quadrati degli scarti dalla media aritmetica è pari a 6,60.
3. La somma dei quadrati degli scarti dal valore 2 è pari a 7,50. Come ci si attendeva tale valore è maggiore di quello ottenuto al punto precedente.

Esercizio D.

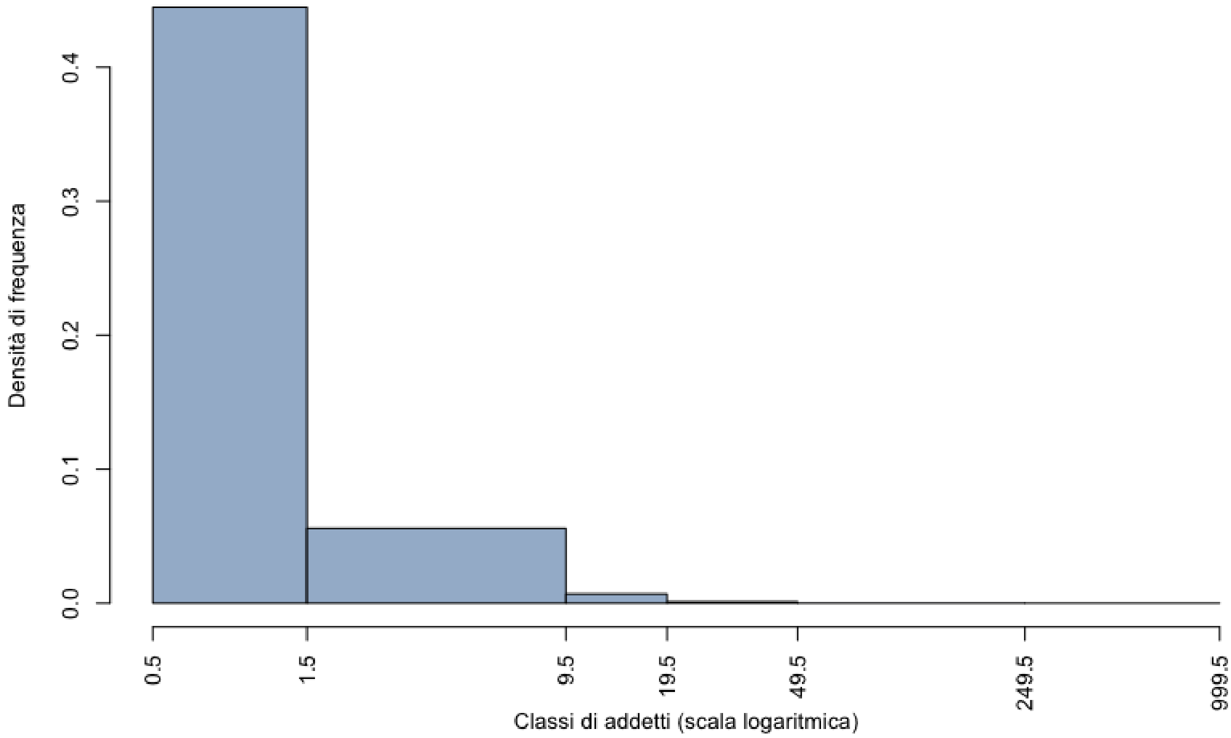
Classi	c_{i-1}	c_i	n_i	f_i	F_i	X_i	μ_i	x_i
0–20	-0,5	20,5	4	0,1333	0,1333	74,305	18,576	10,0
21–30	20,5	30,5	15	0,5000	0,6333	377,055	25,137	25,5
31–40	30,5	40,5	8	0,2667	0,9000	279,476	34,934	35,5
41–50	40,5	50,5	3	0,1000	1,0000	141,085	47,028	45,5

1. Per il calcolo della media geometrica si possono utilizzare i valori centrali di classe x_i oppure, dato che in questo caso sono noti, i valori medi di classe μ_i. Quindi:
$$\mu_g = 18{,}576^{0{,}1333} \times 25{,}137^{0{,}5} \times 34{,}934^{0{,}2667} \times 47{,}028^{0{,}1} = 28{,}062$$
oppure
$$\mu_g = \exp\left(\log(18{,}576)0{,}1333 + \log(25{,}137)0{,}5 + \log(34{,}934)0{,}2667 + \log(47{,}028)0{,}1\right) = 28{,}062$$
2. La classe mediana è la classe 21-30, così come quella in cui cade il primo quartile, mentre il terzo quartile appartiene alla classe 31–40. Quindi:
$$m = q_2 = 20{,}5 + \frac{0{,}5-0{,}1333}{0{,}5}(30{,}5-20{,}5) = 27{,}834$$
$$q_1 = 20{,}5 + \frac{0{,}25-0{,}1333}{0{,}5}(30{,}5-20{,}5) = 22{,}834$$
$$q_3 = 30{,}5 + \frac{0{,}75-0{,}6333}{0{,}2667}(40{,}5-30{,}5) = 34{,}876$$
3. Dalla distribuzione disaggregata del fatturato, dopo aver ordinato i termini in senso non decrescente, essendo $N = 30$ un numero pari occorre individuare le osservazioni che occupano la 15-esima e la 16-esima posizione, cioè 26,065 e 26,500, da cui la mediana $m = (26{,}065 + 26{,}500)/2 = 26{,}282$.
Dal confronto con la mediana calcolata in precedenza possiamo vedere che questa si discosta leggermente dal valore vero.

Esercizio E.

Classi addetti	c_{i-1}	c_i	x_i	n_i	f_i	F_i	d_i	$h_i = f_i/d_i$	X_i	μ_i
1	0,5	1,5	1,0	493669	0,4448	0,4448	1	0,44479	493669	1,00
2-9	1,5	9,5	5,5	495710	0,4466	0,8914	8	0,05583	1849025	3,73
10-19	9,5	19,5	14,5	75805	0,0683	0,9597	10	0,00683	1004476	13,25
20-49	19,5	49,5	34,5	31370	0,0283	0,9880	30	0,00094	928305	29,59
50-249	49,5	249,5	149,5	11796	0,0106	0,9986	200	0,00005	1116329	94,64
250-999	249,5	999,5	624,5	1543	0,0014	1,0000	750	0,00000	1160258	751,95
Totale				1109893	1,0000				6552062	

dove $x_i = (c_{i-1} + c_i)/2$ è il valore centrale di classe, mentre $\mu_i = X_i/n_i$ è il valore medio di classe.

1. 

2. Il numero medio di addetti sotto l'ipotesi di uniforme distribuzione all'interno delle classi è pari a

$$\mu = 1 \times 0{,}4448 + 5{,}5 \times 0{,}4466 + \ldots + 624{,}5 \times 0{,}0014 = 7{,}3238$$

mentre se calcolato in base alla distribuzione di quantità si ottiene:

$$\mu = 1 \times 0{,}4448 + 3{,}73 \times 0{,}4466 + \ldots + 751{,}95 \times 0{,}0014 = 5{,}9033$$

oppure $6552062/1109893 = 5{,}9033$. Il valore calcolato risulta sensibilmente differente; ciò è dovuto in gran parte alla differenza tra valore centrale e valore medio per le classi 20–49 e successive.

3. I quartili richiesti sono i seguenti:

$$m = 1{,}5 + \frac{0{,}5 - 0{,}4448}{0{,}4466}(9{,}5 - 1{,}5) = 2{,}4888$$ in quanto la mediana appartiene alla classe 2–9,

$$q_1 = 1$$ in quanto il primo quartile appartiene alla classe 1,

$$q_3 = 1{,}5 + \frac{0{,}75 - 0{,}4448}{0{,}4466}(9{,}5 - 1{,}5) = 6{,}9671$$ in quanto il terzo quartile appartiene alla classe 2–9.

4. La classe modale è la classe per la più elevata è la densità, cioè la prima classe.

Esercizio F.

1.

Anno	2001	2002	2003	2004	2005	2006	2007
$v_t \times 100$	3,20	3,50	3,60	4,60	2,40	3,10	4,50
$i_t \times 100$	103,20	103,50	103,60	104,60	102,40	103,10	104,50

La variazione percentuale media dei prezzi dal 2003 al 2007 è pari a:

$$\bar{v}_{2007|2003} = \sqrt[4]{104{,}60 \times 102{,}40 \times 103{,}10 \times 104{,}50} - 100 = 13{,}6458 - 100 = 3{,}6458\%$$

2. Dal momento che è noto il valore $a_{2003} = 3{,}25$, possiamo calcolare per $t > 2003$ $a_t = a_{t-1} i_t$, mentre per $t < 2003$ utilizziamo $a_t = a_{t+1}/i_{t+1}$, dove i_t sono i numeri indice a base mobile espressi in base unitaria. La serie dei prezzi dell'olio d'oliva dal 2001 al 2007 è pertanto la seguente:

Anno	2001	2002	2003	2004	2005	2006	2007
a_t	3,03	3,14	3,25	3,40	3,48	3,59	3,75

Esercizio G.

Titolo	Inizio mese		Fine mese		Spesa teorica al tempo t	Spesa effettiva al tempo b
	p_{ib}	q_{ib}	p_{it}	q_{it}	$p_{it} q_{ib}$	$p_{ib} q_{ib}$
A	2,65	500	2,27	400	1135,00	1325,00
B	3,87	400	3,90	400	1560,00	1548,00
C	4,31	200	3,88	100	776,00	862,00
D	2,92	300	2,98	400	894,00	876,00
E	4,01	200	4,25	300	850,00	802,00
					5215	5413

L'indice di Laspeyres si calcola come segue:

$$I_L = \frac{5215}{5413} \times 100 = 96{,}342$$

Quindi, la variazione percentuale media per il portafoglio di titoli azionari nel mese preso in considerazione è stata pari a $96{,}342 - 100 = -3{,}658\%$.