SUMÁRIO

1 Regressão na análise de variância **2**

1.1 Polinômio Ortogonal 2

1.1.1 Constraste Polinomial 5

1.1.2 Regressão por polinômio ortogonal 6

2 Tabelas **10**

1

REGRESSÃO NA ANÁLISE DE VARIÂNCIA

Os tratamentos são denominados quantitativos quando podem ser ordenados segundo um critério quantitativos, por exemplo:

- Determinar a melhor temperatura (0, 10, 20^oC) para quebra de dormência das sementes de uma espécie.
- Avaliar o efeito de diferentes doses de um remédio na concentração de histamina no sangue;
- Avaliar o efeito da adição de água no resistência de um bloco de concreto

Se os tratamentos em estudo no experimento forem quantitativos deve-se utilizar a análise de regressão quando a ANOVA rejeita a hipótese nula.

A regressão é uma técnica de análise que utiliza a relação entre duas ou mais variáveis quantitativas para determinar um modelo matemático de forma que o efeito de uma possa ser previsto por meio de outra variável. Na análise de experimentos, em geral utiliza-se modelos polinomais. Qualquer modelo de regressão linear e não linear pode ser utilizado.

1.1 POLINÔMIO ORTOGONAL

Os modelos de regressão polinomial de grau p, podem ser representado por:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + ... + \beta_p X^p$$

A variável X, ou variável independente, é uma variável não aleatória corresponde aos tratamentos e a variável Y, ou variável dependente, que é a variável resposta. Quando os níveis do tratamento são equidistantes e com igual numero de repetições pode-se utilizar polinômios ortogonais para ajustar o modelo de regressão.

Considerando que os níveis X do tratamento são equidistantes, temos

$$X_1 = X_1;\ X_2 = X_1 + q;\ X_3 = X_1 + 2q; ...; X_p = X_1 + (p-1)q;$$

Assim, o modelo de regressão polinomial pode ser escrito como:

$$Y = B_0 + B_1 P_1(X) + B_2 P_2(X) + ... + B_p P_p(X)$$

em que $P_k(X) = c_1 X_1 + c_2 X_2 + ... + c_n X_n$ é um polinômio ortogonal de ordem $k = 1, 2, ..., p$ que deve atender as seguintes restrições:

- $P_0(X) = 1$
- $\sum\limits_{i=1}^{I} P_k(X_i) = 0$
- $\sum\limits_{i=1}^{I} P_k(X_i) = 0 P'_k(X_i) = 0$ para $k \neq k'$

Os valores dos coeficientes de $P_k(X)$, podem ser obtidos da seguinte forma:

$$\begin{aligned}
C_{i1} &= \lambda_1 z_i \text{ em que } z_i = \frac{(X_i - \overline{X})}{q} \\
C_{i2} &= \lambda_2 \left(z_i^2 - \frac{I^2 - 1}{12} \right) \\
C_{i3} &= \lambda_3 \left(z_i^3 - \frac{3I^2 - 7}{20} z_i \right) \\
C_{i4} &= \lambda_4 \left(z_i^4 - \frac{3I^2 - 13}{14} z_i^2 + \frac{3(I^2 - 1)(I^2 - 9)}{560} \right) \\
C_{i5} &= \lambda_5 \left(z_i^5 - \frac{5(I^2 - 7)}{18} z_i^3 + \frac{15I^4 - 230I^2 + 407}{1008} z_i \right)
\end{aligned}$$

em que

- Os valores dos λ_i são escolhidos de forma que os C_{ii} tenham o menor valor inteiro possível
- $\overline{X} = \frac{\sum_{i=1}^{I} X_i}{I} = \frac{I X_1 + \sum_{i=1}^{I-1} iq}{I}$ é a média dos niveis do tratamentos
- q é a amplitude entre os níveis
- I é quantidade de níveis do tratamento

Exemplo 1.1: Se um experimento tiver um tratamento quantitativo com 4 níveis, podemos ajustar modelos de regressão até grau 3, assim:

$$
\begin{aligned}
\overline{X} &= X_1 + \frac{3}{2}q \\
z_1 &= \frac{X_1 - X_1 - \frac{3}{2}q}{q} = \frac{-3}{2} \\
z_2 &= \frac{X_2 - X_1 - \frac{3}{2}q}{q} = \frac{-1}{2} \\
z_3 &= \frac{X_3 - X_1 - \frac{3}{2}q}{q} = \frac{1}{2} \\
z_4 &= \frac{X_4 - X_1 - \frac{3}{2}q}{q} = \frac{3}{2}
\end{aligned}
$$

Assim $C_{11} = -3, C_{21} = -1, C_{31} = 1$ e $C_{14} = 3$

Para o polinômio do segundo grau temos:

$$
\begin{aligned}
C_{12} &= \left(\frac{-3}{2}\right)^2 + \frac{4^2-1}{12} = 1 \\
C_{22} &= \left(\frac{-1}{2}\right)^2 + \frac{4^2-1}{12} = -1 \\
C_{32} &= \left(\frac{1}{2}\right)^2 + \frac{4^2-1}{12} = -1 \\
C_{42} &= \left(\frac{3}{2}\right)^2 + \frac{4^2-1}{12} = 1
\end{aligned}
$$

Para o polinômio do terceiro grau temos

$$
\begin{aligned}
C_{12} &= \left(\frac{-3}{2} - \frac{3 \times 4^2 - 7}{20}\frac{-3}{2}\right) = -\frac{3}{10} \\
C_{22} &= \left(\frac{-1}{2} - \frac{3 \times 4^2 - 7}{20}\frac{-1}{2}\right) = \frac{9}{10} \\
C_{32} &= \left(\frac{1}{2} - \frac{3 \times 4^2 - 7}{20}\frac{1}{2}\right) = -\frac{9}{10} \\
C_{42} &= \left(\frac{3}{2} - \frac{3 \times 4^2 - 7}{20}\frac{3}{2}\right) = \frac{3}{10}
\end{aligned}
$$

Assim, para $\lambda_3 = \frac{10}{3}$, temos $C_{13} = -1, C_{23} = 3, C_{33} = -3$ e $C_{34} = -1$

1.1.1 Constraste Polinomial

A partir dos coeficientes dos polinômios ortogonais é possível obter a estimativa de um constrastes polinomial de ordem p, dado por:

$$\hat{Y}_p = \sum_{i=1}^{I} C_{ip} y_{i\cdot}$$

Os constrastes polinomiais, são contrastes mutuamente ortogonais, assim, para para testá-los utiliza-se o teste F da análise de variância. Assim, para I tratamentos e J repetições é possível ajustar $I-1$ modelos de regressã.

Cada modelo de regressão uma soma de quadrado associadada a 1 grau de liberdade

$$SQReg_p = \frac{\hat{Y}_p^2}{JK}$$

em que $K = \sum_{i=1}^{I} C_{ip}^2$

Exemplo 1.2: Um experimento foi realizado para estamos da temperatura sobre o tempo de vida de baterias. As taxas de carga utilizadas foram em quatro níveis: (0,5; 1,0, 1,5; 2,0) e foram utilizadas 5 repetições. O tempo de vida das baterias foi medida em termos do numero de ciclos de carga-descarga ate falhar.

Tratamento	Repetição				
	1	2	3	4	5
0,5	230,33	226,67	225,16	228,03	221,1
1,0	160,85	166,35	165,22	167,18	154,28
1,5	102,99	102,3	101,72	98,69	95,41
2,0	33,40	29,11	34,21	38,91	31,74

Aplicando a ANOVA, verifica-se que os tratamentos são significativos

FV	GL	SQ	QM	F
Tratamentos	3	102710,02	34236,67	2159,96
Resíduo	16	253,61	15,85	
Total	19	102963,63		

Para os tratamentos utilizados temos as médias

Tratamento	Média
0,5	226,26
1,0	162,78
1,5	100,22
2,0	33,47

Podemos obter os constrastes polinomiais

Tratamento	$y_{i.}$	C_{i1}	C_{i2}	C_{i3}	linear	quadrático	cúbico
0,5	1131,29	-3	1	-1	-3393,87	1131,29	-1131,29
1,0	813,88	-1	-1	3	-813,88	-813,88	2441,64
1,5	501,11	1	-1	-3	501,11	-501,11	-1503,33
2,0	167,37	3	1	1	502,11	167,37	167,37

- Contraste Linear

$$\hat{Y}_1 = -3204,53 \quad SQReg_1 = \frac{(-3204,53)^2}{5 \times 20} = 102690,13$$

- Contraste Quadrático

$$\hat{Y}_2 = -16,33 \quad SQReg_2 = \frac{(-16,33)^2}{5 \times 4} = 13,33$$

- Contraste Cúbico

$$\hat{Y}_2 = -25,61 \quad SQReg_3 = \frac{(-25,61)^2}{5 \times 20} = 6,56$$

Assim, temos

FV	GL	SQ	QM	F	F_t
Tratamentos	3	102.710,02	34.236,67	2.159,96	3,23
linear	1	102.690,13	102.690,13	6.478,63	4,49
quadrático	1	13,33	13,33	0,84	4,49
cubico	1	6,56	6,56	0,41	4,49
Resíduo	16	253,61	15,85		
Total	19	102.963,63			

Por meio da ANOVA verifica-se que apenas o efeito linear é significativo

1.1.2 Regressão por polinômio ortogonal

Para estimar a equação de regressão, de acordo com o modelo determinado na ANOVA, utiliza-se

$$\begin{aligned} B_0 &= \bar{y} \quad B_p = \frac{\hat{Y}_p}{JK} \\ Y &= B_0 + B_1 P_1(X) + B_2 P_2(X) + \ldots + B_p P_p(X) \end{aligned}$$

O coeficiente de determinação é dado por:

$$r^2 = \frac{SQReg_p}{SQTrat}$$

Exemplo 1.3: No exemplo da bateria temos:

$$
\begin{aligned}
B_0 &= 130,28 \quad B_1 = \frac{-3204,53}{5 \times 20} = -32,04 \\
P_1 &= \lambda_1 z_i = \lambda_1 \frac{X_i - \overline{X}}{q} \\
&= 2\frac{X-1,30}{0,5} = 4X - 5,2 \\
Y &= B_0 + B_1 P_1(X) = 130,28 - 32,04(4X - 5,2) \\
&= 130,28 - 128,16X + 166,60 = 296,60 - 128,16
\end{aligned}
$$

O coeficiente de determinação é dado por:

$$r^2 = \frac{102690,13}{102710,02} = 0,9998$$

Assim temos

$$Y = 296,60 - 128,16X \quad r^2 = 0,9998$$

Exemplo 1.4: Um experimento avaliou o efeito de diferentes doses de hCG no peso do útero ajustados para o peso vivo em camundongas. Foram utilizadas 30 camundongas, com 22 dias de idade, em que os animais foram divididos em 6 grupos Cada grupo recebeu respectivamente 0, 10, 20, 30, 40 e 50 UI de hCG, via intra-peritoneal, no volume de 0,2mL, duas vezes ao dia, com um intervalo de 12 horas entre as aplicações, durante um dia. Aproximadamente 24 horas após a última injeção as camundongas foram individualmente pesadas e eutanasiadas em seguida foram obtidos os pesos dos úteros.

Rep	Doses					
	0	10	20	30	40	50
1	0,52	0,99	1,28	1,37	1,27	1,02
2	0,52	0,99	1,29	1,40	1,31	0,98
3	0,52	0,97	1,30	1,41	1,26	1,06
4	0,50	0,99	1,29	1,40	1,33	0,97
5	0,48	1,02	1,28	1,43	1,33	0,98

FV	GL	SQ	QM	F_c	F_t
Tratamento	5	2,6817	0,53634	869,73	2,62
Residuo	24	0,0148	0,00062		
Total	29	2,6965			

Como $valor - p < 0,05$, rejeita-se H_0, logo podemos concluir ao nível de 5% que o peso é diferente em pelo menos uma das doses de HCg. Assim, podemos obter os constrastes polinomiais

Tratamento	$y_{i.}$	C_{i1}	C_{i2}	C_{i3}	linear	quadrático	cúbico
0	2,54	-5	5	-5	-12,70	12,70	-12,70
10	4,96	-3	-1	7	-14,88	-4,96	34,72
20	6,44	-1	-4	4	-6,44	-25,76	25,76
30	7,01	1	-4	-4	7,01	-28,04	-28,04
40	6,5	3	-1	-7	19,50	-6,50	-45,50
50	5,01	5	5	5	25,05	25,05	25,05

- Contraste Linear

$$\hat{Y}_1 = 17,54 \quad SQReg_1 = \frac{(17,54)^2}{5 \times 70} = 0,8790$$

- Contraste Quadrático

$$\hat{Y}_2 = -27,51 \quad SQReg_2 = \frac{(-27,51)^2}{5 \times 84} = 1,8019$$

- Contraste Cúbico

$$\hat{Y}_2 = -0,71 \quad SQReg_3 = \frac{(-0,71)^2}{5 \times 180} = 0,006$$

FV	GL	SQ	QM	F_c	F_t
Tratamento	5	2,6817	0,53634	869,73	2,62
Linear	1	0,8790	0,8790	1417,74	4,26
Quadrático	1	1,8019	1,8019	2906,29	4,26
Cúbico	1	0,0006	0,0006	0,97	4,26
Desvio	2	0,0002	0,0002	0,32	3,40
Residuo	24	0,0148	0,00062		
Total	29	2,6965			

Pela ANOVA verifica-se que apenas o modelo linear e quadrático foram significativo.Como o $r^2 = 0,67$ modelo quadrático é maior eu o $r^2 = 0,33$ do modelo linear, os dados podem ser representados pelo modelo quadrático.

Assim, temos:

$$
\begin{aligned}
B_0 &= 1,08 \\
B_1 &= \frac{17,54}{5 \times 70} = 0,05 \\
B_2 &= \frac{-27,51}{5 \times 84} = -0,065 \\
z_i &= \frac{X_i - \overline{X}}{q} = \frac{X - 25}{10} = 0,1X - 2,5 \\
P_1 &= \lambda_1 z_i = \lambda_1 \\
&= 2(0,1X - 2,5) = 0,2X - 5 \\
P_2 &= \lambda_2 \left(z_i^2 - \frac{I^2 - 1}{12} \right) = \frac{3}{2} \left((0,1X - 2,5)^2 - \frac{35}{12} \right) \\
&= 0,015X^2 - 0,375X + 5 \\
Y &= 1,08 + 0,05(0,2X - 5) - 0,065(0,015X^2 - 0,375X + 5) \\
&= 1,08 + 0,01X - 0,25 - 0,001X^2 + 0,024X - 0,325 \\
&= 0,50 + 0,06X - 0,001X^2
\end{aligned}
$$

Pelas estimativas do modelo verifica-se:

- Se a dose de HCg for zero o peso médio será de 0,5g
- O peso maximo está em torno da 30 de HCg.
- 67% da variabilidade do peso pode ser explicado pelas doses de HCg.

2

TABELAS

Tabela 2.1: Coeficientes para a interpolação dos polinômios ortogonais

i	$I=3$		$I=4$			$I=5$				$I=6$				
	C_{i1}	C_{i2}	C_{i1}	C_{i2}	C_{i3}	C_{i1}	C_{i2}	C_{i3}	C_{i4}	C_{i1}	C_{i2}	C_{i3}	C_{i4}	C_{i5}
1	-1	1	-3	1	-1	-2	2	-1	1	-5	5	-5	1	-1
2	0	-2	-1	-1	3	-1	-1	2	-4	-3	-1	7	-3	5
3	1	1	1	-1	-3	0	-2	0	6	-1	-4	4	2	-10
4			3	1	1	1	-1	-2	-4	1	-4	-4	2	10
5						2	2	1	1	3	-1	-7	-3	-5
6										5	5	5	1	1
K	2	6	20	4	20	10	14	10	70	70	84	180	28	252
λ	1	3	2	1	10/3	1	1	5/6	35/12	2	3/2	5/3	7/12	21/10