

ESERCIZIO 1

Le conchiglie luminescenti, molto rare ovunque, sono abbondanti in Nuova Zelanda. Alcuni biologi hanno raccolto un campione di conchiglie per studiare le differenze tra conchiglie vive e morte. Hanno misurato le lunghezze (mm) delle conchiglie che sono riportate qui sotto.

VIVE

4,2 4,8 4,8 4,9 6,2 6,4 6,2 6,8 8,6 8,1 10,3 12 12,1
 12,1 12,1 14 14,2 14,5 14,6 16,8 16,9 16,6 18,1 18,7 18 20
 20 20,1 20,7 20,1 20,3 20,5 24,2 24,5 22,8 22,3 22,1 20,9
 4,1 4,1

MORTE

12,4 12,8 16,2 16,6 16,8 16,7 18,3 18,4 18,1 18,9 18,9 18,9 20,1
 20,1 20,2 20,5 20,4 20,4 20,6 22,5 22,6 22,7 22,3 24,9 24,6 24,7
 24,3 22,4 24,9 10,8

- a) per le conchiglie vive, costruite classi di ampiezza pari a 2,5 mm a partire dal valore 3,75. Disegnate l'istogramma riportando sull'asse delle ascisse le classi ottenute e sull'asse delle ordinate le rispettive frequenze relative.
- b) per le conchiglie morte, costruite classi di ampiezza pari a 2,0 mm a partire dal valore 9. Disegnate l'istogramma riportando sull'asse delle ascisse le classi ottenute e sull'asse delle ordinate le rispettive frequenze relative.
- c) riportate sull'ultimo foglio entrambi gli istogrammi. 2+2
- d) descrivete la forma dei due istogrammi delle lunghezze delle conchiglie. 3

Vive	bimodale	
Morte	unimodale approx normale	

- e) calcolate i cinque numeri di sintesi, la media e la deviazione standard per entrambi i campioni di lunghezze. 1,5 + 1,5

Vive	Variable	N	N*	Percent	Mean	SE Mean	StDev	CoefVar	Minimum	Q1
	morte	30	0	100	19,733	0,673	3,686	18,68	10,800	17,775
	vive	40	0	100	14,22	1,04	6,59	46,32	4,10	7,13
Morte	Variable	Median	Q3	Maximum	IQR	Skewness	Kurtosis			
	morte	20,150	22,525	24,900	4,750	-0,64	0,18			
	vive	14,55	20,10	24,50	12,98	-0,20	-1,35			

- f) Quali indici di posizione centrale e di variabilità usereste per descrivere i due

campioni, ossia le due distribuzioni? Spiegate le vostre scelte.

1,5+1,5

Vive	i 5 numeri di sintesi	
Morte	media e dev st	

g) A questo punto siete in grado di confrontare le due distribuzioni e di spiegare le differenze osservate tra le misure delle lunghezze delle conchiglie vive e morte.

3

vive: maggior dispersione con valori da 4 a 24mm	
morte: minor dispersione con valori da 11 a 25	

h) usando la differenza interquartile, identificate qualunque potenziale outlier per la distribuzione di lunghezze delle conchiglie morte. Riportate i calcoli.

3

$\begin{aligned} \text{IQR morte } 4,75 \times 1,75 = 7,125 &\rightarrow 10,8 + 7,1 = 17,9 \quad Q1 = 17,7 \\ &\rightarrow \underline{24,9} + 7,1 = 32 \quad Q3 = 22,52 \\ &\rightarrow \underline{24,7 ; 24,6 ; 24,3} \end{aligned}$	
---	--

i) disegnando il plot dei quantili normali per la variabile lunghezza delle conchiglie morte (senza riportarlo sul foglio) si può affermare che la distribuzione di questa variabile è approssimativamente normale? Giustificate la vostra risposta.

1+1

Si perché i dati si collocano approx lungo una retta	
--	--

l) calcolate il percentile corrispondente alla lunghezza 20,3 mm per le conchiglie vive. Spiegate cosa indica tale valore.

1+1

80° ; ossia l'80% dei valori della distr. cade alla sua sinistra.

ESERCIZIO

1

Se una distribuzione normale ha media pari a 200 e deviazione standard pari a 20, trovare quel valore K tale che la proporzione di valori minori di K sia pari a 0,975.

a b c d e f

- a) 239
- b) 204
- c) 238
- d) 220
- e) 239
- f) 250

ESERCIZIO

2

Supponiamo che vogliate estrarre un campione per determinare l'altezza media degli studenti dell'università "La Sapienza".

Quale dei seguenti metodi di campionamento non introduce una distorsione? a b c d

- a) Poiché siete uno studente della facoltà di lettere, scegliete in modo casuale un campione di studenti di questa facoltà.
- b) Usate le tavole di numeri casuali per scegliere il campione di studenti in base al loro numero di matricola.
- c) Lanciate due dadi per scegliere un campione tra i vostri amici
- d) Nessuno dei metodi di campionamento presenterà una distorsione

ESERCIZIO

1,5

Per ciascuna delle variabili che seguono:

- | | | | | |
|--|---|---|---|---|
| 1) tipo di habitat | a | b | c | d |
| 2) intensità di un terremoto (scala Richter) | a | b | c | d |
| 3) posizione nella graduatoria di una gara: 1°, 10°, ecc... | a | b | c | d |
| 4) presenza o assenza di miopia | a | b | c | d |
| 5) altitudine del luogo di nascita | a | b | c | d |
| 6) Millimetri di pioggia rilevati giornalmente in un certo luogo | a | b | c | d |

indicare se sono:

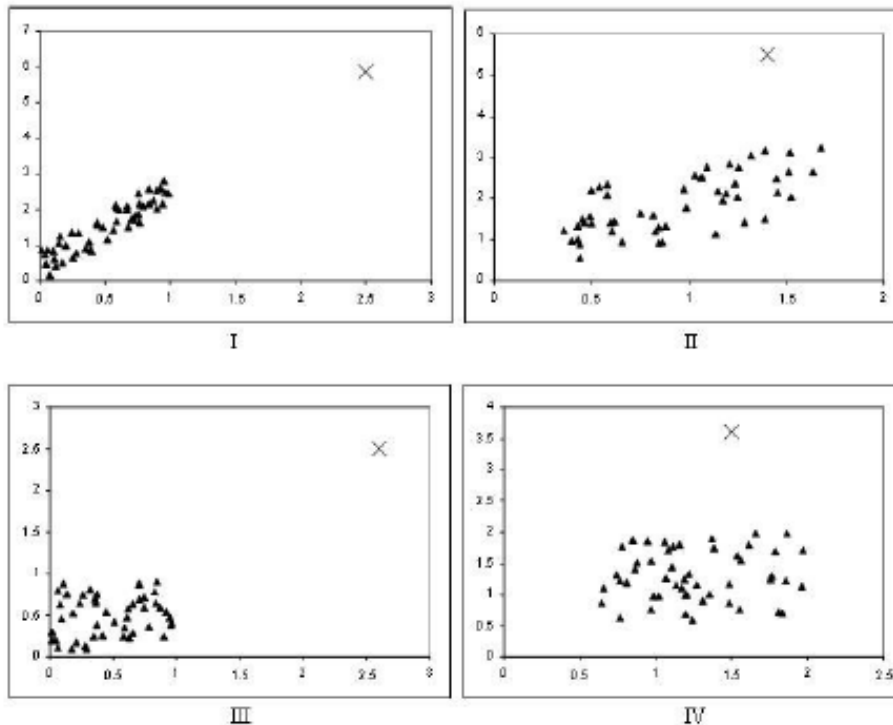
- (a) Variabile qualitativa ordinata
- (b) Variabile qualitativa categorica
- (c) Variabile quantitativa continua
- (d) Variabile quantitativa discreta

ESERCIZIO

2,5

Considerate i 4 grafici di dispersione e i relativi valori dei coefficienti di correlazione riportati qui sotto. Se il punto indicato con una X viene eliminato, quale delle seguenti affermazioni è vera?

a b c d



I $r=0.90$; II $r=0.80$; III $r=0.20$; IV $r=0.10$;

- a) la correlazione di (I) decresce, la correlazione di (II) non cambia
- b) la correlazione di (III) cresce, la correlazione di (IV) cresce
- c) la correlazione di (I) non cambia, la correlazione di (III) decresce
- d) la correlazione di (II) cresce, la correlazione di (III) cresce

ESERCIZIO

2

Individua, tra quelle proposte, le affermazioni false:

a b c d e f

- a. Se c'è maggiore variabilità la curva è più appuntita
- b. Se la distribuzione è asimmetrica positiva la moda è maggiore della media aritmetica
- c. Se la distribuzione è asimmetrica negativa, la mediana è maggiore della media
- d. Quando c'è una asimmetria positiva o una coda destra, esiste un maggior addensamento dei dati in corrispondenza di valori superiori alla media.
- e. Se la distribuzione è simmetrica le due code sono uguali
- f. Quando c'è una asimmetria positiva o una coda destra, esiste un maggior addensamento dei dati in corrispondenza di valori inferiori alla media

ESERCIZIO

1

Quale dei valori, se ne esiste uno, elencati sotto per un coefficiente di correlazione, indica una situazione dove più della metà della variabilità di una variabile può essere spiegata dalla variabilità dell'altra?

a b c d e

- a) $r = -.7$
- b) $r = .3$
- c) $r = -.9$
- d) $r = .6$
- e) $r = 1.0$

ES AGGIUNTIVI 3 CREDITI

g) spiegate se si può dire che la frequenza percentuale cumulata fino alla lunghezza 20,1 è maggiore per le conchiglie vive piuttosto che per quelle morte.

si	vive corrisponde al 77,5%
	morte.....50%

ESERCIZIO 5

2

Il coefficiente di correlazione tra i punteggi a un test neurologico e a un test sull'ansia ha un valore alto e positivo, quale (o quali se ve ne è più di una) delle seguenti affermazioni è vera?

a b c d

- a) quelli con un punteggio basso a un test tendono ad avere un punteggio alto nell'altro
- b) problemi neurologici causano ansia
- c) quelli con un punteggio basso a un test tendono ad avere un punteggio basso nell'altro
- d) non si possono fare previsioni significative a partire da un test sull'altro test

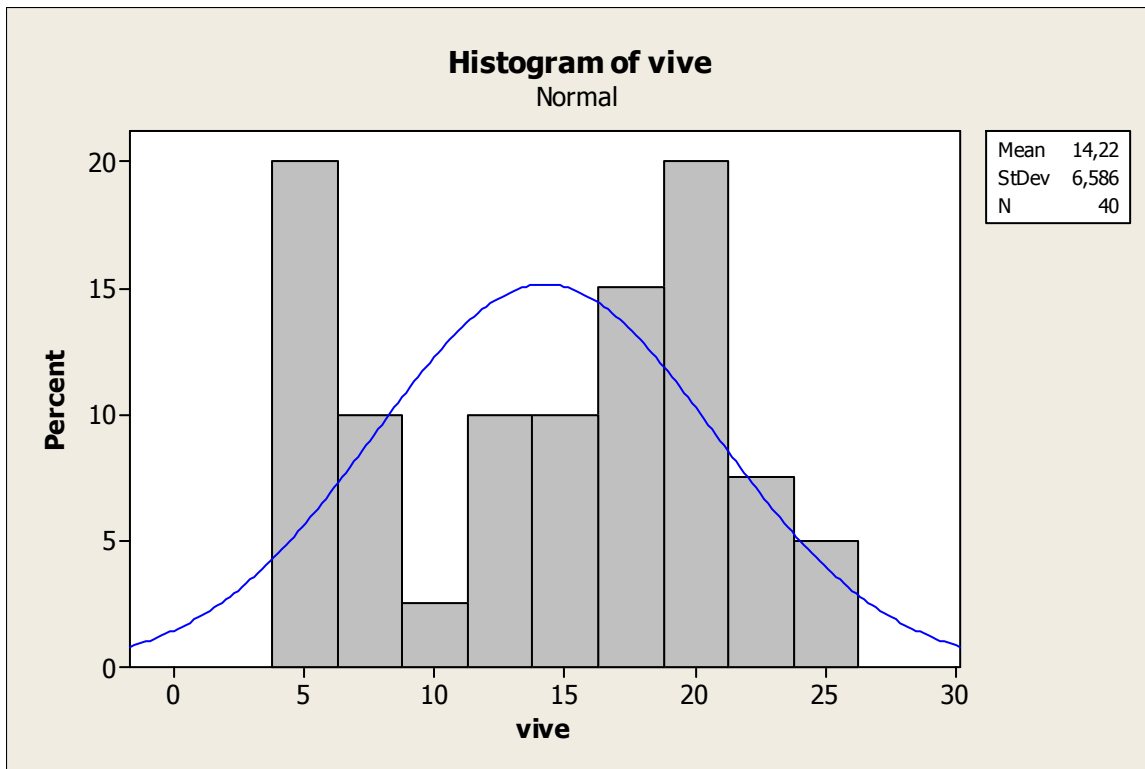
ESERCIZIO 6

2

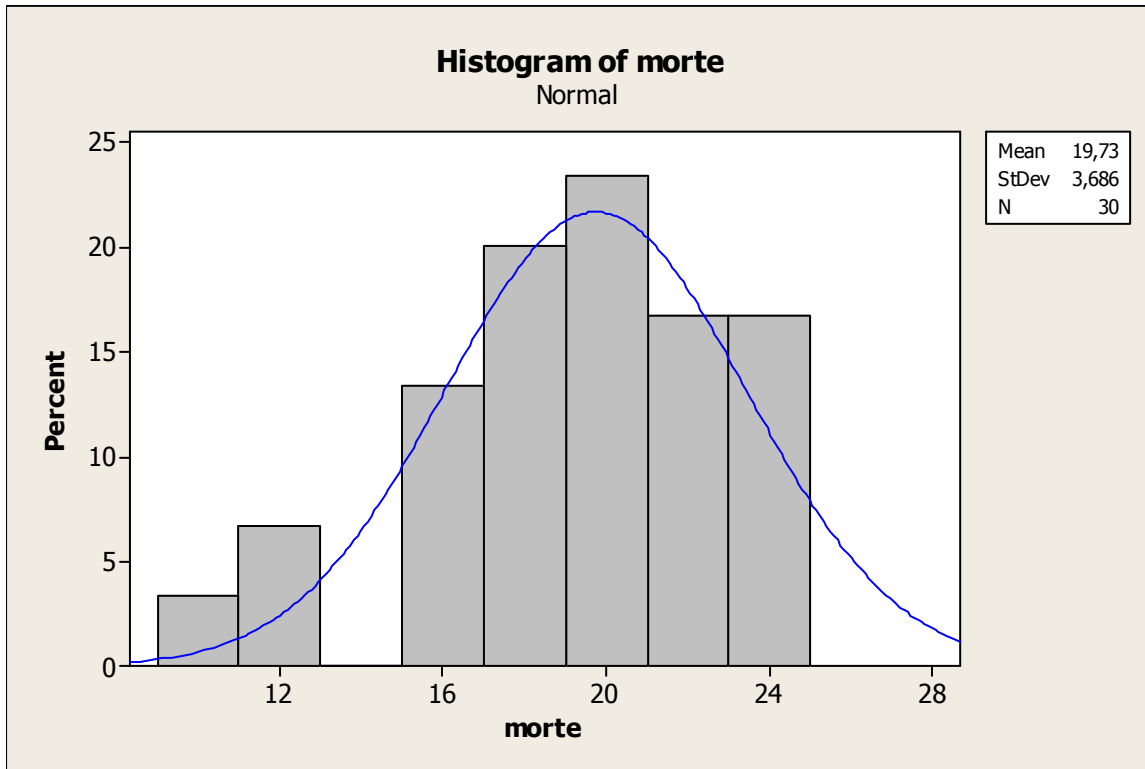
Si supponga che i punteggi di un test abbiano una distribuzione approssimativamente normale con media pari a 500 e deviazione standard pari a 100. Uno studente con un punteggio di 600 ha risposto meglio di quale percentuale di studenti partecipanti al test? a b c d e

- a) 20%
- b) 50%
- c) 57 %
- d) 84%
- e) 97.5%

SOLUZIONI



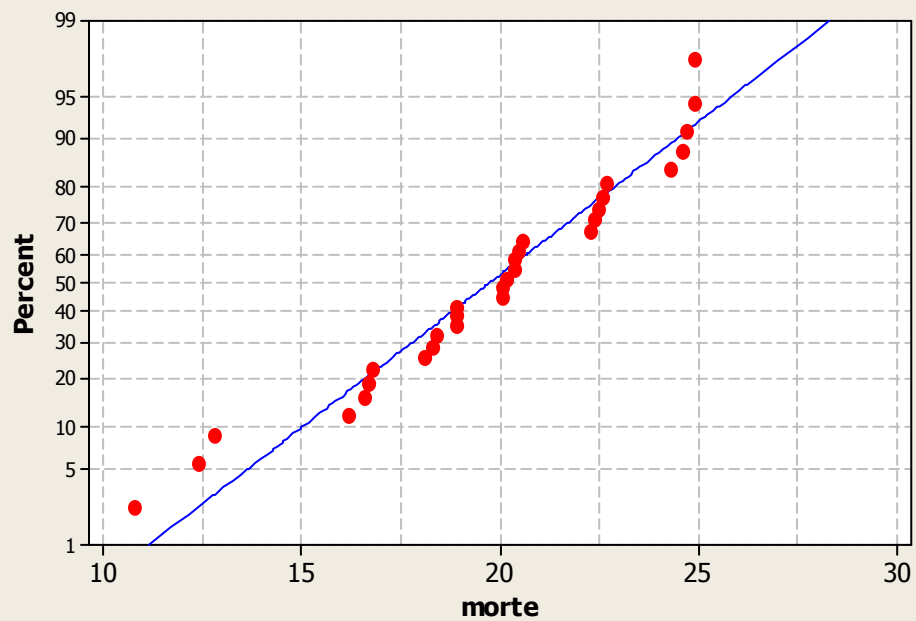
3,75-6,25 0,2
6,25-8,75 0,1
8,75-11,25 0,025
11,25-13,75 0,1
13,75-16,25 0,1
16,25-18,75 0,15
18,75-21,25 0,2
21,25-23,75 0,075
23,75-26,25 0,05



9-11 0,03
 11-13 0,066
 13-15 0
 15-17 0,13
 17-19 0,20
 19-21 0,23
 21-23 0,166
 23-25 0,166

Probability Plot of morte

Normal



Mean	19,73
StDev	3,686
N	30
AD	0,481
P-Value	0,216

