

Planejamento e Análise Estatística de Experimentos em blocos incompletos

Prof. Caio Azevedo

Contexto

- Em muitos experimentos em blocos, não é possível observar todos os tratamentos em todos os blocos.
- As razões podem ser de ordem: econômica, logística, da própria natureza do experimento.
- As vezes, apesar do experimento ter sido delineado como blocos completos, pode haver perda de observações.
- Muitas vezes, tem-se o interesse de preservar um certo balanceamento (número total de vezes em que cada tratamento aparece).

Estrutura geral: blocos incompletos balanceados, $k=4, b=4$

- Temos a seguinte representação:

	Bloco			
	1	2	3	4
1	y_{11}	y_{12}	-	y_{14}
2	-	y_{22}	y_{23}	y_{24}
3	y_{31}	-	y_{33}	y_{34}
4	y_{41}	y_{42}	y_{43}	-

Estrutura geral: blocos incompletos desbalanceados,

$$k=4, b=4$$

- Temos a seguinte representação:

	Bloco			
	1	2	3	4
1	y_{11}	y_{12}	-	-
2	-	y_{22}	y_{23}	y_{24}
3	y_{31}	-	-	y_{34}
4	y_{41}	y_{42}	y_{43}	-

Continuação

- Nos focaremos nos planejamento em blocos incompletos balanceados (BIB).
- Note que o bloco pode ser visto, em termos de modelagem, como um fator.
- Em geral, espera-se observar interação entre o fator principal e o bloco (assim como efeito de bloco).

Continuação

- Contudo, uma abordagem usual consiste em se considerar modelos sem interação. Justificativas: os fatos descritos acima + interesse principal reside no fator de interesse + em geral, temos apenas uma única observação para cada combinação tratamento \times bloco.
- No caso do BIB, o fato de não observarmos todas as combinações tratamento \times bloco, compromete (ainda mais) a possibilidade de se considerar interação.
- Vejamos um exemplo.

Exemplo 8: catalisadores

- Suponha que um engenheiro químico acredita que o tempo de reação relacionado à um processo químico seja influenciado pelo tipo de catalisador empregado.
- Catalisador: substância que aumenta a velocidade de uma reação sem ser consumida.
- O experimento consiste em selecionar porções de matéria-prima e aplicar, em turnos separados, cada catalisador, e observar o tempo de reação.

Exemplo 8: catalisadores (cont.)

- Devido à possíveis variações entre as estruturas das porções de matéria-prima (o que pode afetar o desempenho do catalisador) as porções serão consideradas como blocos.
- Cada porção de matéria -prima é grande o suficiente para poder-se utilizar apenas três catalisadores.
- A ordem em que cada catalisador é empregado, dentro de cada bloco, foi aleatorizada.

Dados amostrais do Exemplo 8

Tratamento (catalisador)	Bloco (porção de matéria prima)			
	1	2	3	4
1	73	74	-	71
2	-	75	67	72
3	73	75	68	-
4	75	-	72	75

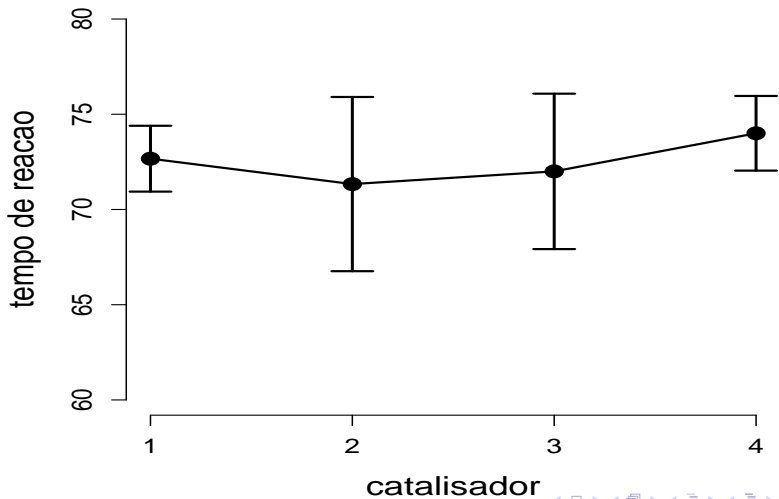
Análise descritiva (por tratamento)

Catalisador	Média	DP	Var	CV(%)	Mínimo	Máximo
1	72,67	1,53	2,33	2,10	71,00	74,00
2	71,33	4,04	16,33	5,67	67,00	75,00
3	72,00	3,61	13,00	5,01	68,00	75,00
4	74,00	1,73	3,00	2,34	72,00	75,00

Análise descritiva (por bloco)

Bloco	Média	DP	Var	CV(%)	Mínimo	Máximo
1	73,67	1,15	1,33	1,57	73,00	75,00
2	74,67	0,58	0,33	0,77	74,00	75,00
3	69,00	2,65	7,00	3,83	67,00	72,00
4	72,67	2,08	4,33	2,86	71,00	75,00

Gráfico de perfis (médios)



Análise dos dados

- Consideraremos um cenário com k tratamento e b blocos.
- Cada bloco contem $r < k$ tratamentos e cada tratamento aparece em $d < b$ blocos.
- Temos um total de $n = k \times b = b \times r$ observações.
- O número de vezes que cada par de tratamentos aparece no mesmo bloco é dado por:

$$\lambda = \frac{d(r-1)}{k-1}$$

Análise dos dados (cont.)

- O parâmetro λ tem de ser inteiro.
- Em nosso exemplo, temos que

$$k = 4, b = 4, r = 3, d = 3, n = 4 \times 3 = 3 \times 4 = 12, \lambda = 3 \times 2/3 = 2.$$

Modelo (casela de referência)

$$Y_{ij} = \mu + \alpha_i + \tau_j + \xi_{ij},$$

(Fator), $i = 1, 2, 3, \dots, k$; (Bloco), $j = 1, 2, 3, \dots, b$

- Erros $\xi_{ij} \stackrel{i.i.d}{\sim} N(0, \sigma^2)$, μ, α_i, τ_j , não aleatórios.
- Restrições : $\alpha_1 = \tau_1 = 0$.
- Contudo, note que para algumas conjuntos de índices (i,j) não heverá valores observados (y_{ij}) .

Decomposição da SQT

- Neste caso, como cada tratamento está em conjuntos diferentes de blocos, a decomposição tem ser feita de modo a separar o efeito de tratamento bloco, da seguinte forma:

$$SQT = SQF_{ajustada} + SQB + SQR$$

Decomposição da SQT

- Relembrando as hipóteses
 - Ausência de efeito de fator principal

$$H_0 : \mu_{1.} = \mu_{2.} = \dots \mu_{k.} \text{ vs } H_1 : \text{ pelo menos uma diferença}$$

- Ausência de efeito de bloco

$$H_0 : \mu_{.1} = \mu_{.2} = \dots \mu_{.b} \text{ vs } H_1 : \text{ pelo menos uma diferença}$$

$$\mu_{i.} = \frac{1}{b} \sum_{j=1}^b \mu_{ij}; \mu_{.j} = \frac{1}{k} \sum_{i=1}^k \mu_{ij}$$

Decomposição da SQT

- $SQT = \sum_{i=1}^k \sum_{j \in \mathbb{1}_i} Y_{ij}^2 - n\bar{Y}_{..}^2$
- $SQB = \frac{1}{k} \sum_{j=1}^b Y_{.j}^2 - n\bar{Y}_{..}^2$, $Y_{.j}$ é o total relacionado ao j ésimio bloco

■

$$SQF_{ajustado} = \frac{r}{k\lambda} \sum_{i=1}^k Q_i^2, Q_i^2 = Y_{i.} - \frac{1}{r} \sum_{j=1}^b n_{ij} \bar{Y}_{.j},$$

em que $n_{ij} = 1$, se o tratamento i aparece no bloco j e 0 caso contrário e $Y_{i.}$ é o total relacionado ao i ésimio tratamento.

- $SQR = SQT - SQF_{ajustada} - SQB$

Tabela de análise de variância

- Temos que:

FV	SQ	GL	QM	Estatística F	pvalor
Fator	$SQF_{ajustada}$	k-1	$QMF_{ajustada} = \frac{SQF_{ajustada}}{(k-1)}$	$F = \frac{QMF_{ajustada}}{QMR}$	$\min(F(f H_0), S(f H_0))$
Bloco	SQB	b-1			
Resíduo	SQR	n-k-b+1	$QMR = \frac{SQR}{[n-k-b+1]}$		
Total	SQT	n-1			

FV: fonte de variação, SQ: soma de quadrados, GL: graus de liberdade, QM: quadrado médio. $F(x|H_0)$, $S(x|H_0)$ fda e fds no ponto x sob H_0 , respectivamente.

Teste para efeito de bloco

- Caso se queira testar efeito de bloco, deve-se usar a $SQB_{ajustada}$, dada por:

$$SQB_{ajustada} = \frac{d}{b\lambda} \sum_{j=1}^b (Q'_j)^2$$

em que $Q'_j = Y_{.j} - \frac{1}{d} \sum_{i=1}^a n_{ij} y_{ij}$

- Assim, deve-se utilizar a estatística $F_B = \frac{SQB_{ajustada}}{SQR}$ que, sob H_0 é tal que $F_B \sim F_{(b-1, n-a-b+1)}$.
- O pacote do R **easynova**, calcula a $SQF_{ajustada}$ e $SQB_{ajustada}$.

Tabela de análise de variância (completa)

- Temos que:

FV	SQ	GL	QM	Estatística F	pvalor
Fator	$SQF_{ajustada}$	k-1	$QMF_{ajustada} = \frac{SQF_{ajustada}}{(k-1)}$	$F = \frac{QMF_{ajustada}}{QMR}$	$\min(F(f H_0), S(f H_0))$
Bloco	$SQB_{ajustada}$	b-1	$QMB_{ajustada} = \frac{SQB_{ajustada}}{(k-1)}$	$F_B = \frac{QMB_{ajustada}}{QMR}$	$\min(F(f H_0), S(f H_0))$
Resíduo	SQR	n-k-b+1	$QMR = \frac{SQR}{[n-k-b+1]}$		
Total	SQT	n-1			

FV: fonte de variação, SQ: soma de quadrados, GL: graus de liberdade, QM: quadrado médio. $F(x|H_0), S(x|H_0)$ fda e fds no ponto x sob H_0 , respectivamente.

Observações

- Note que:

$$SQT = SQF_{ajustada} + SQB + SQR$$

- Ou seja,

$$SQT \neq SQF_{ajustada} + SQB_{ajustada} + SQR$$

- Isso é devido à falta de ortogonalidade entre tratamento e bloco (devido à estrutura do planejamento BIB).

Comparações múltiplas

- Testes de comparações múltiplas podem ser conduzidos de modo semelhante ao apresentado anteriormente. Pequenas modificações são necessárias (veja páginas 157 e 158 do livro do Montgomery).
- Testes do tipo $C\beta = \mathbf{0}$ podem ser utilizados sem modificações.

Análise de resíduos

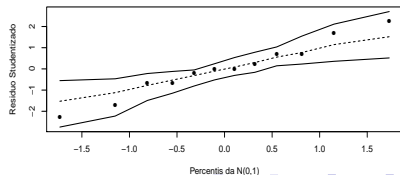
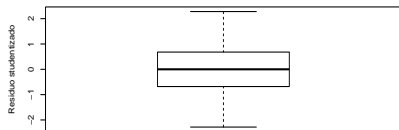
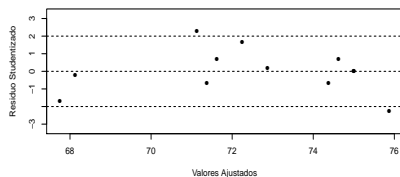
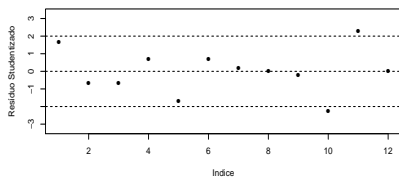


Tabela ANOVA

FV	SQ	GL	QM	Estatística F	pvalor
Catalisador (ajustada)	22,75	3	7,58	11,67	0,0107
Bloco (ajustada)	66,08	3	22,03	33,89	0,0010
Resíduo	3,25	5			
Total	81,00	11			

Efeito de tipo de catalisador.

Estimativas dos parâmetros do modelo

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	pvalor
μ	72,25	0,62	[71,04 ; 73,46]	117,33	<0,0001
α_2	0,25	0,70	[-1,12 ; 1,62]	0,36	0,7349
α_3	0,62	0,70	[-0,74 ; 1,99]	0,90	0,4117
α_4	3,62	0,70	[2,26 ; 4,99]	5,19	0,0035
τ_2	2,12	0,70	[0,76 ; 3,49]	3,04	0,0286
τ_3	-4,75	0,70	[-6,12 ; -3,38]	-6,80	0,0010
τ_4	-0,88	0,70	[-2,24 ; 0,49]	-1,25	0,2655

Possível igualdade entre os catalisadores 1, 2 e 3.

Resultado da aplicação do teste de Tukey

tratamentos	estimativa	erro-padrão	estatística	pvalor
2 - 1 == 0	0,25	0,70	0,36	0,9826
3 - 1 == 0	0,63	0,70	0,90	0,8085
4 - 1 == 0	3,63	0,70	5,19	0,0129
3 - 2 == 0	0,38	0,70	0,54	0,9462
4 - 2 == 0	3,38	0,70	4,83	0,0174
4 - 3 == 0	3,00	0,70	4,30	0,0282

Conclusão: catalisadores 1,2 e 3, equivalentes entre si, e diferentes do catalisador 4.

- Lembrando que $\mu_i = \mu + \alpha_i + \bar{\tau}$, $\bar{\tau} = \frac{1}{4} \sum_{j=1}^4 \tau_j$
- Hipótese 1:

$$H_0 : \begin{cases} \mu_1. = \mu_2. \\ \quad \quad e \\ \mu_1. = \mu_3. \end{cases} \leftrightarrow \begin{cases} \alpha_2 = 0 \\ \quad \quad e \\ \alpha_3 = 0 \end{cases}$$

- Se a hipótese 1 não for rejeitada, podemos testar Hipótese 2:

$$H_0 : \frac{\mu_1. + \mu_2. + \mu_3.}{3} = \mu_4. \leftrightarrow \alpha_2 + \alpha_3 - 3\alpha_4 = 0$$

- Os resultados foram : hipótese 1 0,41 (pvalor = 0,6865) e hipótese 2 34,19(pvalor = 0,0021).
- Conclusão: catalisadores 1,2 e 3, equivalentes entre si, e diferentes do catalisador 4.

Modelo reduzido (casela de referência)

$$Y_{ijk} = \mu + \alpha_i + \tau_j + \xi_{ijk},$$

(Fator), $i = 1, 2, 3, \dots, k$; (Bloco), $j = 1, \dots, b$ e $k = 1, 2, \dots, n_{ij}$ (repetição intra-bloco) .

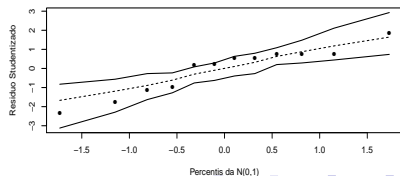
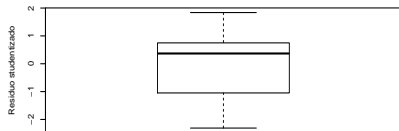
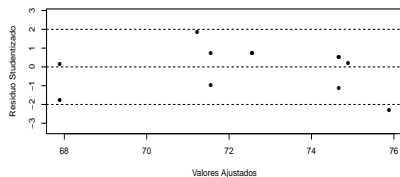
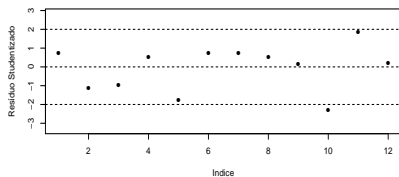
- Os catalisadores 1,2 e 3 são considerados como um único.
- Erros $\xi_{ij} \stackrel{i.i.d}{\sim} N(0, \sigma^2)$, μ, α_i, τ_j , não aleatórios.
- Restrições : $\alpha_1 = \tau_1 = 0$.
- Contudo, note que para algumas conjuntos de índices (i,j,k) não heverá valores observados (y_{ijk}).

Dados amostrais do Exemplo 8 (rearranjados)

- Neste caso, à rigor, em unido-se as observações daos catalisadores 1,2 e 3, teríamos um planejamento em blocos incompletos desbalanceados.
- Apesar de ajustarmos um modelo reduzido, **não consideraremos a Tabela ANOVA gerado por ele (da forma como o programa R nos fornece).**

Tratamento (catalisador)	Bloco (porção de matéria prima)			
	1	2	3	4
1,2,3	73,73	74,75,75	67,78	71,72
4	75	-	72	75

Análise de resíduos



Estimativas dos parâmetros do modelo

Parâmetro	Estimativa	EP	IC(95%)	Estat. t	pvalor
μ	72,56	0,46	[71,66 ; 73,45]	158,38	<0,0001
α_2	3,33	0,52	[2,32 ; 4,35]	6,42	0,0004
τ_2	2,11	0,62	[0,89 ; 3,33]	3,38	0,01117
τ_3	-4,67	0,60	[-5,84 ; -3,49]	-7,78	0,0001
τ_4	-1,00	0,60	[-2,18 ; 0,18]	-1,67	0,1394

As médias dos (novos) tratamentos são diferentes.

Estimativas finais das médias

Grupo	Estimativa	EP	IC(95%)
Grupo 1 (catalisadores 1,2,3)	71,37	0,27	[70,84 ; 71,90]
Grupo 2 (catalisador 4)	74,70	0,47	[73,78 ; 75,63]

Gráfico de perfis médios ajustados via modelo reduzido 2

