Polinômios ortogonais

1- Os dados seguintes referem-se às produções de grãos, em Kg/parcela, de um experimento de adubação de milho instalado segundo um DBC. Os tratamentos foram: as adubações com 0, 25, 50, 75 e 100 Kg/ha de adubo fosfatado.

BLOCOS	0	25	50	75	100	Totais de Blocos
I	8,38	7,15	10,07	9,55	9,14	44,29
II	5,77	9,78	9,73	8,98	10,17	44,43
III	4,90	9,99	7,92	10,24	9,75	42,80
IV	4,54	10,70	9,48	8,66	9,50	42,88
Totais de Tratamentos	23,59	37,62	37,20	37,43	38,56	174,40

a) Fazer a Análise de Variância considerando um delineamento casualizado em blocos

Considerando o modelo estatístico

$y_{ij} = \mu + A_i + B_j + \varepsilon_{ij}$, com *i* =1, ..., 5 e *j* = 1, ..., 4

em que: y_{ij} é a produção de grãos de milho da parcela do j-ésimo bloco que recebeu a i-ésima adubação; μ é uma constante inerente a todas as observações; A_i é o efeito da i-ésima adubação; B_j é o efeito do j-ésimo bloco e ε_{ij} é o erro experimental associado à produção de grãos de milho da parcela do j-ésimo bloco que recebeu a i-ésima adubação, independente e identicamente distribuído de uma Normal com média zero e variância constante σ^2.

Soma de Quadrado Total

$$SQTot = \sum_{i=1}^{5}\sum_{j=1}^{4} y_{ij}^2 - \frac{\left(\sum_{i=1}^{5}\sum_{j=1}^{4} y_{ij}\right)^2}{n} =$$

$$= (8{,}38^2 + \cdots + 9{,}50^2) - \frac{(174{,}4)^2}{20} = 1581{,}8516 - 1520{,}7680 =$$

$$SQTot = 61{,}0836$$

Soma de Quadrado de Blocos

$$SQBloco = \frac{1}{5}\sum_{j=1}^{4} y_{.j}^2 - \frac{\left(\sum_{i=1}^{5}\sum_{j=1}^{4} y_{ij}\right)^2}{n}$$

$$= \frac{1}{5}(44{,}29^2 + \cdots + 42{,}88^2) - \frac{(174{,}4)^2}{20} = 1521{,}2327 - 1520{,}7680 =$$

$$SQBloco = 0{,}4647$$

Soma de Quadrado de Tratamentos

$$SQTrat = \frac{1}{4}\sum_{i=1}^{5} y_{i.}^2 - \frac{\left(\sum_{i=1}^{5}\sum_{j=1}^{4} y_{ij}\right)^2}{n}$$

$$= \frac{1}{4}(23{,}59^2 + \cdots + 38{,}56^2) - \frac{(174{,}4)^2}{20} = 1560{,}8678 - 1520{,}7680 =$$

$$SQTrat = 40{,}0998$$

Soma de Quadrado de Resíduo

$$SQRes = SQTotal - SQBloco - SQTrat = 20{,}5191$$

Quadro da ANAVA

Fonte de Variação	Graus de Liberdade	SQ	QM	Fc
Bloco	*j-1=3*	0,4647	0,1549	0,091
Tratamento	*i-1=4*	40,0998	10,0249	5,863
Resíduo	*(i-1)(j-1)=12*	20,5191	1,7099	
Total	*ij-1=19*	61,0836	---	

```
##############################
#qf(alfa,glreg,glres,lower.tail=F)

# Tratamento
qf(0.05,4,12,lower.tail=F)
Ft= 3,259167
```

Como Fc > Ft rejeita-se a hipótese H0 e concluímos que existe efeito de tratamento sobre a produção de grãos de milho.

b) Desdobrar os graus de liberdade de tratamento (gl=4) em função dos polinômios ortogonais

Quadro da ANAVA

Fonte de Variação	Graus de Liberdade	SQ	QM	Fc
Bloco	*j-1=3*	0,4647	0,1549	0,091
Tratamento	*i-1=4*	40,0998	10,0249	5,863
R. Linear	*1*			
R. Quadrática	*1*			
R. Cúbica	*1*			
R. 4º grau	*1*			
Resíduo	*(i-1)(j-1)=12*	20,5191	1,7099	
Total	*ij-1=19*	61,0836	---	

Para encontrar as somas de quadrados de regressões utilizam-se os coeficientes da tabela do Pimentel

Totais de tratamentos		Coeficientes			
		C1 Linear	C2 Quadrática	C3 Cúbica	C4 4º grau
$y_{1.}$	23,59	-2	2	-1	1
$y_{2.}$	37,62	-1	-1	2	-4
$y_{3.}$	37,20	0	-2	0	6
$y_{4.}$	37,43	1	-1	-2	-4
$y_{5.}$	38,56	2	2	1	1
K		10	14	10	70
M		1	1	5/6	35/12

<u>Regressão Linear</u>

$$SQRLin = \frac{\left(\sum_{i=1}^{5} c_1 y_{i.}\right)^2}{rK_1}$$

$$SQRLin = \frac{[-2(23{,}59) - 1(37{,}62) + 0(37{,}20) + 1(37{,}43) + 2(38{,}56)]^2}{4(10)}$$

$$SQRLin = \frac{885{,}0625}{40} = 22{,}1266$$

<u>Regressão Quadrática</u>

$$SQRQuad = \frac{\left(\sum_{i=1}^{5} c_2 y_{i.}\right)^2}{rK_2}$$

$$SQRQuad = \frac{[2(23{,}59) - 1(37{,}62) - 2(37{,}20) - 1(37{,}43) + 2(38{,}56)]^2}{4(14)}$$

$$SQRQuad = \frac{632{,}5225}{56} = 11{,}2950$$

<u>Regressão Cúbica</u>

$$SQRCub = \frac{\left(\sum_{i=1}^{5} c_3 y_{i.}\right)^2}{rK_3}$$

$$SQRCub = \frac{[-1(23{,}59) + 2(37{,}62) + 0(37{,}20) - 2(37{,}43) + 1(38{,}56)]^2}{4(10)}$$

$$SQRCub = \frac{235{,}6225}{40} = 5{,}8906$$

Regressão 4º grau

$$SQR4grau = \frac{\left(\sum_{i=1}^{5} c_4 y_{i.}\right)^2}{rK_4}$$

$$SQR4grau = \frac{[1(23{,}59) - 4(37{,}62) + 6(37{,}20) - 4(37{,}43) + 1(38{,}56)]^2}{4(70)}$$

$$SQR4grau = \frac{220{,}5225}{280} = 0{,}7876$$

Quadro da ANAVA

Fonte de Variação	Graus de Liberdade	SQ	QM	Fc
Bloco	*3*	0,4647	0,1549	0,091
Tratamento	*4*	40,0998	10,0249	5,863
R. Linear	*1*	22,1266	22,1266	12,9403
R. Quadrática	*1*	11,2950	11,2950	6,6056
R. Cúbica	*1*	5,8906	5,8906	3,4449
R. 4º grau	*1*	0,7876	0,7876	0,4606
Resíduo	*12*	20,5191	1,7099	
Total	*19*	61,0836	---	

Atenção: $SQTrat = SQRLin + SQRQuad + SQRCub + SQR4grau$

```
###############################
#qf(alfa,glreg,glres,lower.tail=F)

# Tratamento
qf(0.05,1,12,lower.tail=F)
Ft= 4,747225
```

Conclusão: As regressões, linear e quadrática, foram significativas e se ajustam aos dados. Devemos considerar a de maior grau que foi significativa, ou seja, o modelo que melhor se ajusta aos dados é um polinômio de 2º grau ou uma equação quadrática.

c) Obtenção da equação de regressão

Para encontrar a equação, calculam-se os coeficientes dos componentes a partir do linear até o último grau significativo

Ex:

Linear significativa: $\hat{y} = \bar{y} + B_1M_1P_1$

Quadrática significativa: $\hat{y} = \bar{y} + B_1M_1P_1 + B_2M_2P_2$

Cúbica significativa: $\hat{y} = \bar{y} + B_1M_1P_1 + B_2M_2P_2 + B_3M_3P_3$

Grau K: $\hat{y} = \bar{y} + B_1M_1P_1 + B_2M_2P_2 + B_3M_3P_3 + \cdots + B_kM_kP_k$

Grau K: $\hat{y} = \bar{y} + B_1M_1P_1 + B_2M_2P_2 + B_3M_3P_3 + \cdots + B_kM_kP_k$

Em que:

i) $\bar{y}$ é a média geral dos dados;

ii) B_k são os coeficientes dos componentes linear, quadrático,...., dados por:

$$B_k = \frac{\left(\sum_{i=1}^{5} c_k y_{i.}\right)}{rK_k}$$

$$B_1 = \frac{\left(\sum_{i=1}^{5} c_1 y_{i.}\right)}{rK_1}$$
$$B_2 = \frac{\left(\sum_{i=1}^{5} c_2 y_{i.}\right)}{rK_2}$$

ATENÇÃO, é muito parecido com as fórmulas das somas de quadrados de regressão

$$SQRLin = \frac{\left(\sum_{i=1}^{5} c_1 y_{i.}\right)^2}{rK_1}$$

iii) M_k são os multiplicadores dados na tabelas de polinômios;

iv) P_k são os polinômios em função dos níveis do tratamento, dados por:

$$P_1 = x$$
$$P_2 = x^2 - \frac{n^2 - 1}{12}$$
$$P_3 = x^3 - \left(\frac{3n^2 - 7}{20}x\right)$$

Em que: $x = \frac{X_i - \bar{X}}{q}$

n é o número de níveis de tratamento= 5 doses de adubo

No exemplo temos que ajustar a equação quadrática, logo, a equação é dada por:

$$\hat{y} = \bar{y} + B_1M_1P_1 + B_2M_2P_2$$

i) $\bar{y} = \frac{174,4}{20} = 8,72$

ii) $B_1 = \frac{\left(\sum_{i=1}^{5} c_1 y_{i.}\right)}{rK_1} = \frac{29,75}{4(10)} = 0,74375$

$$B_2 = \frac{\left(\sum_{i=1}^{5} c_2 y_{i.}\right)}{rK_2} = \frac{-25,15}{4(14)} = -0,44911$$

iii) $P_1 = x = \frac{X_i - \bar{X}}{q} = \frac{X_i - 50}{25}$

$$P_2 = x^2 - \frac{n^2 - 1}{12} = \left(\frac{X_i - 50}{25}\right)^2 - \frac{5^2 - 1}{12} = \frac{X_i^2}{625} - \frac{100X_i}{625} + \frac{2500}{625} - 2$$

$$P_2 = \frac{X_i^2}{625} - \frac{100X_i}{625} + \frac{2500}{625} - \frac{1250}{625} = \qquad \frac{X_i^2}{625} - \frac{100X_i}{625} + \frac{1250}{625}$$

Montando a equação

$$\hat{y} = \bar{y} + B_1M_1P_1 + B_2M_2P_2$$

$$\hat{y} = 8{,}72 + 0{,}74375(1)\left(\frac{X_i - 50}{25}\right) - 0{,}44911(1)\left(\frac{X_i^2}{625} - \frac{100X_i}{625} + \frac{1250}{625}\right)$$

$$\hat{y} = 8{,}72 + 0{,}02975X_i - 1{,}4875 - 0{,}0007186X_i^2 + 0{,}07186X_i - 0{,}89822$$

$$\hat{y} = 6{,}33 + 0{,}10161X_i - 0{,}0007186X_i^2$$

Exercícios

1- Um experimento foi instalado segundo um delineamento em quadrado latino no qual foi avaliado a digestibilidade aparente de carboidratos totais (%) em função da porcentagem de proteína na ração (A=7% de proteína; B=9,5%; C=12%; D=14,5%; E=17%). Foi utilizado o controle do efeito de animais diferentes (linhas) e de períodos diferentes (coluna).

ANIMAIS	PERÍODOS				
	I	II	III	IV	V
B246	46(A)	60(B)	62(C)	69(D)	65(E)
AT14	65(E)	69(A)	72(D)	64(B)	60(C)
NN89	54(C)	80(D)	67(A)	71(E)	63(B)
AG90	63(B)	72(C)	74(E)	66(A)	64(D)
SS45	69(D)	85(E)	74(B)	72(C)	70(A)

1.1. Apresente o modelo estatístico do experimento. Faça a análise de variância de acordo com o modelo

1.2. Faça a análise de regressão utilizando os polinômios ortogonais

2- Um experimento com bananas foi instalado visando a comparação de 4 lâminas de água. O delineamento foi o casualizado em blocos com 8 repetições. Os dados a seguir referem-se aos diâmetros de pseudocaules das plantas filhas. Apresente a análise de variância e análise de regressão caso seja necessário.

Blocos	20%	40%	60%	80%	Blocos	20%	40%	60%	80%
I	5	4	3	8	V	7	6	10	4
II	7	6	4	7	VI	8	5	12	7
III	8	5	5	9	VII	10	8	8	8
IV	10	9	6	13	VIII	12	7	11	6