REGRESSÃO LINEAR E AJUSTES DE CURVAS

Regressão Linear

- *A regressão linear é uma técnica estatística usada para prever o valor de uma variável com base no valor de outra variável.*
- *A ideia por trás do modelo de regressão linear é estimar uma reta que melhor descreva a relação entre variáveis e permita predições.*

Ajuste linear simples

O tipo de relação mais simples entre duas variáveis é a relação linear. Nesta relação, temos uma variável independente "x" relacionada a uma variável resposta ou dependente "y" por meio de um modelo linear, por exemplo:

$y = a + bx$

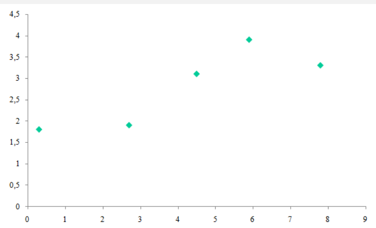
Devemos então estimar os parâmetros b_0 e b_1 para encontrarmos a relação entre x e y.

Uma etapa importante é visualizarmos o gráficos dos pontos em um diagrama de dispersão, de forma a determinarmos se há alguma relação visível entre as variáveis.

Exemplo: Dado o conjunto de pontos abaixo, trace seu diagrama de dispersão.

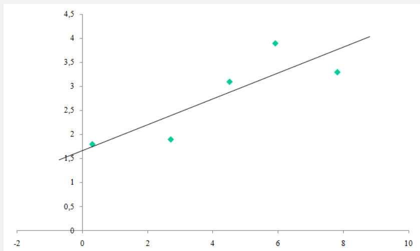
x	0,3	2,7	4,5	5,9	7,8
y	1,8	1,9	3,1	3,9	3,3

Resolução: Marcamos em um gráfico os locais dos pontos x e y:



Aparentemente há relação aproximadamente linear entre as variáveis.

Esta reta seria representada no gráfico de dispersão como:



Para estimarmos os parâmetros b_0 e b_1, devemos recorrer ao método dos quadrados mínimos.

Esta técnica estima os parâmetros de forma que a distância total entre a equação ajustada e os pontos do experimento seja a menor possível.

Exemplo: Dados os pontos do exemplo anterior, calcule a melhor reta de ajuste:

i	x	y	x^2	x.y	y^2
1	0,3	1,8	0,09	0,54	3,24
2	2,7	1,9	7,29	5,13	3,61
3	4,5	3,1	20,25	13,95	9,61
4	5,9	3,9	34,81	23,01	15,21
5	7,8	3,3	60,84	25,74	10,89
Σ	**21,2**	**14,0**	**123,28**	**68,37**	**42,56**

Estatisticamente...

$$b_1 = \frac{\sum x_i . \sum y_i - n.\sum (x_i . y_i)}{\left(\sum x_i\right)^2 - n.\sum \left(x_i^2\right)}$$

$$b_1 = \frac{21{,}2.14{,}0 - 5.68{,}37}{(21{,}2)^2 - 5.123{,}28} = 0{,}2698$$

$$b_0 = \frac{\sum y_i - b_1 . \sum x_i}{n}$$

$$b_0 = \frac{14{,}0 - 0{,}2698.21{,}2}{5} = 1{,}6560$$

y =1,6560 + 0,2698.x

Numericamente...

$$\begin{cases} n.a + \sum x.b = \sum y \\ \sum x\ .a + \sum \left(x^2\right).b = \sum (xy) \end{cases}$$

Numericamente...

$$\begin{cases} 5a + 21.2b = 14 \\ 21.2a + 123.28b = 68.37 \end{cases}$$

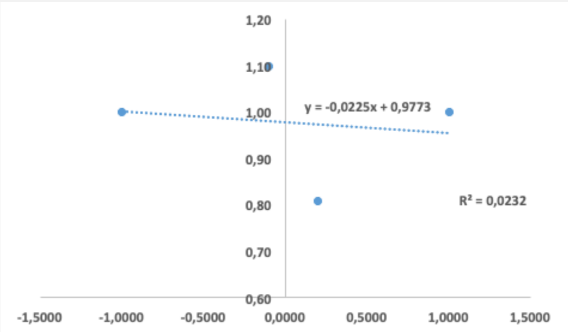
y =1,6558 + 0,2698.x

K	a	b	Erro a	Erro b	
0,0000	2,8000	0,5546	-	-	
1,0000	0,4485	0,4775	2,3515	0,0771	2,3515
2,0000	0,7756	0,4212	0,3270	0,0562	0,3270
3,0000	1,0140	0,3802	0,2385	0,0410	0,2385
4,0000	1,1879	0,3503	0,1739	0,0299	0,1739
5,0000	1,3147	0,3285	0,1268	0,0218	0,1268
6,0000	1,4071	0,3126	0,0924	0,0159	0,0924
7,0000	1,4745	0,3010	0,0674	0,0116	0,0674
8,0000	1,5237	0,2926	0,0491	0,0085	0,0491
9,0000	1,5595	0,2864	0,0358	0,0062	0,0358
10,0000	1,5856	0,2819	0,0261	0,0045	0,0261
11,0000	1,6047	0,2786	0,0190	0,0033	0,0190
12,0000	1,6186	0,2763	0,0139	0,0024	0,0139
13,0000	1,6287	0,2745	0,0101	0,0017	0,0101
14,0000	1,6361	0,2732	0,0074	0,0013	0,0074
15,0000	1,6414	0,2723	0,0054	0,0009	0,0054
16,0000	1,6454	0,2716	0,0039	0,0007	0,0039
17,0000	1,6482	0,2712	0,0029	0,0005	0,0029
18,0000	1,6503	0,2708	0,0021	0,0004	0,0021
19,0000	1,6518	0,2705	0,0015	0,0003	0,0015
20,0000	1,6530	0,2703	0,0011	0,0002	0,0011
21,0000	1,6538	0,2702	0,0008	0,0001	0,0008
22,0000	1,6544	0,2701	0,0006	0,0001	0,0006
23,0000	1,6548	0,2700	0,0004	0,0001	0,0004
24,0000	1,6551	0,2700	0,0003	0,0001	0,0003
25,0000	1,6553	0,2699	0,0002	0,0000	0,0002
26,0000	1,6555	0,2699	0,0002	0,0000	0,0002
27,0000	1,6556	0,2699	0,0001	0,0000	0,0001
28,0000	1,6557	0,2699	0,0001	0,0000	0,0001
29,0000	1,6558	0,2699	0,0001	0,0000	0,0001
30,0000	1,6558	0,2698	0,0000	0,0000	0,0000
31,0000	1,6558	0,2698	0,0000	0,0000	0,0000

Exercício 1

Determine a reta de regressão linear considerando os pontos abaixo.

x_i	-1.0	-0.1	0.2	1.0
y_i	1.000	1.099	0.808	1.000



n	X	Y	X.Y	X²
1	-1,0000	1,0000	-1,0000	1,0000
2	-0,1000	1,0990	-0,1099	0,0100
3	0,2000	0,8080	0,1616	0,0400
4	1,0000	1,0000	1,0000	1,0000
	0,1000	**3,9070**	**0,0517**	**2,0500**

$$\begin{cases} 4a + 0.1b = 3.9070 \\ 0.1a + 2.05b = 0.0517 \end{cases}$$

K	a	b	Erro a	Erro b	
0,0000	0,9768	0,0252	-	-	
1,0000	0,9761	-0,0224	0,0006	0,0476	0,0476
2,0000	0,9773	-0,0225	0,0012	0,0001	0,0012
3,0000	0,9773	-0,0225	0,0000	0,0000	0,0000

y =0,9773 - 0,0225x

Mas se os modelos não forem lineares?

$y=a.x^b \rightarrow ln(y)=ln(a)+b.ln(x)$

$y=a.b^x \rightarrow ln(y)=ln(a)+ln(b).x$

$y=a.e^{b.x} \rightarrow ln(y)=ln(a)+b.x$

$y=e^{(a+b.x1+c.x2)} \rightarrow ln(y)=a+b.x1+c.x2$

$$y = a.x_1^b.x_2^c \rightarrow \ln(y) = \ln(a) + b.\ln(x_1) + c.\ln(x_2)$$

$$y = \frac{1}{a + b.x_1 + c.x_2} \rightarrow \frac{1}{y} = a + b.x_1 + c.x_2$$

$$y = \frac{1}{1 + e^{a + b.x_1 + c.x_2}} \rightarrow \ln\left(\frac{1}{y} - 1\right) = a + b.x_1 + c.x_2$$

Ajuste a uma exponencial

$$y = \alpha e^{bx}$$

$$ln(y) = ln(\alpha e^{bx})$$

$$Y = ln(\alpha) + bx \;\; onde, \;\; ln(\alpha) = a$$

$$Y = a + bx$$

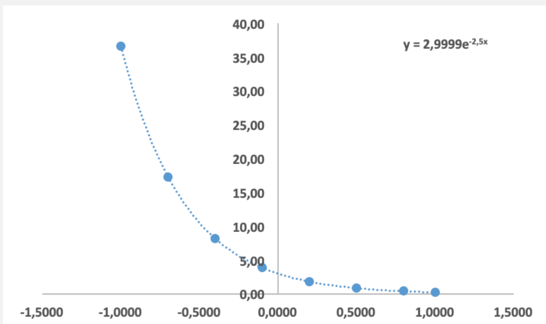
$$\alpha = e^a$$

$$\begin{cases} n.a + \Sigma x.b = \Sigma ln(y) \\ \Sigma x.a + \Sigma(x^2).b = \Sigma x.ln(y) \end{cases}$$

Exercício I

Determine a curva considerando os pontos abaixo.

x_i	-1.0	-0.7	-0.4	-0.1	0.2	0.5	0.8	1.0
y_i	36.547	17.264	8.155	3.852	1.820	0.860	0.406	0.246



n	X	Y	lny	X²	x*lny
1	-1,0000	36,5470	3,5986	1,0000	-3,5986
2	-0,7000	17,2640	2,8486	0,4900	-1,9940
3	-0,4000	8,1550	2,0986	0,1600	-0,8395
4	-0,1000	3,8520	1,3486	0,0100	-0,1349
5	0,2000	1,8200	0,5988	0,0400	0,1198
6	0,5000	0,8600	-0,1508	0,2500	-0,0754
7	0,8000	0,4060	-0,9014	0,6400	-0,7211
8	1,0000	0,2460	-1,4024	1,0000	-1,4024
Somatório	**0,3000**	**69,1500**	**8,0386**	**3,5900**	**-8,6461**

$$\begin{cases} 8a + 0.3b = 8.0386 \\ 0.3a + 3.59b = -8.6461 \end{cases}$$

K	a	b	Erro a	Erro b	
0,0000	1,0048	-2,4084	-	-	
1,0000	1,0951	-2,4999	0,0903	0,0915	0,0915
2,0000	1,0986	-2,5002	0,0034	0,0003	0,0034
3,0000	1,0986	-2,5002	0,0000	0,0000	0,0000
4,0000	1,0986	-2,5002	0,0000	0,0000	0,0000

$$como \;\; \alpha = e^a$$

$$\alpha = e^{1.0986}$$

$$\alpha = 2.9999$$

$$y = 2.999e^{-2.5002x}$$