

## Coefficiente di determinazione, analisi dei residui

1. Per misurare la quantità di proteine in un campione di fegato, i laboratori medici spesso si servono del fatto che la quantità di proteine è correlata alla proporzione di luce assorbita dal campione. Il campione di fegato viene illuminato con luce emessa da uno spettrometro e la proporzione di luce assorbita è utilizzata per stimare la quantità di proteine. Questa procedura viene sperimentata su 5 campioni il cui contenuto di proteine è noto, ottenendo i seguenti dati

Luce assorbita	Quantità di proteine (mg)
0,44	2
0,82	16
1,20	30
1,61	46
1,83	55

- (a) Calcola il coefficiente di determinazione.
  - (b) Pensi che questo sia un modo ragionevole di stimare la proporzione di proteine in un campione di fegato?
  - (c) Qual è la stima della proporzione di proteine quando la luce assorbita è 1,5?
2. I dati che seguono mettono in relazione le età delle mogli e dei mariti al momento del matrimonio.

Età della moglie	18	24	40	33	30	25
Età del marito	21	29	51	30	36	25

- (a) Determinate la retta di regressione considerando l'età della moglie come variabile esplicativa  $x$  e l'età del marito come variabile di risposta  $y$ .
  - (b) Determinate la retta di regressione considerando l'età del marito come variabile esplicativa  $x$  e l'età della moglie come variabile di risposta  $y$ .
  - (c) Determinate i coefficienti di determinazione e di correlazione campionaria nei due casi.
3. I dati seguenti riguardano i trapianti di cuore effettuati a Stanford. Essi mettono in relazione l'età con il coefficiente di rigetto, che viene utilizzato come indicatore di quanto il cuore venga accettato dall'organismo del trapiantato.

Età	Coefficiente di rigetto
51,0	1,32
54,1	0,87
54,6	1,89
42,5	0,61
43,4	1,13
49,5	1,12
55,3	2,76
64,5	0,69
45,0	1,58
52,0	1,62
58,4	0,96
42,8	1,38

- (a) Determina la retta di regressione.
- (b) Determina la percentuale di variabilità del coefficiente di rigetto spiegata dal modello di regressione.
- (c) Stima il coefficiente di rigetto di un trapiantato di 50 anni.
4. Si vuole misurare la crescita delle radici del mais in funzione del contenuto di saccarosio nel nutrimento. I dati si riferiscono alla lunghezza in mm della radice del mais (Y), coltivata in vitro per 10 giorni, in relazione al contenuto di saccarosio nel terreno di coltura (X), misurato in g/l,

X (g/l)	5	10	15	20	25	30	35
Y (mm)	33	4	62	56	74	71	80

- (a) Determina la retta di regressione.
- (b) Qual è la percentuale di variabilità dei valori variabile Y (lunghezza della radice) spiegata da modello di regressione lineare?
- (c) Stima la lunghezza della radice corrispondente a un contenuto di saccarosio nel terreno pari a 17 g/l.
- (d) Disegna il grafico dei residui. Cosa puoi concludere sul modello di regressione lineare?
5. I dati seguenti si riferiscono alla statura, misurata in cm, e al peso, espresso in Kg, di 10 individui

statura (cm)	161	162	164	165	172	179	181	184	186	191
peso (kg)	69	58	55	65	73	76	69	77	80	80

- (a) Determina la retta di regressione usando la statura come variabile esplicativa e il peso come variabile di risposta.
- (b) Calcola i residui. A quale valore della statura corrisponde il residuo più grande in valore assoluto?
- (c) Determina la retta di regressione usando il peso come variabile esplicativa e la statura come variabile di risposta.

(d) Calcola i residui. A quale valore del peso corrisponde il residuo più grande in valore assoluto?

6. I dati che seguono riportano le lunghezze delle zampe (x) e del torace (y) in mm, di due specie di insetti

zampe (mm)	torace (mm)
10	7,46
8	6,77
15	19,74
9	7,11
11	7,81
14	8,84
22	5,50
4	7,39
12	8,15
7	6,42
5	5,73

(a) Disegna il grafico dei residui. Hanno un andamento casuale?

(b) Ci sono osservazioni che distano dalle altre? Quali?

(c) Quanto valgono i residui corrispondenti?

(d) Si tratta di outlier o osservazioni influenti?