

UNIVERSIDADE FEDERAL DE PELOTAS
CENTRO DE DESENVOLVIMENTO TECNOLÓGICO
MÉTODOS ESTATÍSTICOS APLICADOS A RECURSOS HÍDRICOS - 1ª PROVA

Nome: _____

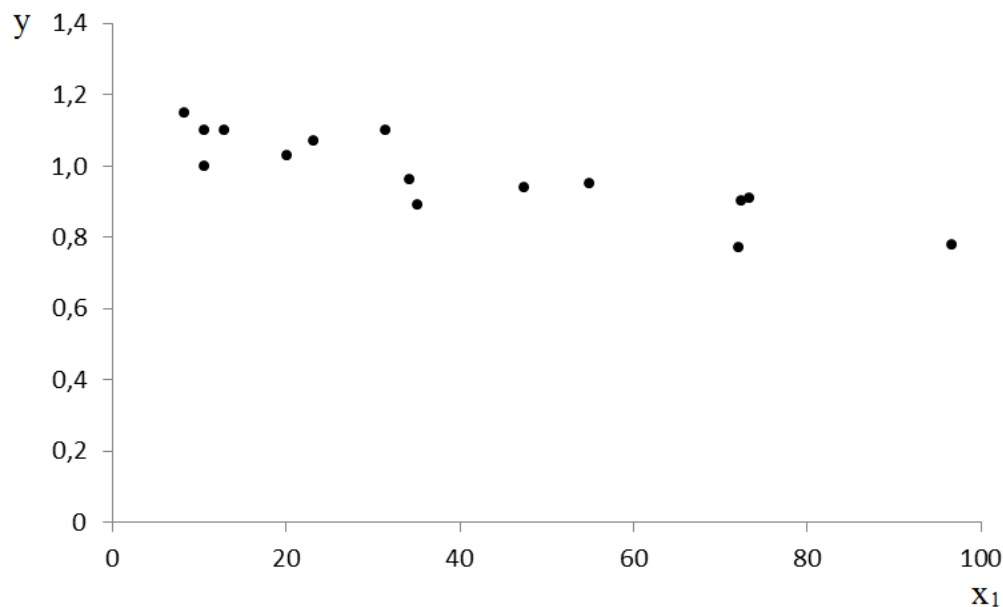
Data: 11/05/2015

Um estudo foi realizado em um caminhão de cargas leves movido a diesel para verificar se a umidade (em %) e a temperatura (em °C) do ar influenciam a emissão de óxido nitroso (em ppm). As medições das emissões foram tomadas em momentos diferentes, com condições experimentais variadas. Os dados observados estão na tabela abaixo.

j	Dados observados			Tabela auxiliar					
	Óxido nitroso (y)	Umidade (x ₁)	Temperatura (x ₂)	y ²	x ₁ ²	x ₂ ²	x ₁ y	x ₂ y	x ₁ x ₂
1	0,90	72,4	24,6	0,81	5241,76	605,71	65,16	22,15	1781,84
2	0,96	34,3	25,1	0,92	1176,49	627,78	32,93	24,05	859,41
3	0,89	35,1	20,0	0,79	1232,01	400,00	31,24	17,80	702,00
4	1,00	10,7	26,1	1,00	114,49	681,79	10,70	26,11	279,39
5	1,10	12,9	19,7	1,21	166,41	386,78	14,19	21,63	253,70
6	1,15	8,3	19,3	1,32	68,89	373,78	9,55	22,23	160,47
7	1,03	20,1	24,9	1,06	404,01	622,23	20,70	25,69	501,38
8	0,77	72,2	25,4	0,59	5212,84	644,60	55,59	19,55	1833,08
9	1,07	23,2	24,9	1,14	538,24	619,46	24,82	26,63	577,42
10	0,94	47,4	30,3	0,88	2246,76	920,11	44,56	28,51	1437,80
11	1,10	31,5	24,9	1,21	992,25	622,23	34,65	27,44	785,75
12	1,10	10,6	30,2	1,21	112,36	910,03	11,66	33,18	319,77
13	0,91	73,3	24,6	0,83	5372,89	605,71	66,70	22,40	1803,99
14	0,78	96,6	25,9	0,61	9331,56	673,11	75,35	20,24	2506,23
15	0,95	54,9	21,6	0,90	3014,01	467,04	52,16	20,53	1186,45
Soma	14,65	603,5	367,6	14,50	35224,97	9160,34	549,96	358,15	14988,68
Média	0,9767	40,23	24,51						

Questão 1 (5,0). Considere apenas as variáveis y e x₁.

(0,2) Interprete o diagrama de dispersão desses dados. [Ver slides](#)



a) (0,5) Calcule o coeficiente de correlação linear simples e interprete-o.

SQX ₁	10944,15
SQX ₂	151,06
SQY	0,1893
SPX ₁ X ₂	195,51
SPX ₁ Y	-39,46
SPX ₂ Y	-0,8760

$r = -0,8669$ Interpretação: ver slides

b) (0,6) Estabeleça a hipótese de interesse a respeito do coeficiente de correlação linear simples, teste-a e redija a conclusão. Use $\alpha = 0,05$.

$t = -6,2715$ Hipóteses e conclusão: ver slides

c) (0,2) Interprete o seguinte intervalo de confiança: IC (ρ ; 0,95): [-0,9551; -0,6380].

Interpretação: ver slides

d) (0,8) Assumindo que o modelo de regressão linear simples é adequado para descrever a relação entre x_1 e y , estabeleça o modelo, diga o significado de cada termo (contextualizando) e indique as pressuposições associadas ao modelo.

Ver slides

e) (0,5) Estime os parâmetros do modelo de regressão linear simples.

Parâmetro	Estimativa
β_0	1,121743
β_1	-0,003606

f) (0,6) Ajuste a equação da reta, calcule $\hat{\mu}_5$ e $\hat{\epsilon}_5$ e trace a reta ajustada no diagrama de dispersão.

$$\hat{\mu}_5 = 1,1652$$

$$\hat{\epsilon}_5 = -0,06523$$

g) (0,8) Estabeleça a hipótese de interesse a respeito do coeficiente de regressão e teste-a utilizando a análise da variância. Use $\alpha = 0,05$. (Construa a tabela da análise da variância e redija a conclusão.)

Fontes	GL	SQ	QM	F
Regressão	1	0,1423	0,1423	39,33
Resíduo	13	0,04703	0,003618	-
Total	14	0,18933	-	-

Hipóteses e conclusão: ver slides

i) (0,4) Faça a **predição** de y para $x_1=60$ (intervalo com 99% de confiança) e interprete o resultado.

$$0,9054 \pm 3,012 \times 0,0003704 \quad \text{IC}(\mu; 0,99): [0,9042; 0,9065]$$

Interpretação: ver slides

j) (0,4) Faça a **previsão** de y para $x_1=60$ (intervalo com 99% de confiança) e interprete o resultado.

$$0,9054 \pm 3,012 \times 0,003988 \quad \text{IC}(y; 0,99): [0,8934; 0,9174]$$

Interpretação: ver slides

Questão 2 (4,0). Considere as variáveis y , x_1 e x_2 .

a) (0,8) Assumindo que o modelo de regressão linear múltipla é adequado para descrever a relação entre y , x_1 e x_2 , estime os parâmetros do modelo e ajuste a equação do plano.

Parâmetro	Estimativa
β_0	1,1510
β_1	-0,003585
β_2	-0,001231

b) (0,2) Estime a média de óxido nitroso emitido para $x_1 = 50\%$ e $x_2 = 28^\circ\text{C}$.

$$E(Y/x_1=50;x_2=28) = 0,9372$$

c) (0,2) Explique o significado da estimativa do coeficiente de regressão parcial β_1 neste modelo.

Ver slides

d) (0,8) Efetue a análise da variância para testar a hipótese geral de linearidade da relação entre y , x_1 e x_2 . (Estabeleça a hipótese de interesse, construa a tabela da análise da variância e redija a conclusão. Use $\alpha = 0,05$).

Fontes	GL	SQ	QM	F
Regressão	2	0,14253	0,07127	18,27
Resíduo	12	0,04680	0,003900	-
Total	14	0,18933	-	-

Hipóteses e conclusão: ver slides

e) (0,4) Calcule o coeficiente de determinação corrigido e explique o seu significado.

$$r^2_c = 0,7116 \quad \text{Interpretação: ver slides}$$

- f) (0,8) Teste as hipóteses parciais, usando $\alpha = 0,05$. (Relacione todos os passos do teste de hipóteses e redija as conclusões.)

Hipótese sob verificação	t calculado
$\beta_1 = 0$	-5,943
$\beta_2 = 0$	-0,2428

Conclusões: ver slides

- g) (0,6) Considerando os resultados do item anterior, qual é o modelo mais adequado para explicar a variação do óxido nitroso emitido por um caminhão de cargas leves movido a diesel? Por quê? Escreva a equação ajustada deste modelo e o seu coeficiente de determinação.

Ver slides

- h) (0,2) Estime a média de óxido nitroso emitido para $x_1 = 50\%$.

$$E(Y/x_1=50) = 0,9372$$

Questão 3 (1,0). Complete as afirmações com V (verdadeiro) ou F (falso) e **corrija** as falsas.

- a) (F) O método científico compreende uma sequência de operações, em que a primeira etapa consiste na **formulação de hipóteses científicas**.
- b) (V) Os métodos de pesquisa mais utilizados na engenharia são: o estudo observacional e o experimento. No experimento não há intervenção do pesquisador na amostra, mas no estudo observacional há.
- c) (F) Processos hidrológicos **determinísticos** ocorrem com alta frequência em rios naturais.
- d) (F) O modelo de regressão linear é adequado quando a parte **aleatória** do modelo é preponderante sobre a parte **fixa**.
- e) (V) O coeficiente de correlação linear não permite estabelecer relação de causa e efeito entre as variáveis estudadas.
- f) (F) Se em um teste de hipóteses o valor p for igual a 0,28 devemos **rejeitar** H_0 .
- g) (F) O método dos mínimos quadrados tem como objetivo obter estimativas de β_1 e β_0 de tal forma que **a soma dos erros** seja o menor valor possível.
- h) (F) A amplitude do intervalo de confiança para μ **diminui** quando x_i se afasta da média \bar{X} .
- i) (V) A temperatura é um exemplo de variável discreta medida em escala intervalar.
- j) (V) Não é recomendada a extrapolação da equação de regressão para além dos limites dos dados amostrais utilizados na estimativa dos parâmetros do modelo de regressão. Uma das razões para isso é que a relação entre as variáveis X e Y pode não ser linear para valores que extrapolam os dados utilizados na regressão.