

**ESERCIZI DI STATISTICA DESCRITTIVA
ALCUNI TRATTI DA PROVE D'ESAME
DA REALIZZARE ANCHE CON L'AUSILIO
DI UN FOGLIO DI CALCOLO**

Esercizio. E' stato tabulato il peso di una certa popolazione di individui.

Peso p (in kg)	F_{ass}
$40 \leq p < 45$	2
$45 \leq p < 50$	12
$50 \leq p < 55$	21
$55 \leq p < 60$	17
$60 \leq p < 65$	18
$65 \leq p < 70$	22
$70 \leq p < 75$	18
$75 \leq p < 80$	7
$80 \leq p < 90$	3

Si determini:

- il peso medio come media aritmetica, poi come moda e infine come mediana;
- lo scarto quadratico medio (o deviazione standard) sulla popolazione e sul campione
- si costruisca infine un istogramma delle frequenze.

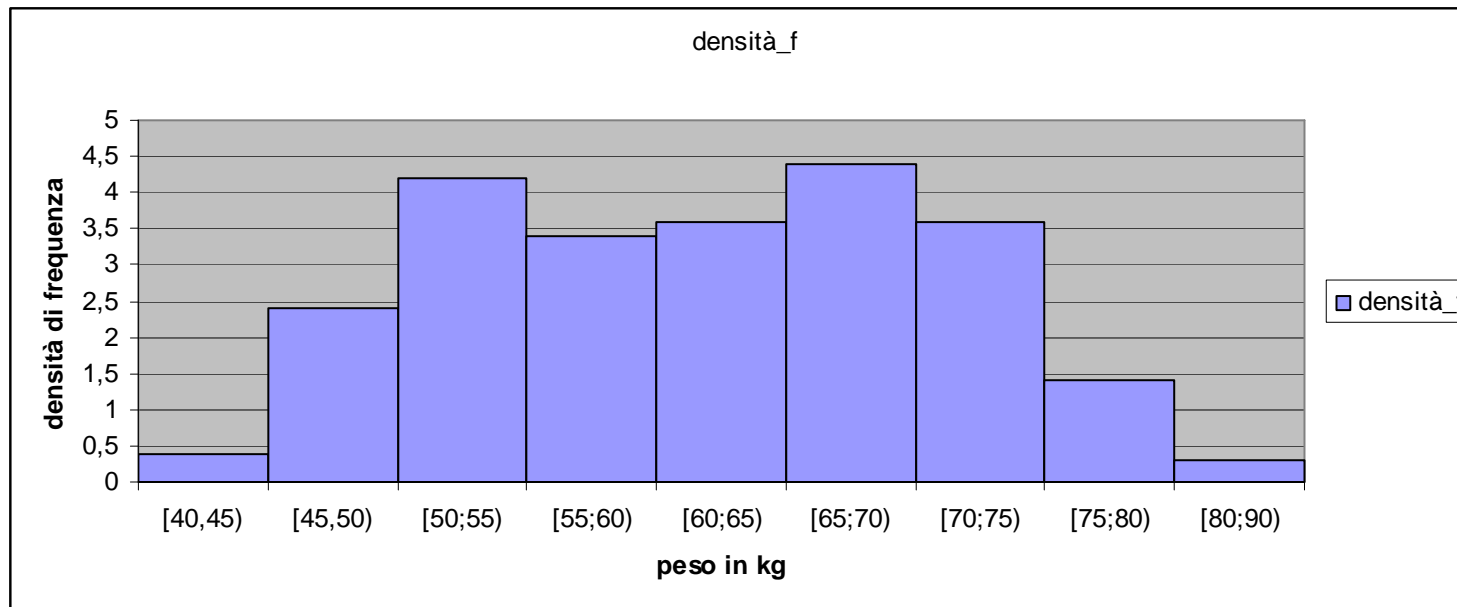
Peso in kg	Fi	densità_f	fr	Fcumulate	fr_cumul	fr_cumul_%	Xci	Xci * Fi	Xci^2	(Xi^2)*Fi
[40;45)	2	0,4	0,02	2	0,02	1,67	42,5	85	1806,25	3612,5
[45;50)	12	2,4	0,1	14	0,12	11,67	47,5	570	2256,25	27075
[50;55)	21	4,2	0,18	35	0,29	29,17	52,5	1102,5	2756,25	57881,25
[55;60)	17	3,4	0,14	52	0,43	43,33	57,5	977,5	3306,25	56206,25
[60;65)	18	3,6	0,15	70	0,58	58,33	62,5	1125	3906,25	70312,5
[65;70)	22	4,4	0,18	92	0,77	76,67	67,5	1485	4556,25	100237,5
[70;75)	18	3,6	0,15	110	0,92	91,67	72,5	1305	5256,25	94612,5
[75;80)	7	1,4	0,06	117	0,98	97,5	77,5	542,5	6006,25	42043,75
[80;90)	3	0,3	0,03	120	1	100	85	255	7225	21675
Totali	120		1					7447,5	37075	473656,25

media	62,06
-------	-------

varianza_pop	95,38
dev_sta_pop	9,77
varianza_camp	128,55
dev_sta_cam	11,34

Per determinare la mediana ricordo che, essendo una distribuzione di frequenze, è necessario guardare le frequenze cumulate. Il 60° e il 61° dato cadono nella classe [60;65) Essendo una distribuzione con classi, la mediana si calcola con la formula:

$$\text{mediana} = x_{\text{min_intervallo}} + \frac{\left(\frac{\sum f_i}{2} - f_{\text{cumul}_{i-1}} \right) \cdot \text{ampiezza}}{f_i} = 60 + \frac{(60 - 52) \cdot 5}{18} = 62,2\text{kg}$$



Esercizio. Tratto dal tema d'esame del 12/09/2011

4) Data la seguente serie di dati statistici:

X	1	2	3	5	6	7
Y	0	3	5	10	11	14

Si trovi la retta di regressione di Y su X ed il valore dell'indice di regressione R^2 di Bravais-Pearson.

Innanzitutto bisogna valutare se è significativo ricercare una regressione lineare dei dati ottenuti. Ciò significa che bisogna innanzitutto accertarsi se esiste o meno una correlazione lineare tra i dati. Tale valutazione parte dall'analisi della "nube" di dati e dal calcolo del

coefficiente di correlazione lineare $r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$. Ricordiamo che per il calcolo di r è

necessario conoscere la covarianza delle variabili x e y e le deviazioni standard di entrambe rispetto alle rispettive medie aritmetiche.

Calcoliamo tutti gli elementi utili: $\bar{x} = \frac{\sum x_i}{n} = 4$ $\bar{y} = \frac{\sum y_i}{n} = 7,1667$

Varianza della X $\sigma^2_x = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{\sum x_i^2}{n} - \bar{x}^2 = 4,667$

Deviazione standard X $\sigma_x = \sqrt{\sigma^2_x} = 2,16$

Varianza della Y $\sigma^2_y = \frac{\sum (y_i - \bar{y})^2}{n} = \frac{\sum y_i^2}{n} - \bar{y}^2 = 23,8$

Deviazione standard Y $\sigma_y = \sqrt{\sigma^2_y} = 4,87$

Covarianza $\sigma_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n} = \frac{\sum x_i y_i}{n} - \bar{x} \cdot \bar{y} = 10,5$

Coefficiente di correlazione $r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} = 0,9962 \Rightarrow r^2 = 0,9924$ (di Bravais - Pearson)

r è prossimo ad 1, pertanto c'è una forte correlazione lineare positiva

A questo punto ricerchiamo l'equazione della retta di regressione mediante il metodo dei minimi quadrati. Sappiamo che la retta avrà equazione $y = mx + q$, con

$$m = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sigma_{xy}}{\sigma_x^2} = 2,25 \quad q = \bar{y} - m\bar{x} = -1,833$$

$$y = 2,25x - 1,833$$

Si ricorda infine che tale retta è ricercabile scrivendo l'equazione del fascio proprio di rette

di centro $(\bar{x}; \bar{y})$ e avente coefficiente angolare $m = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sigma_{xy}}{\sigma_x^2} = 2,25$

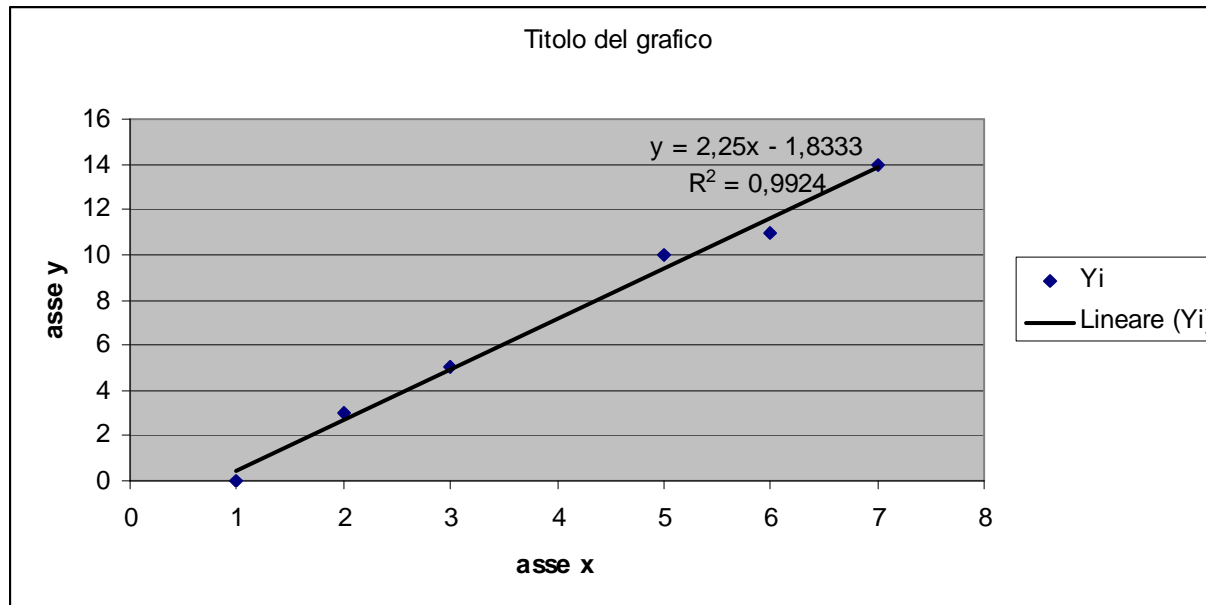
$$y - \bar{y} = m(x - \bar{x})$$

Xi	Yi	Xi^2	Yi^2	XiYi
1	0	1	0	0
2	3	4	9	6
3	5	9	25	15
5	10	25	100	50
6	11	36	121	66
7	14	49	196	98

Totale	24	43	124	451	235
--------	----	----	-----	-----	-----

xm	4
ym	7,166667
xm^2	16
ym^2	51,36111
xm*ym	28,66667

covarianza	10,5
varianza x	4,666667
varianza y	23,80556
dev_sta_x	2,160247
dev_sta_y	4,879094
r	0,9962
r^2	0,992415
m	2,25



Esercizio. Tratto dal tema d'esame del 19/06/2012

4) Si consideri la seguente seriazione di dati:

x_i	10	12	-2	-1	3	dati
f_i	3	5	2	8	1	frequenze

Si calcolino: media aritmetica, scarti, varianza (sul campione) e deviazione standard (sul campione). Si faccia un istogramma delle frequenze.

Esercizio. Tratto dal tema d'esame del 18/07/2011

4) Per la seguente distribuzione univariata della v.a. X

X	1	2	3	4
f_i	3	2	4	3

si calcolino: media aritmetica, media geometrica, media armonica, varianza e scarto quadratico medio.

Esercizio. Tratto dal tema d'esame del 01/03/2011

4) Data la seguente tabella di dati rilevati:

X	-2	-1	0	2	3	4
Y	1	2	3	4	5	6

Si stabilisca la retta di regressione lineare di Y su X con il relativo coefficiente di regressione R^2 (di Bravais - Pearson)

Innanzitutto bisogna valutare il coefficiente di correlazione lineare $r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$.

Calcoliamo tutti gli elementi utili: $\bar{x} = \frac{\sum x_i}{n} = 1$ $\bar{y} = \frac{\sum y_i}{n} = 3,5$

Varianza della X $\sigma_x^2 = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{\sum x_i^2}{n} - \bar{x}^2 = 4,67$

Deviazione standard X $\sigma_x = \sqrt{\sigma_x^2} = 2,16$

Varianza della Y $\sigma_y^2 = \frac{\sum (y_i - \bar{y})^2}{n} = \frac{\sum y_i^2}{n} - \bar{y}^2 = 2,92$

Deviazione standard Y $\sigma_y = \sqrt{\sigma_y^2} = 1,71$

Covarianza $\sigma_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n} = \frac{\sum x_i y_i}{n} - \bar{x} \cdot \bar{y} = 3,67$

Coefficiente di correlazione $r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} = 0,9939 \Rightarrow r^2 = 0,9878$ (di Bravais - Pearson)

r è prossimo ad 1, pertanto c'è una forte correlazione lineare positiva

A questo punto ricerchiamo l'equazione della retta di regressione mediante il metodo dei minimi quadrati. Sappiamo che la retta avrà equazione $y = mx + q$, con

$$m = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{\sigma_{xy}}{\sigma_x^2} = 0,7857 \quad q = \bar{y} - m\bar{x} = 2,7143$$

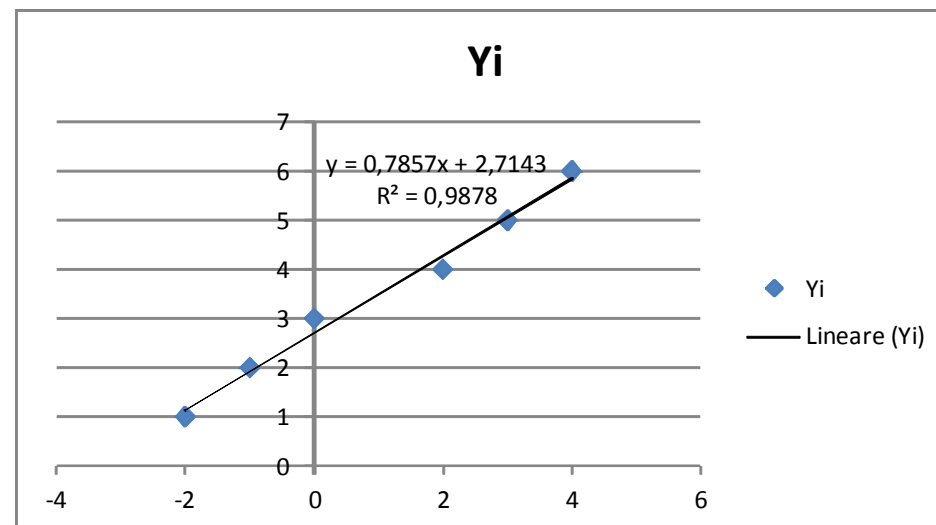
$$y = 0,7857x + 2,7143$$

Xi	Yi	Xi^2	Yi^2	XiYi	Xi-xm	(Xi-xm)^2	Yi-ym	(Yi-ym)^2	(xi-xm)(yi-ym)
-2	1	4	1	-2	-3	9	-2,5	6,25	7,5
-1	2	1	4	-2	-2	4	-1,5	2,25	3
0	3	0	9	0	-1	1	-0,5	0,25	0,5
2	4	4	16	8	1	1	0,5	0,25	0,5
3	5	9	25	15	2	4	1,5	2,25	3
4	6	16	36	24	3	9	2,5	6,25	7,5

Totale	6	21	34	91	43	0	28	0	17,5	22
--------	---	----	----	----	----	---	----	---	------	----

xm	1
ym	3,5
xm^2	1
ym^2	12,25
xm*ym	3,5

covarianza	3,666667	covarianza	3,666667	cov_camp	
varianzax	4,666667	varianzax	4,666667	var_camp_x	5,6
varianzay	2,916667	varianzay	2,916667	var_camp_y	3,5
dev_stan_x	2,160247			dev_s_camp_x	2,366432
dev_stan_y	1,707825			dev_s_camp_y	1,870829
r	0,993859				
r^2	0,987755				
m	0,785714				



Esercizio. Tratto dal tema d'esame del 27/09/2012

4) Data la seguente tabella di dati rilevati:

X	-1	0	1	2	3	4
Y	3	2	0	-3	-6	-7

Si stabilisca la retta di regressione lineare di Y su X con il relativo coefficiente di regressione R^2 (di Bravais - Pearson)

Esercizio. Tratto dal tema d'esame del 19/07/2012

4) Si consideri la seguente seriazione di dati:

x _i	1	2	3	4	5	dati
f _i	4	1	2	1	2	frequenze

Si calcolino: scarti e loro somma, media aritmetica e media geometrica.

Esercizio. Tratto dal tema d'esame del 19/04/2012

4) Si consideri la seguente seriazione di dati:

x_i	2	3	4	-1	7	dati
f_i	3	5	2	18	1	frequenze

Si calcolino: media aritmetica, scarti, varianza e deviazione standard.
Si faccia un istogramma delle frequenze.

Esercizio. Su un campione di 100 pesci di un lago vengono rilevati due caratteri: la lunghezza (X) in cm e la razza (Y). E' data la seguente tabella a doppia entrata delle frequenze congiunte di due caratteri.

X \ Y	Trinche	Carpe	Orate
$0 < x < 40$	10	10	0
$40 < x < 60$	20	15	5
$60 < x < 70$	0	30	10

Si determinino: le distribuzioni marginali di X e Y; la distribuzione di X, condizionata dalla modalità "carpe" di Y; la lunghezza media dei pesci sul campione analizzato; la lunghezza media delle carpe rinvenute nel campione analizzato.

Esercizio. Le vendite di un certo bene e i relativi profitti totali (in euro) sono riportati nella seguente tabella:

Vendite	600	900	1200	1500	1800
Profitti	25000	37000	53000	63000	75000

Si chiede di:

- rappresentare i dati tramite un diagramma cartesiano a dispersione;
- costruire, tramite la retta di regressione dei dati, un modello che esprima il profitto in funzione delle vendite; (*interpolazione*)
- determinare il coefficiente di correlazione lineare tra le vendite e i profitti, commentandone il valore;
- determinare quale possa essere stato il profitto totale relativo ad una vendita di 1400 unità; (*estrapolazione*)
- determinare quale possa essere stato il profitto totale relativo ad una vendita di 2400 unità. (*estrapolazione*)

Esercizio. Nel rilevare le altezze di un gruppo di reclute si è ottenuta la seguente tabella delle frequenze.

Altezza (in cm)	F_{ass}
166	1
168	3
169	6
170	11
171	8
172	6
173	4
174	3
175	1
178	1

Si determini la media aritmetica, gli scarti e la loro somma, la media geometrica e la media armonica. Si calcoli la varianze e la deviazione standard (sul campione). Si determinino la moda e la mediana. Si rappresenti un istogramma delle frequenze. Si valutino gli indici di forma (asimmetria e curtosi) e si confronti quindi l'istogramma trovato con la curva normale di Gauss.