

CORREZIONI

Metodi Matematici e Informatici per la Biologia

Appello del 28- 01- 2013

I compito 3 crediti

ESERCIZIO 1

Sapendo che la durata (in giorni) della gravidanza è approssimativamente distribuita come una normale con media 270g e deviazione standard 10g:

i) calcolare la probabilità che la durata sia superiore a 280g.

PT 2

$$P(X > 280) = 1 - .8413 = .1587$$

Normal with mean = 270 and standard deviation = 10

$$P(X > 280) = 1 - .8413 = .1587$$

ii) a quale percentile corrisponde il valore di 230g?

PT 2

$$P(X < 230) = .0003217 \rightarrow 0.03\%$$

iii) ricordando la definizione di percentile, qual è il significato del valore del percentile ottenuto nel punto ii)?

PT 2

Lo 0.03% dei valori (giorni) cadono a sin di 230 nella distribuzione normale

ESERCIZIO 2

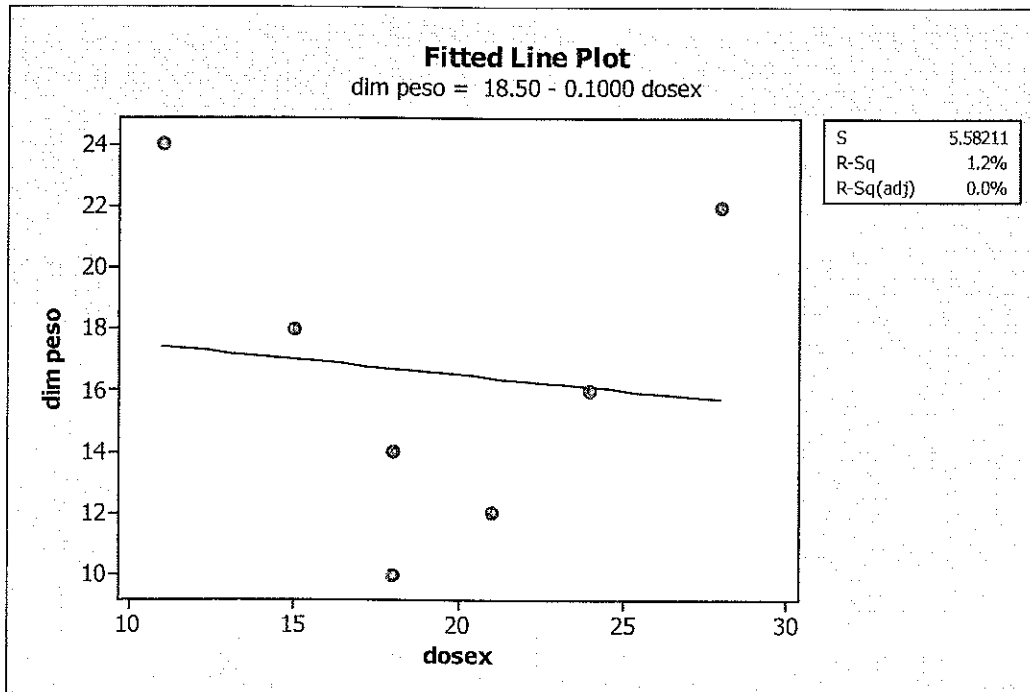
Nella tabella che segue sono riportate le dosi (in mg) di un farmaco contro l'obesità e le relative diminuzioni di peso (Kg) misurate in 7 pazienti dopo 90 giorni di cura.

I dati sono solo una parte dei dati riportati su J of M (1994) per uno studio sull'efficacia di un farmaco contro l'obesità.

Diminuzione (y)	Dose (x)
22	28
10	18
18	15
14	18
16	24
24	11
12	21

i)

Stabilite se esiste una relazione tra la dose (x) e la diminuzione di peso (y) usando la retta di regressione.



a) Riportate il coefficiente angolare e la sua unità di misura.

PT 1,5

R: -0,100 Kg/mg

b)

PT 2

R: variazione della v. di risposta aumentando di un'unità la v. esplicativa

c) Qual è la variazione di peso che ci si può attendere utilizzando un mg aggiuntivo di farmaco?

PT 1

R: -0,10 kg

c)

1) Quale osservazione presenta il residuo maggiore in valore assoluto?

2) Riportate il valore di tale residuo.

PT 1,5 + 1,5 = 3

R: 1) x=18; y=10 2) 6,70

d) riportate le due formule per il calcolo del coefficiente di determinazione che conoscete. Riportate anche il valor di tale coefficiente

R: $R^2 = \frac{\text{var}\hat{y}}{\text{var}y}$ PT 1+ 2=3

2) Riportate anche il valor di tale coefficiente

R: 1.2% PT 1

e) Spiegate la forma del grafico dei residui e dite quali conclusioni possiamo trarne per questo esempio
PT 2.5

R: forma ad imbuto – disposizione non casuale dei pt intorno alla retta dei residui

f) f) Si può concludere che la retta di regressione rappresenta un buon modello per i dati del campione osservato? Spiegate.
PT 1+1.5= 2.5

No Coeffi di determinazione quasi nullo Residui non disposti casualmente intorno alla retta dei residui

ESERCIZIO 3

Ad altri 11 pazienti di un secondo gruppo è stato somministrato un altro tipo di farmaco contro l'obesità e sono state misurate le riduzioni di peso (Kg) che appaiono qui di seguito:

18
22
22
24
23
25
36
39
44
44
69

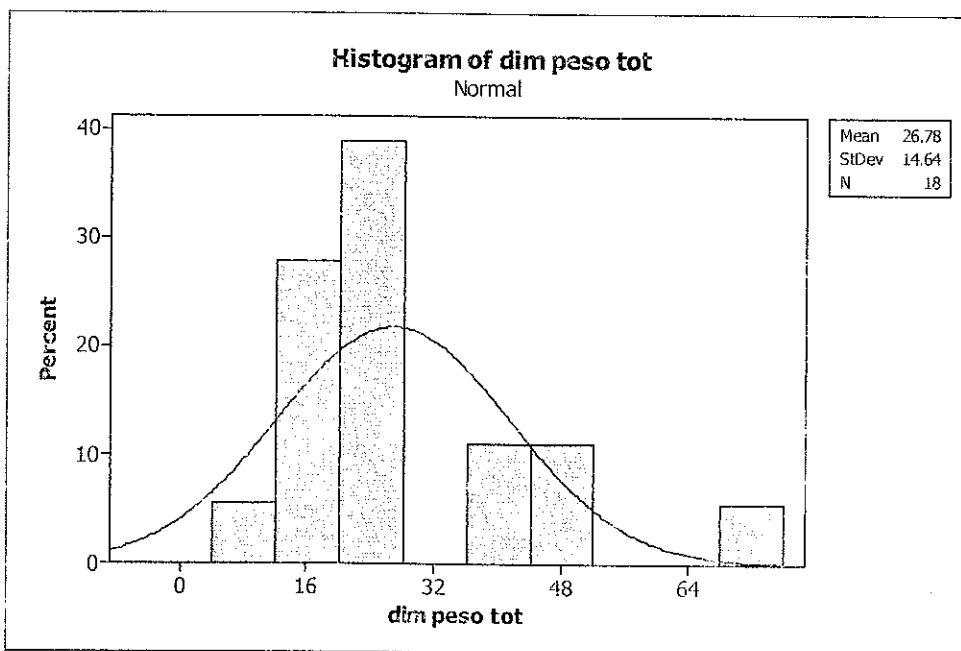
i) Accorpate i valori della riduzione di peso nei 2 gruppi osservati. Per questi 18 dati, considerate 9 classi di ampiezza pari a 8 a partire dal valore 4 (anche se non è un valore compreso nel campione). Per ogni classe calcolate la freq rel perc e riportatele qui sotto.

PT 2.5

R: 4—12	5,5
12---20	27,7
20—28	38,8
36----44	0
44---52	11,1
68----76	11,1
	0
	0
	5,5

ii) Poi disegnate l'istogramma, riportando il grafico sul foglio qui sotto. Scegliete una scala tale che il grafico sia leggibile. Specificate sul grafico cosa riportate su entrambi gli assi.

PT 3

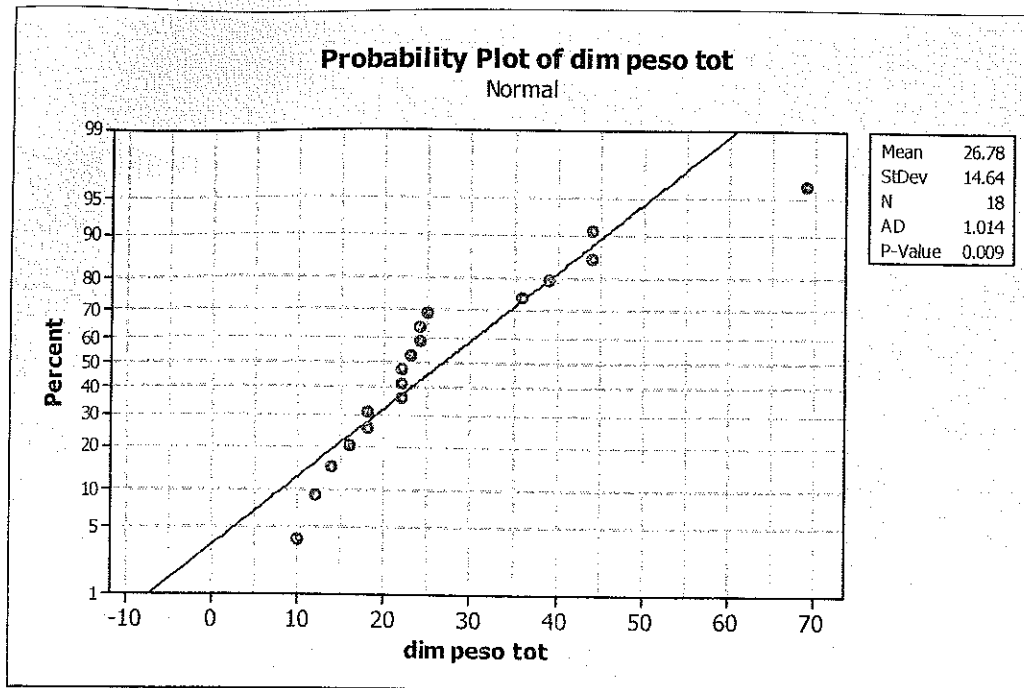


iii) Verificate la normalità di queste misure costruendo un normal plot che non dovete riportare sul foglio:

Questi dati si distribuiscono approssimativamente secondo una normale? Perché sì o perché no? Spiegate. (Può essere d'aiuto il calcolo dei coefficienti di asimmetria e forma)

PT 2

R: NO Si discostano dalla retta (di regressione)



iv) Per questi 18 dati calcolate il 11-mo percentile

PT 1.5

R: 12

dim peso tot	Count	CumCnt	CumPct
10	1	1	5.56
12	1	2	11.11
14	1	3	16.67
16	1	4	22.22
18	2	6	33.33
22	3	9	50.00
23	1	10	55.56

v) Indicate (senza calcolarli) tutti gli indici di centralità e di variabilità che ritenete opportuni per questi dati

PT 1.5

mediana, moda quartili

ESERCIZIO 4

PT 1.5

Quale delle seguenti affermazioni è corretta?

- 1) Un sondaggio basato su un campione è un esempio di esperimento
- 2) Il miglior metodo per cercare una relazione causale è uno studio di osservazione

- a) solo la 1)
- b) solo la 2)
- c) tutte
- d) nessuna

ESERCIZIO 5

Qual è il miglior metodo di raccolta dei dati quando si vogliono isolare le variabili nascoste o di confondimento?

PT 1.5

ESPERIMENTO