

PRO 3200 - Estatística

Lista 6: Correlação e Regressão

1) Uma regressão de y =volume de cálcio (g/L) em x =material dissolvido (mg/cm^2) foi descrita num artigo. A equação da reta de regressão estimada era $y=3,678 + 0,144x$, onde $r^2=0,860$, com base em $n=23$.

a. Interprete o coeficiente angular estimado e o coeficiente de determinação 0,860. R: 1 mg/cm^2 de aumento na quantidade dissolvida está associado a uma mudança de 0,144 no volume de cálcio. Aproximadamente 86% da variação observada pode ser atribuída a uma relação linear.

b. Calcule uma estimativa pontual do volume médio real de cálcio quando a quantidade de material dissolvido for igual a 50 mg/cm^2 . R: 10,88

c. O valor total da soma dos quadrados era $SQT=320,398$. Calcule uma estimativa do desvio padrão do erro no modelo de regressão linear simples. R: 1,46

2) Os autores de um artigo apresentaram uma análise de correlação para investigar a relação entre o nível máximo de lactato x e a resistência muscular y . Os dados a seguir foram obtidos

x	400	750	770	800	850	1025	1200	1250	1300	1400	1475	1480	1505	2200
y	3,80	4,00	4,90	5,20	4,00	3,50	6,30	6,88	7,55	4,95	7,80	4,45	6,60	8,90

$$S_{xx} = 36,9839$$

$$S_{yy}=2628930,375$$

$$S_{xy}= 7377,704$$

O gráfico de dispersão mostra um padrão linear:

a. Verifique se existe uma correlação positiva na população da qual esses dados foram obtidos. R: $r=0,748$, $t=3,9$, valor $P=0,001$. Usando tanto $\alpha=0,05$ ou $\alpha=0,01$, existe.

b. Se uma análise de regressão tivesse de ser realizada, que proporção da variação observada podia ser atribuída à relação linear aproximada? R: $r=0,56$, pouco deve ser atribuído

3) As propriedades físicas de seis amostras de tecidos antichamas foram investigadas em um artigo. Use os dados a seguir e teste em um nível de significância 0,05 se existe uma relação linear entre rigidez x (mg/cm) e espessura y (mm). O resultado do teste é surpreendente levando em consideração o valor de r ? R: $r=0,773$, sendo $t=2,44 < 2,776$; portanto $H_0: p=0$ não pode ser rejeitada.

X	7,98	24,52	12,47	6,92	24,11	35,71
Y	0,28	0,65	0,32	0,27	0,81	0,57

4) Sejam y os valores relativos à porosidade (%) e x o peso unitário de cimento queimado.

a) ajuste um modelo de regressão linear simples aos dados fornecidos. R. $y=118,91-0,905x$

b) calcule a estimativa da variância σ^2 . R. 1,93 c) encontre um intervalo de confiança de 95% para a inclinação da linha de regressão. R. $-1,181 \leq b \leq -0,629$ d) construa um intervalo de confiança de 95% para a . $109,648 \leq a \leq 128,72$

Obs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
X	99	101,1	103	103	105	107	108,7	110,8	112,1	112,4	113,6	113,8	115,1	115	120
y	28,8	27,9	27	22,8	21,5	20,9	19,6	17,1	18,9	16	16,7	13	13,6	10,8	25,2

5) Considere uma pesquisa para avaliar a influência de uma determinada substância na pureza da água. Os dados a seguir demonstram os resultados aferidos.

N. Observação	Nível substância x (%)	Pureza y 1,15(%)	N. Observação	Nível substância x (%)	Pureza y (%)
1	0,99	90,01	11	1,19	93,54
2	1,02	89,05	12	1,15	92,52
3	1,15	91,43	13	0,98	90,56
4	1,29	93,74	14	1,01	89,54
5	1,46	96,73	15	1,11	89,85
6	1,36	94,45	16	1,20	90,39
7	0,87	87,59	17	1,26	93,25
8	1,23	91,77	18	1,32	93,41
9	1,55	99,42	19	1,43	94,98
10	1,40	93,65	20	0,95	87,33

a) Ajuste um modelo de regressão linear simples aos dados de pureza da água ($\hat{Y}=74,20+14,97X$)

b) Calcule a variância σ^2 R. 1,17

c) Encontre um intervalo de confiança de 95% para a inclinação da linha de regressão R. $12,21 \leq b \leq 17,73$

d) Construa um intervalo de confiança de 95% para a R. $70,85 \leq a \leq 77,55$

e) Teste a significância da regressão, considerando nível de significância de 1%. R. $t_{cal}=11,41$ e $t_{crit}=2,88$

6) Considere que uma empresa esteja testando métodos motivacionais e deseja verificar a relação do gasto com métodos motivacionais aplicados aos funcionários e o resultado em vendas. Sejam y os valores relativos à quantidade vendida de um produto e x as despesas relativas aos valores investidos em métodos motivacionais.

a) ajustar um modelo de regressão linear simples R. $y=121,03+10,77x$

b) calcular a estimativa da variância σ^2 R. 2740,74

c) Encontrar um intervalo de confiança de 95% para a inclinação da linha de regressão. R. $0,932 \leq b \leq 20,61$

d) construir um intervalo de confiança de 95% para a. R. $44,25 \leq a \leq 197,81$

e) teste a significância da inclinação da linha de regressão, considerando o nível de significância de 2%. R. $t_{cal}=2,476$ e $t_{crit}=2,821$ f) teste a significância do aumento de vendas (y) quando acontece investimento em propaganda (x) ao nível de 2,5%. R. $t_{cal}=2,476$ e $t_{crit}=2,262$

x	1,5	5,5	10	3	7,5	5	13	4	9	12,5	6
y	120	190	240	140	180	150	280	110	210	220	310

7) Considere uma empresa que está avaliando a influência de um detergente na remoção de manchas. Sendo x o nível de detergente usado no teste e y um índice de qualidade da remoção das manchas: a) Calcule um intervalo de confiança de 95% para a qualidade média de todos os testes com nível de detergente de 45. R. 12,18;15,40 b) Calcule um intervalo de previsão de 95% para a qualidade que pode resultar da seleção de um único tipo com nível de detergente de 45 R. (7,51;20,07)

X	8	15	1,5	20	20	27,5	30	30	35
Y	22,8	27,2	23,7	17,1	21,5	18,6	16,1	23,4	13,4

X	38	40	45	50	50	55	55	59	60
Y	19,5	12,4	13,2	11,4	10,3	14,1	9,7	6,8	