

Il volume:

A. Cerioli – M.A. Milioli – M. Riani “Esercizi di statistica parte I”, Uni.nova, Parma, 2006

contiene una raccolta di temi d'esame di statistica descrittiva, adeguati per la preparazione dell'esame di Statistica – Analisi dei dati.

Di seguito si riportano alcuni ulteriori esercizi per autovalutazione.

ESERCIZIO 1 (16 punti)

Si conosce la seguente matrice dei dati riferita a 10 impiegati ed ai fenomeni X=sesso ed Y=stipendio mensile (in euro):

Impiegati	X	Y
A	M	1350
B	M	1420
C	F	1230
D	M	1510
E	F	1360
F	F	1400
G	F	1210
H	F	1330
I	M	1400
L	F	1350

- i) Si costruisca la corrispondente tabella a doppia entrata, considerando le seguenti classi per il reddito: 1200-1300, 1300-1500, 1500-1600.
- ii) Si calcolino separatamente per i maschi e per le femmine (a partire dalla matrice dei dati) lo stipendio mensile medio e lo scostamento quadratico medio.
- iii) Si commentino i quattro valori ottenuti.
- iv) Si calcoli la media dello stipendio e la si commenti.
- v) Si calcoli la moda dello stipendio e la si commenti.
- vi) Si calcoli la mediana dello stipendio e la si commenti.

Esercizio 2 (15 punti)

Un'associazione di consumatori, per valutare la puntualità dei voli in partenza da un importante aeroporto internazionale, ha rilevato il ritardo (in minuti) con cui sono decollati i voli di linea in una determinata settimana. La tabella seguente riporta la corrispondente distribuzione di frequenze.

Ritardo (in minuti) dei voli in partenza	Frequenze
Sino a 5	126
5 – 15	420
15 – 30	233
30 – 60	94
60 – 120	24
Oltre 120 (in media 185)	3

- i) Si calcoli la media aritmetica e la si commenti.
- ii) Si calcoli la mediana e la si commenti.
- iii) Si calcoli il novantacinquesimo percentile e lo si commenti.

- iv) Motivando adeguatamente la risposta, si dica se i valori degli indici calcolati al punto i) sono uguali a quelli che sarebbero stati ottenuti partendo dalla matrice dei dati originaria (cioè contenente le informazioni sui ritardi dei singoli voli in partenza), anziché dalla distribuzione di frequenze.

ESERCIZIO 3 (15 punti)

Nella seguente tabella sono riportati per 4 ipermercati dell'area di Milano il fronte espositivo (in metri) relativi ai detergenti per la casa (X) e il numero di referenze (Y), fonte: Mark up.

Ipermercato	X	Y
Auchan	48	178
Carrefour	70	222
Ipercoop	80	192
Esselunga	70	159

- i) Si determini il fronte espositivo per referenza di ciascun ipermercato.
 ii) Si scriva l'espressione della media aritmetica del fronte espositivo per referenza e la si calcoli.
 iii) Si scriva l'espressione scostamento quadratico medio del fronte espositivo per referenza e lo si calcoli.
 iv) Si illustri il significato degli indici calcolati ai punti ii) e iii).

ESERCIZIO 4 (15 punti)

Nella seguente tabella sono riportati i rendimenti, in quintali per ettaro, di una certa varietà di frumento, per 100 appezzamenti di terreno:

rendimenti	n. di appezzamenti
54-58	16
58-62	23
62-66	25
66-70	21
70-74	15
	100

- i) Si calcolino i quartili dei rendimenti per gli appezzamenti di terreno
 ii) Si commenti il significato del terzo quartile dei rendimenti dei terreni.
 iii) Si costruisca il box-plot.
 iv) Si illustrino tutte le informazioni che si possono trarre dalla suddetta rappresentazione.

ESERCIZIO 5 (14 punti)

Nella seguente matrice dei dati, riferita a 7 famiglie, sono riportate le variabili:

- ❖ Spesa annua (in euro) per la partecipazione a spettacoli e manifestazioni culturali;
- ❖ Reddito mensile (in migliaia di euro) del capofamiglia:

Famiglia	Spesa annua (in euro) per spettacoli e manifestazioni culturali	Reddito mensile (in migliaia di euro) del capofamiglia
A	200	1,9
B	420	4,0
C	250	2,5
D	70	1,6
E	180	2,2
F	300	2,8
G	100	1,5

- i) Si costruisca la corrispondente matrice degli scostamenti standardizzati.
- ii) Si illustri il significato degli scostamenti standardizzati ottenuti per la famiglia B.
- iii) Si confronti la variabilità della spesa annua per spettacoli e manifestazioni culturali con quella del reddito mensile del capofamiglia.

ESERCIZIO 6 (14 punti)

I supermercati di una regione sono stati classificati in base allo spazio sugli scaffali, in metri, riservato al cibo per animali (X) e all'ammontare delle vendite settimanali, in migliaia di euro, (Y):

X \ Y	≤ 1	1 – 2	2 – 4	Totale
≤ 2	8	4	0	12
2 – 5	6	8	4	18
5 – 10	2	6	4	12
10 – 20	0	6	2	8
Totale	16	24	10	50

- i) Si calcolino le medie parziali di Y.
- ii) si commenti il significato dell'ammontare medio delle vendite quando lo spazio sugli scaffali è minore di 2.
- iii) Si riportino le medie ottenute al punto precedente in un grafico opportuno.
- iv) Si ricavi la media generale dell'ammontare delle vendite settimanali a partire dalle medie parziali calcolate al punto i).

ESERCIZIO 7 (15 punti)

La tabella seguente riporta la serie annuale del reddito mensile di un individuo, dal 1995 al 2000.

	Reddito (euro)
1995	1175
1996	1190
1997	1205
1998	1220
1999	1250
2000	1310

- i) Si calcoli la serie dei numeri indice a base fissa (base 1995=100).
- ii) Si calcoli la serie dei numeri indici a base mobile.
- iii) Si calcoli il valore del tasso medio di variazione annuale dello stipendio e lo si commenti.
- iv) Conoscendo la serie dei numeri indice dei prezzi al consumo per le famiglie di operai ed impiegati (FOI), si calcolino i valori deflazionati del reddito mensile.
- v) Si commenti il valore ottenuto per il 1999.

	FOI
1995	100
1996	103,9
1997	105,8
1998	107,7
1999	109,5
2000	112,2

ESERCIZIO 8 (16 punti)

Un'azienda vinicola commercializza due diverse etichette di vino, uno rosso ed uno bianco. Le variazioni percentuali rispetto all'anno precedente dei prezzi delle due etichette di vino sono risultate le seguenti:

Anno	Variazioni % prezzi vino rosso	Variazioni % prezzi vino bianco
2000	–	–
2001	+3,3%	+2,5%
2002	+7,6%	+6,4%
2003	+1,8%	+0,2%
2004	+2,9%	+3,1%

- i) Si calcolino le serie storiche dei numeri indici a base fissa dei prezzi delle due etichette di vino, con base 2000 = 100.
- ii) Si illustri l'interpretazione dei numeri indici calcolati nel 2004.
- iii) Sapendo che il fatturato dell'etichetta di vino rosso è triplo rispetto a quello dell'etichetta di vino bianco, si determini la corrispondente serie storica dei numeri indici composti dei prezzi, con base 2000 = 100.
- iv) Si illustri l'interpretazione del numero indice composto calcolato nel 2004.
- v) Si calcoli il tasso medio annuo di variazione dei numeri indici composti e lo si commenti.

ESERCIZIO 9 (12 punti)

Le variazioni percentuali rispetto all'anno precedente del prezzo di una marca di caffè sono risultate le seguenti:

Anno	Variazioni %
1999	–
2000	+2,5%
2001	+4,8%
2002	–1,7%
2003	+0,9%

- i) Si determini la serie storica dei numeri indici a base fissa del prezzo della marca di caffè, con base 2000 = 100.
- ii) Si illustri l'interpretazione del numero indice calcolato nel 2003.
- iii) Si calcoli il tasso medio annuo di variazione del prezzo della marca di caffè nell'intero periodo considerato (1999-2003).
- iv) Si commenti il significato dell'indice calcolato al punto iii).

ESERCIZIO 10 (13 punti)

La seguente tabella riporta la serie storica del fatturato, in migliaia di euro, di una panetteria e la serie storica dei numeri indici dei prezzi al consumo riferiti al pane, con base 2000 = 100 (Fonte: ISTAT).

Anno	Fatturato (migliaia di €)	N.i. dei prezzi al consumo per il pane (base 2000)
2000	162	100
2001	168	104,5
2002	169	107,6
2003	171	110,2

- i) Si determini la corrispondente serie storica del fatturato deflazionato della panetteria (ai prezzi del 2000).
- ii) Si commenti il valore calcolato per l'anno 2003.
- iii) Si calcoli il tasso medio annuo di variazione del fatturato a prezzi correnti.
- iv) Si calcoli il tasso medio annuo di variazione del fatturato a prezzi costanti.
- v) Si commentino in termini comparati i risultati ottenuti ai punti iii) e iv).

ESERCIZIO 11 (15 punti)

La seguente tabella riporta la serie storica, dal 2002 al 2006, della retribuzione (in euro) percepita da un lavoratore dipendente, il signor XY.

Anno	Retribuzione signor XY (€)
2002	20.500
2003	20.700
2004	21.200
2005	21.400
2006	21.700

- i) Si illustri che cosa significa *deflazionare* la retribuzione del signor XY.
- ii) Si descrivano sinteticamente i passaggi della corrispondente procedura di calcolo.
- iii) Disponendo delle variazioni percentuali rispetto all'anno precedente dell'indice dei prezzi al consumo per le famiglie di operai e di impiegati (indice FOI) determinate dall'ISTAT:

Anno	Variazioni % rispetto all'anno precedente dell'indice FOI (fonte: ISTAT)
2002	–
2003	+2,5%
2004	+2,0%
2005	+1,7%
2006	+2,0%

si determini la retribuzione del signor XY ai prezzi costanti del 2002.

- iv) Si commentino i risultati ottenuti.

ESERCIZIO 12 (12 punti)

La seguente serie storica riporta i profitti/perdite (in miliardi di lire) dell'Alitalia, nel periodo 1994–2003:

anni	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Y	-288	1,3	-1.217	438	408	0	-935	-1.756	178	-775

- i) Si calcolino le medie mobili di 5 termini dei profitti/perdite dell'Alitalia.
- ii) Si rappresentino in un unico grafico la serie storica originaria e le medie mobili ottenute al punto precedente, commentando il risultato ottenuto.
- iii) Si dica quale altra metodologia si potrebbe applicare per individuare il trend dei profitti/perdite dell'Alitalia.

ESERCIZIO 13 (17 punti)

La seguente tabella riporta, a cadenza biennale, la serie storica delle quantità estratte di idrocarburi (in milioni di tonnellate) in Italia dal 1984 al 1998:

Anno	Idrocarburi estratti (milioni di tonnellate)
1986	15,4
1988	18,3
1990	18,3
1992	18,6
1994	19,8
1996	19,7
1998	19,1

- i) Adottando un'opportuna scala dei tempi, si calcolino i parametri della funzione interpolante lineare che esprime la quantità di idrocarburi estratti in funzione del tempo.
- ii) Si illustri il significato dei parametri calcolati al punto precedente con riferimento alla serie storica degli idrocarburi.
- iii) Si determini la bontà di adattamento di tale funzione interpolante lineare.
- iv) Si spieghi il significato dell'indice ottenuto al punto precedente nel caso in esame.
- v) Si calcoli la stima della quantità di idrocarburi estratti nel 2004.
- vi) Si dica, con opportune motivazioni, se tale stima può essere ritenuta attendibile oppure no.

ESERCIZIO 14 (15 punti)

Nella seguente tabella sono riportati i numeri indici a base fissa con base 2000=100 del numero di brevetti per invenzioni depositate presso la Camera di Commercio di Parma nel periodo 2000-2004:

anni	n. i. base 2000=100
2000	100
2001	136,2
2002	118,8
2003	166,7
2004	143,5

- i) Sapendo che il numero di brevetti depositati nel 2000 è stato pari a 69, si ricavi la corrispondente serie storica del numero di brevetti (senza cifre decimali).
- ii) Si calcolino i parametri della funzione interpolante lineare del numero di brevetti in funzione del tempo.
- iii) Si illustri il significato dei parametri nel caso in esame.
- iv) Si valuti la validità del modello ottenuto.

- v) Si commenti il significato dell'indice ottenuto al punto precedente nel caso in esame.

ESERCIZIO 15 (15 punti)

E' stato rilevato per 6 famiglie l'ammontare della spesa annua (in euro) per l'acquisto di due generi di largo consumo: latte fresco e biscotti.

Famiglia	Spesa annua per l'acquisto di latte fresco (€)	Spesa annua per l'acquisto di biscotti (€)
A	105	65
B	190	130
C	80	160
D	120	90
E	240	220
F	60	50

- i) Si calcoli il coefficiente di correlazione lineare tra la spesa per l'acquisto di latte fresco e quella per l'acquisto di biscotti.
- ii) Si commenti il significato dell'indice ottenuto a punto precedente nel caso in esame.
- iii) Si tracci il diagramma di dispersione (in termini di valori originari).
- iv) Si dica se le informazioni che fornisce il diagramma ottenuto al punto precedente sono in accordo con il valore del coefficiente di correlazione.
- v) Si dica perché, dal punto di vista logico, nel problema in esame il calcolo del coefficiente di correlazione è preferibile all'adattamento di una retta di regressione

ESERCIZIO 16 (16 punti)

Nella tabella che segue sono riportate le serie storiche del tasso di indebitamento delle famiglie, in percentuale, (X) e del fabbisogno di energia elettrica, in migliaia di megawatt, (Y) in Italia nel periodo 1998 – 2002 (fonte: Mark up e Il Mondo):

anni	X	Y
1998	27,8	279
1999	31,1	286
2000	32,6	299
2001	32,6	305
2002	35,1	311

- i) Conoscendo il valore dei seguenti indici:
 $COV(X,Y) = 26,82$ $VAR(X) = 5,73$ $VAR(Y) = 140,80$
 si calcoli il coefficiente di correlazione lineare tra le due serie storiche.
- ii) Si illustri il significato del coefficiente calcolato al punto precedente nel caso in esame.
- iii) Si dica, motivando la risposta, quali trasformazioni di ciascuna serie storica sarebbe opportuno introdurre.
- iv) Si calcoli il coefficiente di correlazione sui dati trasformati.
- v) Si illustri il significato del coefficiente calcolato al punto iv) nel caso in esame.
- vi) Si confrontino i risultati ottenuti ai punti i) e iv).

ESERCIZIO 23 (17 punti)

Con riferimento a 9 aziende dei settori della moda e del lusso, si sono considerate le seguenti variabili, relative ai risultati consolidati alla prima metà del 2004:

X = fatturato, in milioni di euro

Y = utile operativo, in milioni di euro

Z = indebitamento finanziario netto, in milioni di euro

La matrice di varianza/ covarianza tra le variabili è risultata la seguente:

	X	Y	Z
X	212.939	36.333	-184.431
Y		6.619	-31.316
Z			177.227

- i) Si determini la matrice di correlazione tra le variabili.
- ii) Si commentino i valori ottenuti nel caso in esame.
- iii) Volendo “spiegare” l’utile operativo, si individui la miglior variabile esplicativa.
- iv) Si calcoli la corrispondente retta di regressione (in termini di scostamenti dalla media).
- v) Si commenti il significato del coefficiente di regressione della retta individuata al punto iv).
- vi) Si valuti la validità del modello ottenuto.
- vii) Si commenti il significato dell’indice ottenuto al punto precedente nel caso in esame.

ESERCIZIO 17 (16 punti)

Nella che segue viene riportata la serie storica del totale dei cosiddetti “crediti problematici” che gravano sul sistema finanziario italiano (in milioni di euro).

Tempi	Totale dei “crediti problematici”
Giugno 2001	63,0
Dicembre 2001	64,3
Giugno 2002	64,5
Dicembre 2002	66,9
Giugno 2003	69,3
Dicembre 2003	71,9
Giugno 2004	74,0

- i) Assumendo come unità temporale il semestre, si calcolino i parametri della funzione interpolante lineare del totale dei crediti problematici in funzione del tempo.
- ii) Si illustri il significato dei parametri ottenuti al punto precedente con riferimento alla serie storica in esame.
- iii) Si valuti la bontà di adattamento del modello ottenuto.
- iv) Si commenti il significato dell’indice ottenuto al punto precedente con riferimento alla serie storica in esame.
- v) Senza ripetere i calcoli, si ricavino dal precedente modello i valori dei parametri della funzione interpolante in cui l’unità temporale è l’anno, mostrando tutti i passaggi effettuati.