**CORREZIONI 1 A 27 –9—2012**

**ES 2**

i) per il campione di ali della I specie, costruite 7 classi di ampiezza pari a 2 a partire dal valore 11.0 e calcolate le frequenze relative percentuali (con una cifra decimale), per ogni classe, riportandole qui sotto.

11-13 6 .66  **2.0 PT**

13-15 20

15-17 0

17-19 13.3

19-21 40

21-23 6.6

23-25 13.33

ii) RIPORTARE SUGLI ASSI LE GRANDEZZE CHE RAPPRESENTANO **2 + 0.5 PT**



iii) Dite quali sono tutti gli altri grafici che si possono usare per descrivere questa distribuzione di dati, senza disegnarli**. 1.5 PT**

ramo-foglia, box-plot

iv) Calcolate la media e la deviazione standard per i dati di questo campione riportando una cifra decimale.

 **0.5 PT**

m=19 s=3.7

v) Supponete che questo campione di dati provenga da una distribuzione approssimativamente normale con media e deviazione standard pari a quelle calcolate sul campione (nel punto iv).

Usate le tavole e dite:

a) Qual è la percentuale di insetti con una lunghezza dell’ala maggiore di 21mm? **1.5 PT**

b) Quale valore della lunghezza dell’ala corrisponde al 60-mo percentile? **1.5 PT**

 a) 0.295

 b) 19.93

vi) spiegate cosa indica il valore corrispondente al 60-mo percentile **1 PT**

i**il 60% delle ali è inf a 20.2**

vi) Considerando i due campioni di lunghezze delle ali, quale ha maggior dispersione? Giustificate la vostra risposta.

Variable N N\* Mean StDev CoefVar Minimum Q1 Median Q3

C1 15 0 19.013 3.729 19.61 12.000 14.600 20.100 20.700

C2 11 0 20.00 9.54 47.70 7.00 12.00 20.00 29.00

Variable Maximum Skewness Kurtosis

C1 24.500 -0.42 -0.58

C2 32.00 -0.16 -1.81

**II campione più disperso 🡪 coeff var = 47.7 % 1.5 PT**

vii) Disegnate il grafico ramo-foglia per la lunghezza delle ali degli insetti della II specie **2.5 PT**

 2 0 77

 4 1 23

 5 1 5

(1) 2 0

 5 2 8899

 1 3 2

viii calcolate la misura di dispersione che ritenete più opportuna per descrivere questa distribuzione di dati (II specie), giustificando la vostra risposta

**9.54 = Dev. stand. (o varianza), i dati sono approx simmetrici 2 PT**

**ESERCIZIO 3**

Si vuole testare un nuovo farmaco contro l’asma. Allo studio partecipano 250 soggetti sani della stessa età e dello stesso sesso e con caratteristiche simili.

i) Che tipo di disegno degli esperimenti usereste? Spiegate, con precisione, i passi necessari per svolgere l’esperimento. Se volete potete disegnare un grafico.

**ESP COMPLETAMENTE RANDOMIZZATO**  **2.5 PT**

**Si considerano 2 trattamenti: farmaco e placebo. 125 Donne e125 uomini devono essere assegnati a caso ai trattamenti. ESP COMPLET: RANDOMIZ**

ii) Perché è necessario che i soggetti siano il più simile possibile tra loro?  **2 PT**

**PER EVITARE L’EFFETTO DI VARIABILI NASCOSTE CONFONDENTI**

iii) In questo esperimento quali sono la variabile esplicativa e la variabile di risposta?

V. esplicat: **FARMACO**  v. di risposta: **riduzione effetti asma 0.5 PT**

**ESERCIZIO 4**

I due grafici che seguono rappresentano le età, alla nascita del primo figlio, rispettivamente,

di un campione di madri e di un campione di padri.

 **i)** I due grafici sono simmetrici. **V F 0.5 PT**

 **ii)** Per le età dei padri il miglior indice di posizione centrale e il miglior indice

 di variabilità sono: **a b c d e 1.5 PT**

 (a) la media e il 1° e il 3° quartile

 (b) la media e la deviazione standard

 (c) la mediana e la deviazione standard

 (d) la moda e la deviazione standard

 (e) nessuna delle precedenti risposte



**ESERCIZIO 3**

Nella tabella che segue sono riportate le dosi (in mg) di un farmaco contro l’obesità e le relative diminuzioni di peso (kg) misurate in 8 pazienti dopo quattro mesi.

 dose(x) diminuzione(y)

8 12

20 24

 30 6

10 22

18 21

14 14

20 25

16 26

i) Calcolate a quale dose corrisponde il residuo positivo più piccolo nel modello di regressione lineare  **10 1.5 PT**

iii) Riportate le due definizioni, che avete studiato, del coefficiente di determinazione **3 PT**

 **variabilità spiegata**

 **r2 = ------------------------- = coefficiente di determinazione**

**variabilità totale**

 **R2 = quadrato del coeff di correlaz**

iv) Sul grafico di dispersione appare un’osservazione con il residuo più alto (in valore assoluto)..

 Si tratta di un outlier o di un’osservazione influente? Spiegate le vostre conclusioni.

**oss infl** eliminandola, cambierebbe profondamente il risultato (v FIG 2)

 **2,0 PT**



v) Si può dire che la retta di regressione lineare rappresenta un buon modello per i dati? Perché si o perché no, spiegate.

**R2 = 5,9% -🡪NO 1.5 PT**

vii) Disegnate il grafico dei residui della regressione (senza riportarlo sul compito) e dite a quale scopo viene usato **2 PT**

**Per stabilire la bontà del modello**