

PROBLEMES TEMA 2

Exemple Es disposa de $N=13$ dades, que expressen la relació entre la pressió de vapor de β -trimetilborazol (Y , en mm de mercuri) i la temperatura (X , en graus centígrads):

X	13.0	19.5	22.5	27.2	31.8	38.4	45.7	56.1	64.4	71.4	80.5	85.7	91.5
Y	2.9	5.1	8.5	10.3	14.6	21.3	30.5	51.4	74.5	10.2	143.7	176.9	216.9

1. Decidiu, gràficament, si hi ha algun outlier.

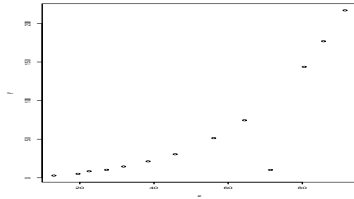
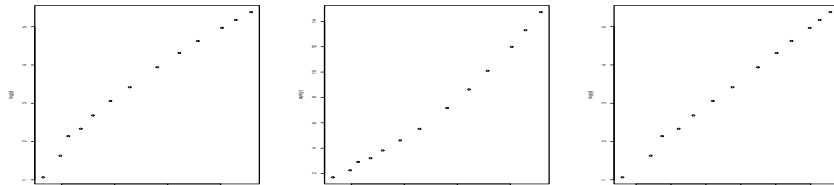


Figura 1: Diagrama bivariant

Hi ha un outlier, el punt (71.4,10.2), sembla un error de transcripció, seria (71.4,102).

2. Dibuixeu diagrames bivariants amb les dades corregides (si hi havia outliers), per a diferents transformacions de les dades (logaritmes, arrels quadrades ...), fins que el nuvol s'acoste a una recta. En eixe cas, calculeu la recta de regressió i determineu la qualitat de l'ajust.



El millor diagrama és el de l'arrel de X (AX) i el logaritme de Y (LY). Calculem la recta de regressió:

$$LY = c + b \cdot AX,$$

$$b = \frac{N \sum ax_i ly_i - (\sum ax_i)(\sum ly_i)}{N \sum ax_i^2 - (\sum ax_i)^2} = \frac{13 \cdot 337.65 - 88.34 \cdot 44.72}{13 \cdot 647.7 - 88.34^2} = 0.712$$

$$c = \frac{(\sum ly_i - b \sum ax_i)}{N} = \frac{44.72 - 0.712 \cdot 88.34}{13} = -1.39$$

Per tant, $LY = -1.39 + 0.712 AX$.

$$R^2 = r^2 = \frac{s_{axly}}{s_{ax} s_{ly}} = \frac{\frac{\sum ax_i ly_i - N \bar{ax} \bar{ly}}{N-1}}{\sqrt{\frac{\sum ax_i^2 - N \bar{ax}^2}{N-1}} \sqrt{\frac{\sum ly_i^2 - N \bar{ly}^2}{N-1}}} = \frac{\frac{337.65 - 13 \cdot 6 \cdot 3.44}{12}}{\sqrt{\frac{647.7 - 13 \cdot 46.17}{12}} \sqrt{\frac{177.93 - 13 \cdot 11.83}{12}}} = \frac{2.8}{\sqrt{3.96 \cdot 2.01}} = 0.99 \text{ és quasi}$$

1, és un ajust molt bo.

1. La resistència del paper emprat en la fabricació d'unes caixes (Y) se sap que està relacionat amb la concentració de fusta dura en la polpa original (X). S'han tret les següents mostres, a partir de les quals volem conèixer:

- la mitjana de X i Y
- la recta de regressió de la variable Y sobre la X
- la qualitat de l'ajust
- la resistència d'una caixa fabricada amb polpa amb concentració de 2.3

X	1	1.5	1.5	1.5	2	2	2.2	2.4	2.5	2.5	2.8	2.8	3	3	3.2	3.3
Y	101.4	117.4	117.1	106.2	131.9	146.9	146.8	133.9	111	123	125.1	145.2	134.3	144.5	143.7	146.9

(Càlculs: $\bar{x}=2.325$, $\bar{y}=129.706$, $\sum x_i^2=93.66$, $\sum y_i^2=272841.3$, $\sum x_i y_i=4937.22$)
 (Sol. : $Y=93.34+15.64X$, $R^2=0.479$, 129.312)

2. El temps que triga un programa a realitzar un determinat càlcul depèn de la mida de l'arxiu tractat. En 10 observacions s'han obtingut les següents dades:

X (Kb.)	352	387	254	317	428	231	276	324	441	510
Y (segons)	22	25	20	22	28	17	19	23	25	29

Calculeu:

- la mitjana de X i Y
- la recta de regressió de la variable Y sobre la X
- la qualitat de l'ajust
- quan es trigaria a tractar un arxiu de 300 Kb.

(Càlculs: $\bar{x}=352$, $\bar{y}=23$, $\sum x_i^2=1310956$, $\sum y_i^2=5422$, $\sum x_i y_i=83895$)
 (Sol. : $Y=8.63+0.0408X$, $R^2=0.907$, 20.87)

3. Es fa un estudi de la quantitat de sucre refinat que s'obté en variar la temperatura d'un procés determinat. Les dades es mostren a la taula:

X (Temp.)	1	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
Y (Sucre)	8.1	7.8	8.5	9.8	9.5	8.9	8.6	10.2	9.3	9.2	10.5

Determineu:

- la mitjana de X i Y
- la recta de regressió de la variable Y sobre la X
- la qualitat de l'ajust
- quant sucre s'obtindria per a una temperatura d'1.25.

(Càlculs: $\bar{x}=1.5$, $\bar{y}=9.127$, $\sum x_i^2=25.85$, $\sum y_i^2=923.58$, $\sum x_i y_i=152.59$)
 (Sol. : $Y=6.414+1.809X$, $R^2=0.4999$, 8.675)

4. A partir d'una mostra de 10 peces de llautó, es vol estudiar la influència de la força de tensió (X , en lliures per polzada quadrada) sobre la duresa (Y en unitats Rockwell) del material. Les dades són:

X	64	65	66	69	73	74	76	79	80	83
Y	45	46	49	49	51	54	57	57	58	59

Determineu:

- la mitjana de X i Y
- la recta de regressió de la variable Y sobre la X
- la qualitat de l'ajust
- quina duresa es prediria per a $X=70$.

(Càlculs: $\bar{x}=72.9$, $\bar{y}=52.5$, $\sum x_i^2=53549$, $\sum y_i^2=27803$, $\sum x_i y_i=38576$)

(Sol. : $Y=-2.143+0.75X$, $R^2=0.9459$, 50.357)

5. Un dels aspectes d'un programa de protecció de residus consisteix en mesurar el contingut d'un dipòsit. La determinació del seu volum es realitza indirectament mesurant la diferència de pressió entre la part més alta i més baixa del tanc. Per la geometria del tanc, se sap que la relació entre la pressió i el volum és aproximadament lineal. Amb l'objectiu de calibrar la pressió respecte al volum, es colliquen al tanc quantitats conegudes de líquid i es prenen lectures de la pressió. Les dades són (P : pressió en Pascals, V : volum en Kilolitres):

P						V					
215	218	633	629	1034	1033	0.189	0.19	0.377	0.379	0.567	0.568
1474	1475	1925	1922	2372	2374	0.757	0.758	0.946	0.947	1.135	1.136
2377	2819	3263	3262	3268	3712	1.137	1.327	1.514	1.515	1.516	1.705

Calculeu:

- la mitjana de V i P
- la recta de regressió de la variable Volum sobre la Pressió.
- la qualitat de l'ajust
- quin seria el volum si la pressió val 2000 Pascals.

(Càlculs: $\bar{x}=1889.167$, $\bar{y}=0.9257$, $\sum x_i^2=85380065$, $\sum y_i^2=19.366$, $\sum x_i y_i=40605.07$)

(Sol. : $Y=0.1102+0.00043X$, $R^2=0.9997$, 0.97)

6. La qualitat d'un sabó és determinada pel contingut d'àcid sebàtic, que pot ser mesurat mitjançant tècniques químiques. Per a l'ús en control de la qualitat en fàbriques de sabó, s'ha suggerit determinar el percentatge d'àcid sebàtic mesurant la conductivitat elèctrica del sabó. La conductivitat és fàcil de mesurar i pot ser mesurada en el lloc de producció. A la taula següent es mostren una sèrie de mesures de la conductivitat en mS (Milli-Siemens) per a un determinat sabó i els corresponents percentatges d'àcid sebàtic.

Calculeu:

C						A					
81.3	81.3	81.3	81.3	81.3	81.3	1.20	0.90	1.00	1.08	1.03	0.98
81.3	82.2	82.2	82.2	82.2	82.2	0.88	1.75	1.50	1.70	1.80	1.34
82.2	82.2	82.2	82.3	82.3	82.3	1.44	1.49	1.24	1.52	1.52	1.67
82.3	82.3	82.3	82.3	82.3	83.0	1.67	1.35	1.50	1.30	1.45	2.10
83.0	83.0	83.0	83.0	83.0	83.0	1.95	1.85	1.90	2.35	2.22	2.00

- la mitjana de C i A
- la recta de regressió de la variable A (% àcid) sobre la C (conductivitat).
- la qualitat de l'ajust
- quin seria el % àcid si la conductivitat és 82 mS.

(Càlculs: $\bar{x}=82.2$, $\bar{y}=1.548$, $\sum x_i^2=216230.6$, $\sum y_i^2=82.6035$, $\sum x_i y_i=4078.397$)

(Sol. : $Y=-47.93+0.662X$, $R^2=0.79$, 1.43)

Al final està la sol.

7. Durant la investigació de la pol·lució d'un fiord, es prengueren diverses mostres d'aigua a diferents profunditats. Per mesurar el grau de pol·lució es determina la concentració d'una bactèria. La següent taula mostra el logaritme de la concentració de la bactèria (B) a diferents profunditats (P).

P						B					
0	0	0	0	0	0	1.95	2.42	2.56	2.08	2.15	2.42
0	0	0	0	4	4	1.90	2.15	1.95	2.42	2.15	2.15
4	4	4	4	4	4	2.15	1.95	2.18	1.95	2.42	1.78
4	4	8	8	8	8	2.15	2.26	2.15	2.26	1.78	1.90
8	8	8	8	8	8	2.15	1.78	2.15	2.56	1.78	1.95

Calculeu:

- la mitjana de P i B
- la recta de regressió de la variable B sobre la P .
- la qualitat de l'ajust
- Dibuixeu un diagrama de les variables P i B . Hi ha relació entre elles?

(Càlculs: $\bar{x}=4$, $\bar{y}=2.12$, $\sum x_i^2=800$, $\sum y_i^2=136.3354$, $\sum x_i y_i=248.24$)

(Sol. : $Y=2.197-0.019X$, $R^2=0.079$, No)

8. Es realitza un experiment per tal de determinar la duració de vida de certs circuits electrònics (Y) en funció de certa variable de fabricació X . S'han obtingut els següents resultats:

X	-10	-15	20	-10	5	5
Y	11	8	73	21	46	30

Determineu:

- (a) la mitjana de X i Y
- (b) la recta de regressió de la variable Y sobre la X
- (c) la qualitat de l'ajust
- (d) quant durarà si $X=0$.

(Càlculs: $\bar{x}=-0.833$, $\bar{y}=31.5$, $\sum x_i^2=875$, $\sum y_i^2=8971$, $\sum x_i y_i=1400$)
 (Sol. : $Y=32.99+1.788X$, $R^2=0.92$, 32.99)

9. En un treball sobre seguretat vial s'estudia la possible relació entre la velocitat d'un determinat vehicle i la seua distància de frenat. Un dels objectius de l'anàlisi és determinar si la relació entre ambdues variables és aproximadament lineal o si la velocitat està relacionada linealment amb l'arrel quadrada de la distància de frenat, com sugereix una llei física. Les dades següents mostren diferents velocitats del vehicle (en Km/h) amb les seues corresponents distàncies de frenat (en metres).

V	F			
32	3.74	5.1	4.4	
48	8.2	10.58	9.55	
64	11.67	21.46	22.95	20.95
80	36.81	39.52	37.45	35.45
97	54.48	52.17	46	50.59

- (a) Calcula la mitjana de V i F
- (b) Calcula la recta de regressió estimada de la distància de frenat (Y) sobre la velocitat (X).
- (c) Calcula la recta de regressió estimada de l'arrel quadrada de la distància de frenat (Z) sobre la velocitat.
- (d) Quina sembla més adequada?

(Càlculs: $\bar{x}=66.889$, $\bar{y}=26.171$, $\bar{z}=4.769$, $\sum x_i^2=89604$, $\sum y_i^2=17833.2$, $\sum z_i^2=471.07$, $\sum x_i y_i=38366.12$, $\sum x_i z_i=6476.414$)
 (Sol. : $Y=-24.398+0.756X$, $R^2=0.94$; $Z=-0.6498+0.081X$, $R^2=0.96$, la segona)

10. En un estudi per relacionar la longitud de la línia de la vida a la ma esquerra i la vida d'una persona, s'han obtes dades de 50 persones amb els següents resultats (X =longitud de la línia en cm; Y =edat al morir en anys):

$\sum y_i=3333$, $\sum x_i=459.9$, $\sum x_i^2=4308.57$, $\sum y_i^2=231933$, $\sum x_i y_i=30549$
 Calcula la recta de regressió i determina la qualitat de l'ajust.
 (Sol. : $Y=66.66-1.38(X-9.20)$, $R^2=0.015$)

Problema 6:
 Càlculs: 82.2, 1.52, 202731.9, 74.041, 3761.227
 Sol: $Y=-48.128+0.604X$, $R^2=0.832$, 1.4