

ESERCIZIARIO

1) Sono state raccolte le seguenti sei coppie (x,y) di dati ottenuti rilevando variabili quantitative X e Y su 6 unità sperimentali: (1,3) (2,7) (5,5) (7,6) (10,8) (11,13).

- Calcolare la media, la mediana e la varianza di X e Y
- Misurare opportunamente il legame lineare tra X e Y e determinare l'equazione della retta di regressione di Y rispetto a X
- Detti Z e W variabili così definite $Z=2X+3$ e $W=4Y$, calcolare la media, la mediana e la varianza di Z e W

2) A causa di un errore informatico sono state modificate le età X osservate su un gruppo di 200 persone. In particolare è emerso che ogni età ha subito la seguente modifica: $Z=2X+3$; dei valori modificati si sono calcolati media, pari a 35 anni, mediana di 34 anni e scarto quadratico medio di 10.3 anni.

Calcolare media, mediana e scarto quadratico medio, se possibile, delle età corrette.

3) In uno studio atto a rilevare le relazioni fra età e pulsazioni cardiache al minuto e sotto sforzo, sono state effettuate le seguenti rilevazioni:

età (anni) X	15	20	25	30	35	40	45	50	55	60	65	70
puls al min Y	200	195	190	190	180	185	180	175	170	170	165	160

- Fornire:
- a) il numero mediano, modale e medio di X e Y
 - b) un opportuno indice di dispersione (eventualmente relativizzato) di X e Y
 - c) una misura di legame lineare fra X e Y.

4) Su un campione di 150 famiglie è stata rilevata la statura in cm del padre (X) e del figlio primogenito (Y), ottenendo i seguenti risultati

Y \ X	160-170	170-180	180-190	190-200	TOTALE
150-160	20	5	0	0	25
160-170	10	15	10	3	38
170-180	7	13	12	5	37
180-190	3	10	10	6	29
190-200	5	2	6	8	21
TOTALE	45	45	38	22	150

- a) costruire una tabella con le frequenze relative
- b) costruire una tabella con i profili riga e una con i profili colonna
- c) determinare la media aritmetica, la moda e la mediana di X e Y (attenzione: i dati sono raggruppati in classi, utilizzare come rappresentante della classe il punto medio dell'intervallo di valori)

5) La variabile quantitativa X, lunghezza di chiodi, rilevata su 20 unità statistiche ha fornito i seguenti valori, espressi in cm:

2.2 2.4 2.6 2.8 2.3 2.5 2.1 2.9 3.0 3.0 3.1 3.2 2.3 2.1 2.6 2.9 2.8 2.9 3.1 3.2

Determinare moda, mediana, media e scarto quadratico medio.

Se tutti i dati fossero aumentati di 0.8 cm, come si modificherebbero i risultati appena ottenuti, senza rifare i calcoli ex novo?

6) In una classe di 15 allievi si sono conseguiti i seguenti risultati delle prove di algebra e di fisica:

algebra	3	5	6	6	6	7	7	7	8	8	8	8	9	9	10
fisica	2	3	5	6	6	6	5	6	6	8	9	8	8	9	9

- a) Calcolare media aritmetica, moda, mediana e un indice di dispersione di entrambe le materie.
- b) Fornire un indice che misura il legame di dipendenza lineare fra le due valutazioni
- c) Supponendo che a questa classe si aggiunga un 16° allievo di cui si conosce il risultato della prova di algebra, 7, discutere come si può stabilire il voto mancante di fisica scegliendo fra le seguenti possibilità quella che ritenete più favorevole all' allievo:
 - media aritmetica dei voti dei suoi compagni nella stessa materia
 - mediana dei voti dei suoi compagni nella stessa materia (perchè NO?)

→ valore che si ottiene dalla retta di regressione dei voti di fisica rispetto a quelli di algebra

d) Se il professore di fisica decide di aumentare tutti i voti di 1 punto, come cambiano i valori di centralità, di dispersione e quello di dipendenza lineare?

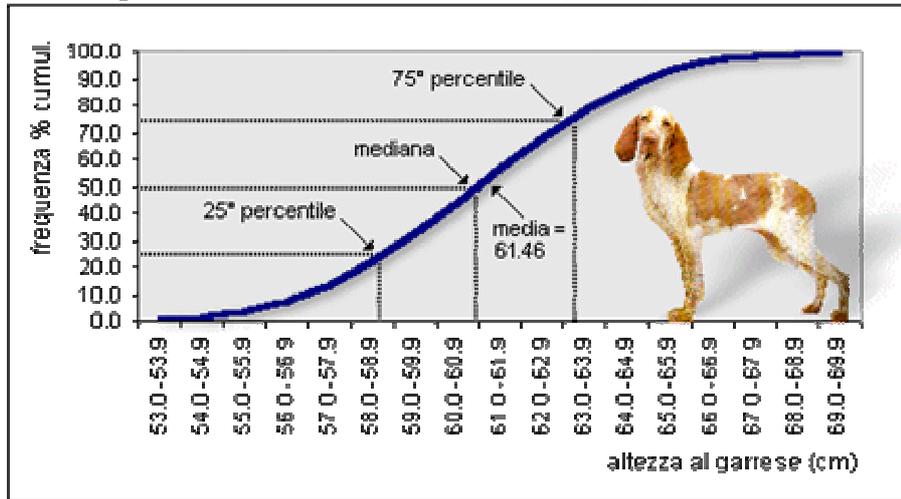
e) Valutare la veridicità delle seguenti frasi:

- si hanno prove negative in fisica perché non si conosce l'algebra
- mediamente gli allievi vanno meglio in algebra piuttosto che in fisica
- c'è meno dispersione relativa in algebra piuttosto che in fisica

7) Il seguente grafico (tratto dal sito http://www2.unipr.it/~bottarel/epi/varbio/freq_cu.htm) mostra la funzione di distribuzione cumulata dell'altezza al garrese di 659 cani di razza Bracco italiano.

Disegnare il box-plot corrispondente.

Altezza al garrese di 659 cani "Bracco italiano": frequenze percentuali cumulative



Prof. Ezio Battarelli - Università di Parma

8) La seguente tabella riporta la distribuzione sul territorio italiano (Y) divisa per sesso (X) della popolazione italiana nell'anno 2004 (fonte ISTAT):

Popolazione sul territorio 2004			
	M	F	TOT
Nord-Ovest	7.493.203	7.945.238	15.438.441
Nord-Est	5.380.080	5.650.570	11.030.650
Centro	5.415.770	5.830.189	11.245.959
Sud	6.854.041	7.230.151	14.084.192
Isole	3.233.710	3.429.423	6.663.133
Italia	28.376.804	30.085.571	58.462.375

a) Costruire la tabella di contingenza con i valori percentuali della distribuzione congiunta (X,Y).

b) Costruire le due distribuzioni di frequenza marginali e due diagrammi a barre che le rappresentino.

c) Date le due distribuzioni marginali è possibile, risalire alla distribuzione delle frequenze congiunte? Come si fa?

9) La seguente tabella riporta la distribuzione dei decessi per fasce di età (Y) e per sesso (X) della popolazione italiana nell'anno 2002 (fonte ISTAT):

Morti per età e sesso – Anno 2002			
	M	F	TOT
0-4 anni	1563	1211	2774
5-19 anni	1343	572	1915
20-39 anni	8379	3266	11645
40-54 anni	16320	9238	25558
55-69 anni	57232	30715	87947
70-79 anni	88331	64105	152436
80-89 anni	79272	105664	184936
90 e oltre	26856	66323	93179
TOTALE	279296	281094	560390

- a) Costruire la tabella di contingenza con i valori percentuali della distribuzione congiunta (X,Y).
 b) Costruire le due distribuzioni di frequenza marginali e due diagrammi a barre che le rappresentino.
 c) Date le due distribuzioni marginali è possibile, risalire alla distribuzione delle frequenze congiunte? Come si fa?

10) Un esperimento scientifico ha fornito la misura di una grandezza A con i seguenti risultati:

6.45	6.52	6.50	6.46	6.49	6.52	6.41	6.49	6.51	6.50
6.59	6.50	6.52	6.51	6.50	6.49	6.51	6.49	6.46	6.55

- a) Costruite la funzione di distribuzione cumulata
 b) Indicare il primo quartile, la mediana, il secondo quartile e la distanza interquartile e costruire il box-plot.

11) In un Liceo scientifico ci sono 5 classi seconde, i cui alunni sono distribuiti per sezione e per sesso in base alla seconda tabella:

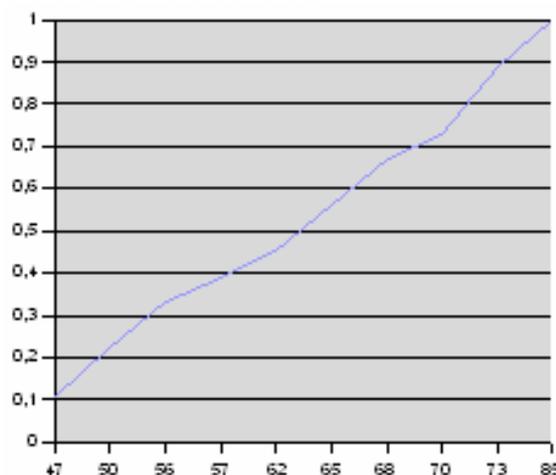
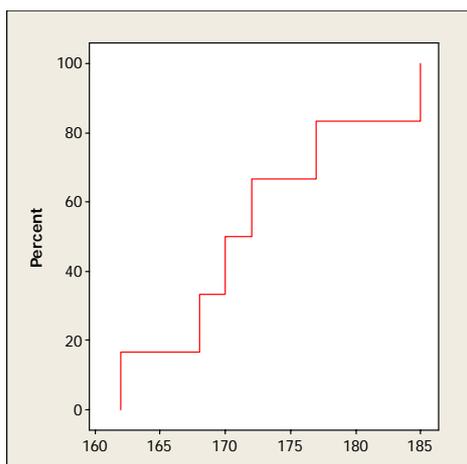
Sesso\sezione	A	B	D	E	F
M	16	18	7	5	8
F	10	7	13	16	11

Calcolare:

- le distribuzioni marginali degli studenti per sezione e per sesso.
- la tabella di distribuzione congiunta.
- la tabella dei profili riga.
- la tabella dei profili colonna.

12) Date le seguenti funzioni di distribuzione cumulata di due variabil, trovare di ciascuna di esse la mediana, il primo quartile, il terzo quartile, l'intervallo interquartile.

Discutere le due diverse rappresentazioni grafiche della funzione di distribuzione cumulata.



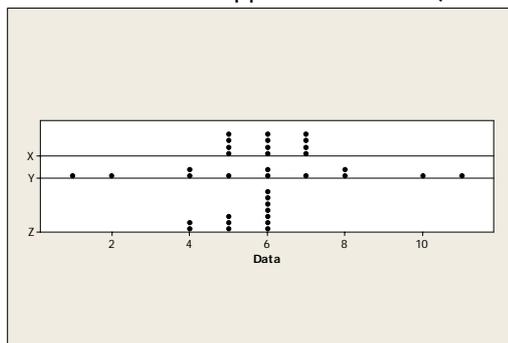
13) Date le seguenti tabelle, che rappresentano le frequenze cumulate dei pesi (X) e delle altezze (Y) degli studenti di una classe, calcolare la media del peso e delle altezze, la varianza, lo scarto quadratico medio, l'IQR(inter quartile range) e l'intervallo di variazione.

X	F(x)
45	0.038
48	0.077
49	0.115
50	0.192
52	0.231
54	0.308
55	0.346
58	0.385
60	0.615
63	0.654
65	0.885
66	0.923
83	0.962
84	1.000

Y	F(y)
156	0.038462
158	0.076923
160	0.153846
163	0.192308
165	0.346154
170	0.538462
172	0.692308
173	0.730769
175	0.846154
176	0.923077
183	0.961538
184	1.000000

14) Sapendo che i dati nella tabella si riferiscono ai dotplot a sinistra, associarli opportunamente. (motivare la risposta)

	media	Deviazione standard
A	5.417	0.793
B	6.000	0.853
C	6.000	3.015



15) Si misura 100 volte una quantità e si rilevano i seguenti dati:

$$\sum_{i=1}^{100} x_i = 2480.82$$

$$\sum_{i=1}^{100} x_i^2 = 125147$$

Calcolare media e scarto quadratico medio dei dati.

16) La seguente tabella rappresenta dei dati osservati in una classe dal 1/9/05 al 22/12/05

Alunno	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Giorni assenza	5	1	3	1	1	6	7	3	3	2	12	4	4	1	2	2	4	3	1	1	13	5	1

- Raccogliere i dati in una tabella di frequenza.
- Calcolare media e moda dei dati.
- Calcolare la funzione di distribuzione cumulata e rappresentarla graficamente.
- Calcolare la mediana dei dati.

17) Sotto sono riportati i pesi di bambini di 2 anni espressi in kg:

8.9 8.8 8.7 8.9 8.6 8.6 8.8 9.0 9.1 9.0 9.2 9.2 9.3 9.3
9.4 9.5 9.3 9.4 9.5 9.3 8.5 9.0 9.4 9.5 9.6 9.3 9.7 9.8

- Costruire una tabella con le frequenze assolute, la distribuzione dei pesi, la funzione cumulata assoluta e la funzione di distribuzione cumulata.
- Determinare i valori dei quartili Q1, Q2 e Q3 e l'ampiezza dell'intervallo interquartile.
- Calcolare la media e lo scarto quadratico medio del peso.

18) Per il seguente insieme di dati $A = \{2, 4, 4, 4, 5, 5\}$ indica quali delle seguenti relazioni sono vere:

- mediana = moda
- moda < media
- mediana > media
- media = moda

19) Sono rilevati i seguenti valori di una variabile X su 5 unità sperimentali: 3,6,2,1,7,5. Si costruisce una variabile Y aggiungendo 5 a ciascuno dei valori e si ottiene 8,11,7,6,12,10.

- tracciare un diagramma di dispersione di X e di Y.
- mostrare che X e Y hanno lo stesso scarto quadratico medio ma medie diverse.
- come sono collegate tra loro le medie di X e di Y?

20) Sono rilevati i seguenti valori di una variabile Z su 5 unità sperimentali: 3,6,2,1,7,5. Si costruisce una variabile T moltiplicando per -2 i dati e aggiungendo 5 e si ottiene -1,-7,1,3,-9,-5.

- quale relazione c'è tra le medie di Z e T? E tra gli scarti quadratici medi?
- quali proprietà della media e degli s.q.m. si deducono dai punti b), c), d)?
- dire – senza fare i calcoli espliciti – qual è il segno della covarianza fra X e Y e quello Z e T? Perché?
- dire – senza fare i calcoli espliciti – quanto vale l'indice di correlazione fra X e Y e quello Z e T? Perché?
- verificare con il calcolo le risposte f) e g)

21) In un totale di n numeri il numero 1 si presenta p volte mentre il numero 0 si presenta q volte, con $q = n - p$.

- dimostrare che lo scarto quadratico medio vale $\sqrt{\frac{pq}{n^2}}$

b) trovare lo scarto quadratico medio dei numeri 0,0,0,0,0,1,1,1 utilizzando le due formule.