**Metodi Matematici e Informatici per la Biologia**

 **Appello 02-02- 2016**

 1 compito 3 crediti

Nome: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Cognome: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Matricola: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Specificate se su Infostud dovete verbalizzare**:

 a) 9 + 3 crediti Nuovo Ordinamento (Calcolo, Biostatistica+MMIB)

 oppure

 b) SOLO MMIB 3 crediti Vecchio Ordinamento

**ESERCIZIO 1**

Si vuole studiare la durata di vita media µ della popolazione di malati di tumore alla prostata con 3 diverse diagnosi di gravità (livello di gravità da 1 a 3) dopo la prima visita in ospedale. Per poter estrarre un campione di 150 individui con tale diagnosi, vengono considerati i database di 5 ospedali a Roma negli anni 2013-2014.

i) dite che tipo di campionamento usereste

**ii) spiegate** come fareste, **in pratica**, per estrarre il campione.

iii) µ indica una statistica campionaria **V F**

**ESERCIZIO 2**

In una ricerca genetica si è scoperto che 400 soggetti, su 1200 esaminati, sono portatori di una mutazione che può indurre una malattia rara. Il campione esaminato è stato estratto casualmente.

i) Quale modello (distribuzione) probabilistico può descrivere questo esperimento aleatorio?

ii) Quali **condizioni** sono necessarie per poter applicare questa distribuzione?

ii) dite **quali sono e quali valori** assumono i parametri di questo modello (distribuzione)

ESERCIZIO 3

Considerate queste coppie di dati.

 **x y**

80.7 2.00

71.8 3.64

75.6 3.42

72.7 2.84

75.3 1.83

73.6 2.82

75.1 4.04

74.0 2.56

73.8 0.80

70.4 4.19

Dopo aver osservato il grafico di dispersione e calcolato l’equazione della retta di regressione

i) dite a **quale valore x** corrisponde il residuo maggiore in valore assoluto

ii) quanto vale tale residuo

i) ii)

iii) il punto corrispondente a tale valore x può essere considerato un outlier o un’osservazione influente rispetto alla retta di regressione? **Spiegate la vostra scelta**

iv) quale percentuale di variabilità delle y è spiegata dal modello di regressione dei minimi quadrati?

v) **spiegate** come avete calcolato tale percentuale

vi) spiegate come si interpreta questo risultato relativamente al modello di regressione

 **ESERCIZIO 4**

Osservate il seguente box-plot

|  |  |  |  |
| --- | --- | --- | --- |
|  |   |   |   |
|  |   |   |  |
|

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|   |   |   |   |   |   |   |   |   |   |
| 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 |  |

 |

 i) dite che forma ha la distribuzione dei dati rappresentati dalla figura

ii) riportate tutti gli indici di centralità e di variabilità che sono opportuni per questi dati

 iii) riportate la formula della differenza interquartile e calcolatela per questi dati

iv) in un box-plot cosa rappresenta la scatola che compare nella fig.?

v) dite quanto vale approssimativamente la mediana

vi) la media è più grande della mediana per questi dati V F

vii) quale percentuale di osservazioni è approssimativamente inferiore o uguale a 12?

a- 20%

b- 40%

c- 25%

d- nessuna delle precedenti

viii) il box-plot può essere usato per variabili qualitative ordinate V F

ESERCIZIO 5

i) Se una distribuzione normale ha media pari a 200 e deviazione standard pari a 50, trovare quel valore K tale che la proporzione di valori minori di K sia pari a 0.975.

a) 239

b) 238

c) 298

d) 300

e) 286

ii) per la stessa distribuzione calcolate il percentile corrispondente al valore 157

**ESERCIZIO 6**

Il numero di anelli sulle scaglie indica l’età di un pesce. In un allevamento di trote è stato osservato un campione grande, ed è stata determinata l’età (anni) di ogni pesce. I dati sono riportati nella tabella.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Età in anni  | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| N° di pesci catturati | 16 | 19 | 13 | 12 | 10 | 6 | 1 |

i) Calcolate media, mediana e moda dell’età per questo campione.

ii) Considerando l’età come variabile discreta, disegnate qui sotto, il grafico che ritenete più opportuno per rappresentare l’età

iv) Dite quale o quali altri grafici sarebbe possibile usare (senza disegnarli) per rappresentare questo campione di dati.