

## *ESERCIZIO 1*

$X_1$	7	13	16	27	44	61	85	100	125
$X_2$	3	5	5	7	9	10	12	13	14
$X_3$	4	8	10	15	24	37	51	62	80

*a) Si determini la matrice di correlazione e, in base ai risultati ottenuti si indichi, giustificando la risposta, quale delle due rette ai minimi quadrati*

$$X_1 = a + bX_2 \text{ o } X_1 = c + dX_3$$

presenta il migliore adattamento.

b) Si confrontino  $r_{12.3}$  e  $r_{13.2}$  con i rispettivi coefficienti di correlazione grezzi e commentare.

c) Determinare l'equazione del piano ai minimi quadrati  $X_1 = \alpha_1 + \alpha_2 X_2 + \alpha_3 X_3$ , e valutarne la bontà di adattamento.

d) Si effettui l'analisi dei residui relativamente al piano di cui al punto c, commentando i risultati ottenuti.

$$\sum_{i=1}^9 x_{1i} = 478, \quad \bar{X}_1 = 53, \bar{1}, \quad \sum_{i=1}^9 x_{1i}^2 = 39710$$

$$\sum_{i=1}^9 x_{2i} = 78, \quad \bar{X}_2 = 8, \bar{6}, \quad \sum_{i=1}^9 x_{2i}^2 = 798$$

$$\sum_{i=1}^9 x_{3i} = 291, \quad \bar{X}_3 = 32, \bar{3}, \quad \sum_{i=1}^9 x_{3i}^2 = 15195$$

$$\sigma_{11} = \sigma^2(X_1) = 1591,4321 \quad \sigma_1 = 39,8928$$

$$\sigma_{22} = \sigma^2(X_2) = 13,5 \quad \sigma_2 = 3,6818$$

$$\sigma_{33} = \sigma^2(X_3) = 642,8 \quad \sigma_3 = 25,3553$$

$$\sum_{i=1}^9 x_{1i} x_{2i} = 5431 \quad \sigma_{12} = \sigma(X_1, X_2) = +143,148$$

$$\sum_{i=1}^9 x_{1i} x_{3i} = 24545 \quad \sigma_{13} = \sigma(X_1, X_3) = +1009,9630$$

$$\sum_{i=1}^9 x_{2i} x_{3i} = 3331 \quad \sigma_{23} = \sigma(X_2, X_3) = +89,8$$

a)  $r_{ij} = \frac{\sigma_{ij}}{\sigma_i \cdot \sigma_j}$ . Matrice di correlazione C:

	$X_1$	$X_2$	$X_3$
$X_1$	1	$r_{12}$	$r_{13}$
$X_2$	$r_{21}$	1	$r_{23}$
$X_3$	$r_{31}$	$r_{32}$	1

	$X_1$	$X_2$	$X_3$
$X_1$	1	0,9746	0,9985
$X_2$		1	0,9629
$X_3$			1

$$I_{1.2}^2 = r_{12}^2 = 0,9498$$

$$I_{1.3}^2 = r_{13}^2 = 0,9970$$

Poiché  $I_{1.3}^2 > I_{1.2}^2$  è da preferirsi la retta

$$X_1 = c + dX_3.$$

b)

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}} = 0,8851 < r_{12} = 0,9746$$

Rispetto al coefficiente di correlazione grezzo, che è positivo e di valore elevato, la correlazione tra  $X_1$  e  $X_2$  al netto del contributo

lineare di  $X_3$  si mantiene positiva ed elevata, pur riducendosi leggermente.

$$r_{13.2} = \frac{r_{13} - r_{12}r_{23}}{\sqrt{(1-r_{12}^2)(1-r_{23}^2)}} = 0,9950 < r_{13} = 0,9985.$$

La correlazione parziale tra  $X_1$  e  $X_3$  al netto del contributo lineare di  $X_2$  è elevata e dello stesso ordine di grandezza di quella grezza.

c)

$$\alpha_2 = \frac{\sigma_{12}\sigma_{33} - \sigma_{13}\sigma_{23}}{\sigma_{22}\sigma_{33} - \sigma_{23}^2} = 1,9598$$

$$\alpha_3 = \frac{\sigma_{13}\sigma_{22} - \sigma_{12}\sigma_{23}}{\sigma_{22}\sigma_{33} - \sigma_{23}^2} = 1,2970$$

$$\alpha_1 = \bar{x}_1 - \alpha_2 \bar{x}_2 - \alpha_3 \bar{x}_3 = -5,8102$$

$$\begin{aligned} X_1 &= \alpha_1 + \alpha_2 X_2 + \alpha_3 X_3 = \\ &= -5,8102 + 1,9598 X_2 + 1,2970 X_3 \end{aligned}$$

$$I_{1.23}^2 = \frac{\text{Varianza Spiegata}}{\text{Varianza Totale}} = \frac{\hat{\alpha}_2 \sigma_{12} + \hat{\alpha}_3 \sigma_{13}}{\sigma_{11}} = 0,9994$$

adattamento molto buono del piano interpolatore.

Variazione della varianza spiegata dalla miglior retta  $X_1 = c + dX_3$  al piano ai minimi quadrati: (la differenza indica la frazione della varianza totale che viene spiegata nel passare dalla retta al piano)

$$I_{12.3}^2 - I_{13}^2 = 0,0024.$$

Grado di miglioramento come riduzione relativa della varianza residua:  $\frac{I_{1.23}^2 - I_{1.3}^2}{1 - I_{1.3}^2} = 0,8$

(indica la frazione della varianza residua dalla retta che viene spiegata nel passare dalla retta al piano).

d)

$X_1$	$\hat{X}_1$	$X_1 - \hat{X}_1$	$(X_1 - \hat{X}_1)^2$
7	5,2572	1,7428	3,0374
13	14,3648	-1,3648	1,8627
16	16,9588	-0,9588	0,9193
27	27,3634	-0,3634	0,1321
44	42,956	1,044	1,0900
61	61,7768	-0,7768	0,6034
85	83,8544	1,1456	1,3124

100	100,0812	-0,0812	0,0066
125	125,387	-0,387	0,1498
<b>478</b>	<b>477,9996</b>		<b>9,1137</b>

$$\sum_{i=1}^9 |x_{1i} - \hat{x}_{1i}| = 7,8644$$

$$A_1 = \frac{1}{N} \sum_{i=1}^9 |x_{1i} - \hat{x}_{1i}| = \frac{7,8644}{9} = 0,8738 < 1$$

$$A_1' = \frac{A_1}{\bar{x}_1} = \frac{0,8738}{53,1} = 0,0165.$$

L'ordine medio di grandezza dei residui in valore assoluto è l'1,65% del valore medio di  $\bar{X}_1$ .

$$A_2' = \frac{A_2}{\bar{x}_1} = \frac{\left[ \frac{1}{N} \sum (x_{1i} - \hat{x}_{1i})^2 \right]^{\frac{1}{2}}}{\bar{x}_1} =$$

$$= \frac{\sqrt{\text{Var Residua}}}{\bar{x}_1} = 0,0189.$$

**ESERCIZIO:** Condurre la medesima analisi per la retta  $\hat{X}_1 = c + dX_3$  e confrontare gli indici  $A_1, A_1', A_2, A_2'$  con quelli del piano.

## ESERCIZIO 2 *Tema d'esame del 12.06.01.*

Su un collettivo di Paesi vengono rilevate le seguenti variabili:

$X_1$  = *speranza di vita della popolazione femminile*

$X_2$  = *speranza di vita della popolazione maschile*

$X_3$  = *calorie medie assunte giornalmente dalla popolazione*

$X_4$  = *percentuale della popolazione alfabetizzata.*

Si ottiene la seguente matrice di correlazione:

	$X_1$	$X_2$	$X_3$	$X_4$
$X_1$	1	0,982	0,775	0,865
$X_2$		1	0,765	0,809
$X_3$			1	0,682
				1

e le seguenti informazioni

	<i>Media</i>	$\sigma$
$X_1$	70,16	10,57
$X_2$	64,92	9,27
$X_3$	2753,83	567,83
$X_4$	78,34	22,88

- a) Si determinino i parametri della retta interpolante  $\hat{X}_1 = a + \alpha_{12}X_2$  e se ne fornisca la relativa interpretazione.
- b) Si determinino i parametri del piano interpolante  $\hat{X}_1 = b + \alpha_{12.3}X_2 + \alpha_{13.2}X_3$  e se ne fornisca la relativa interpretazione. Si confrontino  $\alpha_{12}$  e  $\alpha_{12.3}$ .
- c) Si calcoli la varianza spiegata del piano interpolante di cui al punto b e si valuti la bontà di adattamento del piano.
- d) Si calcolino i coefficienti di correlazione parziale  $r_{12.3}$  e  $r_{13.2}$  fornendone l'interpretazione e confrontandoli con i rispettivi coefficienti grezzi.
- e) Si valuti il miglioramento che si ottiene passando dal piano  $\hat{X}_1 = b + \alpha_{12.3}X_2 + \alpha_{13.2}X_3$  all'interpolante che include anche  $X_4$  tra le variabili esplicative.

### SOLUZIONI:

a)

$$\alpha_{12} = \frac{\sigma_{12}}{\sigma_{22}} = \frac{\sigma(X_1, X_2)}{\sigma^2(X_2)} = \frac{(0,982)(10,57)(9,27)}{(9,27)^2} = 1,1197$$



$$a = \overline{X}_1 - \alpha_{12} \overline{X}_2 = 70,16 - (1,1197)(64,92) = -2,5309$$

$$\hat{X}_1 = -2,5309 + 1,1197 X_2$$

Il parametro  $\alpha_{12}$  della retta indica la variazione di  $\hat{X}_1$  che si ha in corrispondenza di un incremento unitario di  $X_2$  al lordo delle variazioni di  $X_3$ .

b)

$$\sigma_{13} = r_{13} \cdot \sigma_1 \cdot \sigma_3 = (0,775)(10,57)(567,83) = 4651,5214$$

$$\sigma_{23} = r_{23} \cdot \sigma_2 \cdot \sigma_3 = (0,765)(9,27)(567,83) = 4026,7948$$

$$\sigma_{11} = (10,57)^2 = 111,7249$$

$$\sigma_{33} = (567,83)^2 = 322430,9089$$

$$\hat{X}_1 = b + \alpha_{12.3} X_2 + \alpha_{13.2} X_3$$

dove

$$\alpha_{12.3} = \frac{\sigma_{33} \sigma_{12} - \sigma_{13} \sigma_{23}}{\sigma_{22} \sigma_{33} - \sigma_{23}^2} = 1,0697$$

$$\alpha_{13.2} = \frac{\sigma_{22} \sigma_{13} - \sigma_{12} \sigma_{23}}{\sigma_{22} \sigma_{33} - \sigma_{23}^2} = 0,0011$$

$$b = \overline{X}_1 - \alpha_{12.3} \overline{X}_2 - \alpha_{13.2} \overline{X}_3 = -2,3076$$

$$\hat{X}_1 = -2,3076 + 1,0697 X_2 + 0,0011 X_3$$

Il parametro  $\alpha_{12.3}$  (*coefficiente di regressione netto*) del piano indica la variazione di  $\hat{X}_1$  che si ha in corrispondenza di un incremento unitario di  $X_2$  se  $X_3$  resta costante. Analogamente ha il significato il parametro  $\alpha_{13.2}$ .

c) *Varianza spiegata*:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N (\hat{x}_{1i} - \bar{x})^2 &= \text{Var}(\hat{X}_1) = \alpha_{12.3} \sigma_{21} + \alpha_{13.2} \sigma_{31} = \\ &= 108,0434. \end{aligned}$$

Bontà di adattamento del piano interpolatore:

$$\begin{aligned} I_{1.23}^2 &= \frac{\text{Varianza spiegata}}{\text{Varianza totale}} = \frac{\frac{1}{N} \sum_{i=1}^N (\hat{x}_{1i} - \bar{x})^2}{\frac{1}{N} \sum_{i=1}^N (x_{1i} - \bar{x})^2} = \\ &= \frac{108,0434}{111,7249} = 0,9670. \end{aligned}$$

Il piano interpolatore spiega circa il 96,7% della variabilità totale della speranza di vita della popolazione femminile.

d)

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{\{1 - r_{23}^2\} \cdot \{1 - r_{13}^2\}}} = 0,9561 < r_{12} = 0,982$$

Se le variabili  $X_1$  e  $X_2$  fossero incorrelate con  $X_3$  la correlazione parziale  $r_{12.3}$  sarebbe uguale a quella grezza  $r_{12}$ . Negli altri casi può esservi notevole differenza.

$$r_{24.3} = \frac{r_{24} - r_{23}r_{34}}{\sqrt{\{1 - r_{23}^2\} \cdot \{1 - r_{34}^2\}}} = 0,6099$$
$$< r_{24} = 0,809$$

La correlazione parziale tra  $X_2$  e  $X_4$  al netto del contributo lineare di  $X_3$  è minore di quella grezza.

e) Riduzione relativa della varianza residua che si ha nel passare dal piano ai minimi quadrati

$$\hat{X}_1 = \alpha + \alpha_{12.3}X_2 + \alpha_{13.2}X_3$$

all'iperpiano ai minimi quadrati

$$\hat{X}_1 = \alpha + \alpha_{12.34}X_2 + \alpha_{13.24}X_3 + \alpha_{14.23}X_4:$$

$$\begin{aligned} r_{14.23}^2 &= \frac{(r_{14.3} - r_{12.3} \cdot r_{24.3})^2}{(1 - r_{12.3}^2)(1 - r_{24.3}^2)} = \\ &= \frac{[0,7279 - (0,9561)(0,6099)]^2}{(1 - 0,9561^2)(1 - 0,6099^2)} = 0,3886 \end{aligned}$$

essendo

$$r_{14.3} = \frac{r_{14} - r_{13}r_{34}}{\sqrt{(1 - r_{13}^2)(1 - r_{34}^2)}} = 0,7279.$$

La diminuzione relativa della varianza residua che si è avuta aggiungendo  $X_4$  rappresenta circa il 38,86% della variabilità residua del modello  $\hat{X}_1 = \alpha + \alpha_{12.3}X_2 + \alpha_{13.2}X_3$ .

### ESERCIZIO 3

Una società produttiva di beni di largo consumo ha recentemente lanciato sul mercato

un nuovo prodotto. Volendo analizzare i risultati delle vendite sono state raccolte le informazioni seguenti in 20 aree geografiche:

$X_1 =$  vendite del nuovo prodotto nell'ultimo mese (migliaia di euro).

$X_2 =$  spese per la campagna pubblicitaria (migliaia di euro).

$X_3 =$  numero di punti vendita

$X_4 =$  quota di mercato di una società concorrente (in percentuale).

Sulla base di tali informazioni sono state ricavate le medie aritmetiche:

$$\overline{X_1} = 145 \quad \overline{X_2} = 37,89 \quad \overline{X_3} = 5,31 \quad \overline{X_4} = 25,8;$$

la matrice di varianza e covarianza:

	$X_1$	$X_2$	$X_3$	$X_4$
$X_1$	2131,78	488,04	81,18	213,69
$X_2$		122,06	21,39	62,87
$X_3$			4,28	12,54
$X_4$				45,01

a) Si calcoli il coefficiente di correlazione parziale  $r_{14.2}$  e lo si confronti con il corrispondente coefficiente di correlazione grezzo, commentando il risultato ottenuto.

b) Calcolare con il metodo dei minimi quadrati i parametri dei modelli

$$1. \hat{X}_1 = a + \alpha_2 X_2$$

$$2. \hat{X}_1 = b + \alpha_{12.4} X_2 + \alpha_{14.2} X_4$$

c) Valutare la bontà di adattamento di ciascuno dei modelli previsti al punto b e il miglioramento ottenuto passando dalla retta al piano.

d) Sapendo che per il modello

$$3. \hat{X}_1 = c + \alpha_{12.3} X_2 + \alpha_{13.2} X_3$$

risulta

$$\sum_{i=1}^{20} (x_{1i} - \hat{x}_{1i})^2 = 2968,52$$

scegliere tra il modello 2. e il modello 3. quello che presenta il migliore adattamento.

## SOLUZIONI:

### a) *Matrice di correlazione*

	$X_1$	$X_2$	$X_3$	$X_4$
$X_1$	1	0,9567	0,8499	0,6899
$X_2$		1	0,9358	0,8482
$X_3$			1	0,9035
$X_4$				1

$$r_{14.2} = \frac{r_{14} - r_{12}r_{42}}{\sqrt{\{1 - r_{12}^2\} \cdot \{1 - r_{42}^2\}}} = -0,7886$$

Il coefficiente di correlazione parziale  $r_{14.2}$  informa che, al netto dell'influenza della spesa per la campagna pubblicitaria  $X_2$ , all'aumentare di  $X_4$  tende a diminuire  $X_1$ . Osserviamo differenza di segno tra  $r_{14.2}$  e  $r_{14}$ .

b)

$$\alpha_2 = \frac{\sigma_{12}}{\sigma_{22}} = \frac{Cov(X_1, X_2)}{\sigma^2(X_2)} = \frac{488,04}{122,06} = 3,9984$$

$$a = \bar{x}_1 - \alpha_2 \bar{x}_2 = -6,4992$$

$$\boxed{\hat{X}_1 = -6,4992 + 3,9984 X_2}$$

$$\alpha_{12.4} = \frac{\sigma_{44}\sigma_{12} - \sigma_{14}\sigma_{24}}{\sigma_{22}\sigma_{44} - \sigma_{24}^2} = 5,5356$$

$$\alpha_{14.2} = \frac{\sigma_{22}\sigma_{14} - \sigma_{12}\sigma_{24}}{\sigma_{44}\sigma_{22} - \sigma_{24}^2} = -2,9846$$

$$b = \bar{x}_1 - \alpha_{12.4}\bar{x}_2 - \alpha_{14.2}\bar{x}_4 = 12,2588$$

$$\hat{X}_1 = 12,2588 + 5,5356X_2 - 2,9846X_4$$

c) Per il modello 1.  $I_{1.2}^2 = r_{12}^2 = 0,9153$ .

Per il modello 2. si ha

$$\frac{\text{Varianza Spiegata}}{\text{Varianza Totale}} = \frac{\text{Var}(\hat{X}_1)}{\text{Var}(X_1)} = \frac{\alpha_{12.4}\sigma_{21} + \alpha_{14.2}\sigma_{14}}{\sigma_{11}} =$$

$$= I_{1.24}^2 = \frac{2063,8151}{2131,78} = 0,9681.$$

Miglioramento:  $I_{1.24}^2 - I_{1.2}^2 = 0,0528$  (indica la frazione di varianza totale che viene spiegata nel passare dalla retta 1. al piano 2.).

Grado di miglioramento:  $\frac{I_{1.24}^2 - I_{1.2}^2}{1 - I_{1.2}^2} = 0,6233$

(il 62,33% della varianza residua della retta viene spiegata nel passare dalla retta 1. al piano 2.)



d)Modello 3.

$$\frac{\text{Varianza Spiegata}}{\text{Varianza Totale}} = \frac{\text{Var}(\hat{X}_1)}{\text{Var}(X_1)} = \frac{\alpha_{12.3}\sigma_{21} + \alpha_{13.2}\sigma_{13}}{\sigma_{11}} =$$
$$= I_{1.23}^2 = \frac{1876,76}{2131,78} = 0,8804.$$

Si sceglie il modello 2. poiché presenta un migliore adattamento.

#### *ESERCIZIO 4 (LINEARIZZAZIONE)*

*Si consideri il seguente insieme di valori*

$X_1$	53	36	65	44	31	30	14	2	3
$X_2$	1	2	3	4	5	6	7	8	9
$X_3$	3	5	4	6	8	9	12	19	20

a) Si calcolino i coefficienti di correlazione parziale  $r_{12.3}$  e  $r_{13.2}$  e li si confronti con i rispettivi coefficienti grezzi.

b) Si determinino i parametri del modello  $\hat{X}_1 = \alpha\beta^{X_2}\gamma^{X_3}$ .

c) Si valuti la bontà di adattamento del modello ai dati.

d) Si effettui l'analisi (puntuale e grafica) dei residui del modello di cui al punto b evidenziando eventuali tendenziosità.

## SOLUZIONI:

a)

$$\hat{X}_1 = \alpha \beta^{X_2} \gamma^{X_3}$$

$$\ln \hat{X}_1 = \ln(\alpha \beta^{X_2} \gamma^{X_3})$$

$$\ln \hat{X}_1 = \ln \alpha + X_2 \ln \beta + X_3 \ln \gamma$$

$$\hat{Y}_1 = a + bX_2 + cX_3$$

dove

$$\hat{Y}_1 = \ln X_1 \quad a = \ln \alpha \quad b = \ln \beta \quad c = \ln \gamma.$$

Ci si è ricondotti a una funzione lineare nei parametri  $a, b, c$ , che prevede i logaritmi di  $X_1$  in funzione lineare di  $X_2$  e  $X_3$ . Si ottiene la seguente tabella:

$Y_1$	3,97	3,58	4,17	3,78	3,43	3,40	2,64	0,69	1,10
$X_2$	1	2	3	4	5	6	7	8	9
$X_3$	3	5	4	6	8	9	12	19	20

$$\bar{Y}_1 = 2,97 \quad \bar{X}_2 = 5 \quad \bar{X}_3 = 9,5$$

$$\text{Var}(Y_1) = \sigma_{11} = 1,43$$

$$\text{Var}(X_2) = \sigma_{22} = 6,6$$

$$\text{Var}(X_3) = \sigma_{33} = 34,91$$

$$\text{Cov}(Y_1, X_2) = \sigma_{12} = -2,6235$$

$$\text{Cov}(Y_1, X_3) = \sigma_{13} = -6,86198$$

$$\text{Cov}(X_2, X_3) = \sigma_{23} = 14,3$$

$$b = \frac{\sigma_{33}\sigma_{12} - \sigma_{13}\sigma_{23}}{\sigma_{22}\sigma_{33} - \sigma_{23}^2} = 0,25 = \ln \beta \rightarrow \beta = \exp(b)$$

$$c = \frac{\sigma_{22}\sigma_{13} - \sigma_{12}\sigma_{23}}{\sigma_{22}\sigma_{33} - \sigma_{23}^2} = -0,30 = \ln(\gamma) \rightarrow \gamma = \exp(c)$$

$$a = \bar{y}_1 - b\bar{x}_2 - c\bar{x}_3 = 4,59 = \ln \alpha \rightarrow \alpha = \exp(a)$$

$$\boxed{\hat{Y}_1 = 4,59 + 0,25X_2 - 0,30X_3}$$

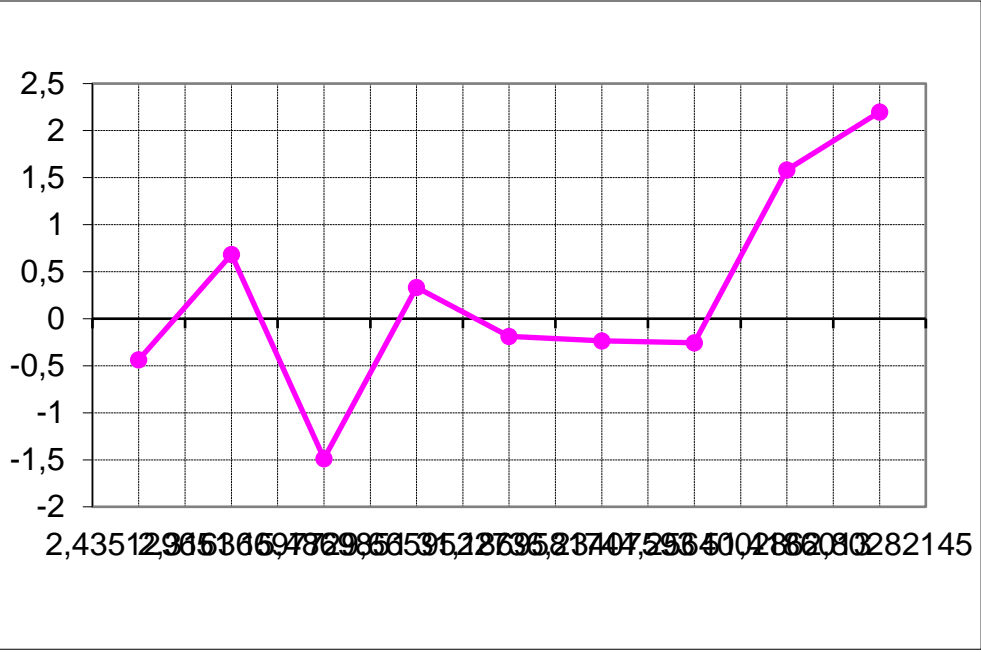
Sostituendo le coppie  $(x_{2i}, x_{3i})$  nel piano si ottengono i valori di  $\hat{Y}_1 = \ln X_1$  e infine i valori di  $X_1 = \exp(\hat{Y}_1)$  secondo quanto in tabella.

$X_1$	$\hat{X}_1$	$X_1 - \hat{X}_1$
2	2,43513	-0,43513
3	2,316367	0,683633
14	15,48699	-1,48699
30	29,66595	0,334048
31	31,18696	-0,18696
36	36,23408	-0,23408
44	44,2564	-0,2564
53	51,4186	1,581399
65	62,80282	2,197179

$$A_1 = \frac{1}{N} \sum_{i=1}^9 |\hat{x}_{1i} - x_{1i}| = 0,822$$

$$A_1' = \frac{A_1}{\bar{x}_1} = \frac{0,822}{30,8}$$

Come si può vedere dal grafico seguente i residui non presentano tendenziosità.



## *ESERCIZIO 5 (LINEARIZZAZIONE)*

Si consideri il seguente insieme di valori:

$X_1$	$X_2$	$X_3$
145	1	30
82	2	21
78	3	18
300	4	41
194	5	32
161	6	29
310	7	38
159	8	27
132	9	23
107	10	19

$X_1$ : spese annue di manutenzione di un veicolo  
(in Euro)

$X_2$ : anno di vita del veicolo

$X_3$ : chilometri percorsi/anno (in migliaia)

- a) Si determinino i parametri del modello  $\hat{X}_1 = \alpha \cdot \beta^{X_2} \cdot \gamma^{X_3}$  e se ne dia il significato statistico.
- b) Si valuti la bontà d'accostamento del modello e s'interpreti il risultato ottenuto.
- c) Si svolga l'analisi grafica dei residui, precisando se esista tendenziosità e se vi siano dati anomali.

*Soluzione*

$$\hat{X}_1 = \alpha \cdot \beta^{X_2} \cdot \gamma^{X_3}$$
$$\ln \hat{X}_1 = \ln(\alpha \cdot \beta^{X_2} \cdot \gamma^{X_3}) =$$
$$= \ln \alpha + (\ln \beta)X_2 + (\ln \gamma)X_3$$

$$\ln X_1 \equiv Y_1$$

$$\hat{Y}_1 \equiv A + BX_2 + CX_3$$

Ci si è ricondotti ad un modello lineare nei parametri che prevede i logaritmi di  $X_1$  in funzione lineare di  $X_2$  e  $X_3$ .

Pertanto si origina la seguente matrice dei dati:

$Y_1$	$X_2$	$X_3$
4,9767	1	30
4,4067	2	21
4,3567	3	18
5,7038	4	41
5,2679	5	32
5,0814	6	29
5,7366	7	38
5,0689	8	27
4,8828	9	23
4,6728	10	19
50,1543	55	278

$$\bar{Y}_1 = 5,01543 \quad \bar{X}_2 = 5,5 \quad \bar{X}_3 = 27,8$$



$Y_1^2$	$X_2^2$	$X_3^2$	$Y_1 \cdot X_2$	$Y_1 \cdot X_3$	$X_2 \cdot X_3$
24,7675	1	900	4,9767	149,3010	30
4	4	441	8,8134	92,5407	42
19,4190	9	324	13,0701	78,4206	54
1	16	168	21,4268	233,8558	164
18,9808	25	1	26,3395	168,5728	160
3	36	102	30,4884	147,3606	174
32,5333	49	4	40,1562	217,9908	266
3	64	841	40,5512	136,8603	216
27,7507	81	144	43,9452	112,3044	207
7	10	4	46,7280	88,7832	190
25,8206	0	729			
3		529			
32,9085		361			
8					
25,6937					
4					
23,8417					
4					
21,8350					
6					
253,551	38	827	276,495	1425,990	1503
2	5	4	5	2	

$$\text{Var}(Y_1) = \sigma_{11} = 25,35512 - (5,01543)^2 = 0,2006$$

$$\text{Var}(X_2) = \sigma_{22} = 38,5 - (5,5)^2 = 8,25$$

$$\text{Var}(X_3) = \sigma_{33} = 827,4 - (27,8)^2 = 54,56$$

$$\begin{aligned} \text{Cov}(Y_1, X_2) &= \sigma_{12} = 27,64955 - 5,01543 \cdot 5,5 = \\ &= 0,0647 \end{aligned}$$

$$\begin{aligned} \text{Cov}(Y_1, X_3) &= \sigma_{13} = 142,59902 - 5,01543 \cdot 27,8 = \\ &= 3,1701 \end{aligned}$$

$$\text{Cov}(X_2, X_3) = \sigma_{23} = 150,3 - 5,5 \cdot 27,8 = -2,6$$

$$\hat{B} = \frac{0,0647 \cdot 54,56 - 3,1701 \cdot (-2,6)}{8,25 \cdot 54,56 - (-2,6)^2} = \frac{11,7723}{444,1025} =$$

$$= 0,0265$$

$$\hat{C} = \frac{3,1701 \cdot 8,25 - 0,0647 \cdot (-2,6)}{444,1025} = \frac{26,3215}{444,1025} =$$

$$= 0,0593$$

$$\hat{A} = 5,01543 - 0,0265 \cdot 5,5 - 0,0593 \cdot 27,8 = 3,2211$$

$$\hat{Y}_1 \equiv 3,2211 + 0,0265 \cdot X_2 + 0,0593 \cdot X_3$$

Sostituendo le coppie  $(x_{2i}, x_{3i})$  nel piano si ottengono i logaritmi naturali dei valori

previsti e, infine, i valori previsti, secondo quanto in tabella.

$X_1$	$\hat{X}_1$	$X_1 - \hat{X}_1$
145	152,4139	-7,4139
82	92,3606	-10,3606
78	78,8857	-0,8857
300	316,8410	-16,8410
194	190,7951	+3,2049
161	163,9891	-2,9891
310	287,1486	+22,8514
159	153,5767	+5,4233
132	124,3997	+7,6003
107	100,7659	+6,2341

$$A_1 = \frac{1}{N} \sum_{i=1}^N |x_{1i} - \hat{x}_{1i}| = \frac{83,8043}{10} = 8,3804$$

$$A'_1 = \frac{A_1}{\bar{x}_1} = \frac{8,3804}{166,8} = 0,0502$$

$X_1$	$\hat{X}_1$	$X_1 - \hat{X}_1$
78	78,8857	-0,8857
82	92,3606	-10,3606
107	100,7659	+6,2341
132	124,3997	+7,6003
145	152,4139	-7,4139
159	153,5767	+5,4233
161	163,9891	-2,9891
194	190,7951	+3,2049
300	287,1486	+22,8514
310	316,8410	-16,8410

Non vi è tendenziosità, né dati anomali.  
*(L'analisi grafica è lasciata per esercizio)*

## *Esempio*

Si considerino i dati in tabella:

Anno	$X_1$	$X_2$	$X_3$
1948	100	100	100
1949	106	104	99
1950	107	106	110
1951	120	111	126
1952	110	111	113
1953	116	115	103
1954	123	120	102
1955	133	124	103
1956	137	126	98

$X_1$ : numero indice delle importazioni di beni e servizi nel Regno Unito, a prezzi costanti (1948)

$X_2$ : numero indice del prodotto lordo del Regno Unito ai prezzi 1948

$X_3$ : rapporto tra indici dei prezzi delle importazioni e del prodotto totale del Regno Unito.

$$n = 9 \quad \sum x_{1i} = 1052 \quad \sum x_{2i} = 1017 \quad \sum x_{3i} = 954$$

$$\bar{x}_1 = 116,9 \quad \bar{x}_2 = 113 \quad \bar{x}_3 = 106$$

$$\sum x_{1i}^2 = 124228 \quad \sum x_{2i}^2 = 115571 \quad \sum x_{3i}^2 = 101772$$

$$\sum x_{1i}x_{2i} = 119750 \quad \sum x_{1i}x_{3i} = 111433$$

$$\sum x_{2i}x_{3i} = 107690.$$

$$\sigma_{11} = 140,09\bar{8} \quad \sigma_{22} = 72,2\bar{2} \quad \sigma_{33} = 72$$

$$\sigma_{12} = 97,1\bar{1} \quad \sigma_{13} = -8,7\bar{7} \quad \sigma_{23} = -12,4\bar{4}$$

ottiene il seguente piano:

$$\hat{X}_1 = -49,3297 + 1,3642X_2 + 0,1139X_3$$

$$I_{1.23}^2 = \frac{\hat{\alpha}_2\sigma_{12} + \hat{\alpha}_3\sigma_{13}}{\sigma_{11}} = \frac{1,3642 \cdot 97,1\bar{1} - 0,1139 \cdot 12,4\bar{4}}{140,09\bar{8}} =$$

$$= 0,9355$$

$$H_0 : \alpha_2 = \alpha_3 = 0$$

$$H_1 : \alpha_2 \neq 0 \cup \alpha_3 \neq 0$$

$$V = \frac{I_{1.23}^2/2}{(1 - I_{1.23}^2)/(9 - 3)} = 43,51 > F_{0,01}(2,6) = 10,925$$

Rifiuto l'ipotesi nulla. I coefficienti di regressione sono congiuntamente significativamente diversi da zero.

$$\sum (x_{1i} - \hat{x}_{1i})^2 = n\sigma_{11}(1 - I_{1.23}^2) = 9 \cdot 140,09\bar{8}(1 - 0,9355) =$$
$$= 81,3276$$

$$\hat{\sigma}^2 = \frac{81,3276}{9 - 3} = 13,5546$$

$$H_0 : \alpha_2 = 0$$

$$H_1 : \alpha_2 \neq 0$$

$$\begin{aligned} \text{Var}(\hat{\alpha}_2) &= \frac{\sigma_{33}}{n \cdot (\sigma_{22}\sigma_{33} - \sigma_{23}^2)} \cdot \sigma^2 = \\ &= \frac{72}{9(72,2 \cdot 72 - 12,4^2)} \sigma^2. \end{aligned}$$

Si rifiuta  $H_0$  se

$$\frac{|1,3642 - 0|}{\sqrt{13,5546 \cdot 0,001585}} = 9,3072 > t_{1-\frac{\alpha}{2}}(6)$$

$$t_{0,9995}(6) = 5,959.$$

*Il coefficiente di regressione di  $X_2$  è significativamente diverso da zero.*