UNIVERSIDADE DE SÃO PAULO
Faculdade de Zootecnia e Engenharia de Alimentos

ZAB0229 – ESTATÍSTICA EXPERIMENTAL

AULA 9
DELINEAMENTO CASUALIZADO EM BLOCOS (DCB)

Material preparado pelo Prof. Dr. César Gonçalves de Lima (FZEA/USP) – ZAB0229

DELINEAMENTO CASUALIZADO EM BLOCOS (DCB)

1. INTRODUÇÃO

- No DCB estão envolvidos os três princípios básicos da experimentação: **repetição**, **casualização** e o **controle local**, com um único fator de controle.

- O DCB é usado quando o material (ou área) experimental é heterogêneo e se conhece a principal causa desta heterogeneidade.

  Neste caso, o material será reagrupado em novos grupos homogêneos (*blocos*) cada um dos quais constitui uma repetição.

- O erro experimental dentro de cada *bloco* deve ser o menor possível, mas entre os *blocos* pode variar à vontade.

- O número de unidades experimentais dentro de cada *bloco* deve ser igual ao número de tratamentos.

  Cada tratamento aparece somente uma vez em cada *bloco*.

- O número de repetições de cada tratamento é igual ao número de blocos.

- O sorteio dos tratamentos é realizado dentro de cada bloco.

Exemplo: O problema de fertilidade do solo nos experimentos agrícolas é o principal motivo para o uso do DCB. Dentro de cada bloco, as parcelas devem ter a mesma fertilidade, mas entre os blocos, a fertilidade pode variar bastante.

Exemplo: Nos experimentos zootécnicos cada bloco deve ser constituído de animais com características semelhantes.

- **Gado de leite**: mesma idade, peso, época de parição, ordem de parto, com produções de leite semelhantes *etc*.
- **Gado de corte**: pesos ou idades semelhantes, de mesmo sexo, de mesma raça *etc*.
- **Suínos**: mesma leitegada, mesmo sexo, pesos iniciais semelhantes *etc*.
- **Poedeiras**: mesma idade, mesma linhagem *etc*.
- **Frangos de corte**: mesma linhagem, mesmo sexo, mesma idade *etc*.

Importante:

- Lembrar que o manejo dos animais deve ser o mesmo em todas as parcelas.
- Os galpões experimentais devem ter boas condições de ambiência.

Caso geral:

- O fator usado para realizar o controle local (criação de blocos) não é de interesse principal no estudo.
- O fator é incluído no estudo, <u>exclusivamente</u>, para ajudar a controlar o erro experimental.

Vantagens do DCB:

- Com o agrupamento das parcelas similares, geralmente, se obtém resultados mais precisos do que os obtidos num DIC.
- Desde que exista material experimental suficiente, o delineamento casualizado em blocos (DCB) será sempre balanceado.
- A análise estatística ainda é bastante simples.

Desvantagens do DCB:

- A perda de uma ou mais unidades experimentais dificulta a análise, mas não impossibilita a sua realização.
- Apesar de existirem métodos para estimar o(s) valor(es) perdido(s), há uma perda de eficiência na comparação de médias envolvendo esses tratamentos.

Aleatorização

Exemplo: Num experimento com 5 tratamentos (T_1, T_2, T_3, T_4 e T_5) e 4 blocos (repetições) os tratamentos são sorteados às parcelas dentro de cada bloco, como no croqui:

Bloco 1	Bloco 2	Bloco 3	Bloco 4
T_2	T_1	T_3	T_4
T_1	T_2	T_1	T_3
T_4	T_3	T_5	T_5
T_3	T_5	T_4	T_2
T_5	T_4	T_2	T_1

Note que cada tratamento aparece uma única vez no mesmo bloco.

Dados de um experimento em DBC com a tratamentos e n blocos

	Blocos					
Tratamento	1	2	…	n	Total	Média
T_1	y_{11}	y_{12}	$\cdots$	y_{1n}	$y_{1\bullet}$	$\bar{y}_{1\bullet}$
T_2	y_{21}	y_{22}	$\cdots$	y_{2n}	$y_{2\bullet}$	$\bar{y}_{2\bullet}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
T_a	y_{a1}	y_{a2}	$\cdots$	y_{an}	$y_{a\bullet}$	$\bar{y}_{a\bullet}$
Total	$y_{\bullet 1}$	$y_{\bullet 2}$	$\cdots$	$y_{\bullet n}$	$y_{\bullet\bullet}$	$-$
Média	$\bar{y}_{\bullet 1}$	$\bar{y}_{\bullet 2}$	$\cdots$	$\bar{y}_{\bullet n}$	$-$	$\bar{y}_{\bullet\bullet}$

Em que: y_{ij} é a observação do i-ésimo tratamento feita no j-ésimo bloco; $y_{i\bullet}$ é o total do i-ésimo tratamento; $\bar{y}_{i\bullet}$ é a média do i-ésimo tratamento; $y_{\bullet j}$ é o total do j-ésimo bloco; $\bar{y}_{\bullet j}$ é a média do bloco j, $y_{\bullet\bullet}$ é o total geral e $\bar{y}_{\bullet\bullet}$ é a média geral.

2. MODELO ESTATÍSTICO

Admitimos o modelo:

$$y_{ij} = \mu + t_i + b_j + \varepsilon_{ij} \quad (1)$$

em que:

μ é uma constante comum a todas as observações

t_i é o efeito do i-ésimo tratamento, para $i = 1, 2, \ldots, a$

b_j é o efeito do j-ésimo bloco, para $j = 1, 2, \ldots, n$

ε_{ij} é o erro aleatório, não observável, atribuído a y_{ij}

Na estimação dos parâmetros usamos o Método dos Mínimos Quadrados, impondo as seguintes restrições: $\sum_{i=1}^{a} \hat{t}_i = 0$ e $\sum_{j=1}^{n} \hat{b}_j = 0$. Os estimadores de mínimos quadrados dos parâmetros do modelo restrito (1) são:

$$\hat{\mu} = \bar{y}_{\bullet\bullet} \qquad \hat{t}_i = \bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet} \qquad \hat{b}_j = \bar{y}_{\bullet j} - \bar{y}_{\bullet\bullet}$$

para $i = 1, \cdots, a$ e $j = 1, \cdots, n$

Suposições do modelo (que deverão ser confirmadas!)

- Os efeitos do modelo são aditivos.
- Os erros ε_{ij} são independentes, homocedásticos e têm distribuição normal, isto é, $\varepsilon_{ij} \sim N(0; \sigma^2)$.

3. HIPÓTESES A SEREM TESTADAS NA ANOVA

H_0: $\mu_1 = \mu_2 = \ldots = \mu_a$
H_1: pelo menos duas médias diferem entre si.

Que também pode ser escrita como:

H_0: $t_1 = t_2 = \cdots = t_a = 0$
H_1: $t_i \neq 0$, para algum i = 1, …, a.

Nota: Não é comum realizar um teste comparando as médias dos blocos. O teste pode ser realizado para indicar a necessidade do uso do DCB, quando se pretende repetir o experimento em condições similares ao que está sendo analisado

4. Decomposição da soma de quadrados total

$$SQTotal = \sum_{j=1}^{n}\sum_{i=1}^{a}\left(y_{ij} - \bar{y}_{\bullet\bullet}\right)^2$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{a}\left[\left(y_{ij} - \bar{y}_{i\bullet} - \bar{y}_{\bullet j} + \bar{y}_{\bullet\bullet}\right) + \left(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet}\right) + \left(\bar{y}_{\bullet j} - \bar{y}_{\bullet\bullet}\right)\right]^2$$

$$= \underbrace{n\sum_{i=1}^{a}\left(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet}\right)^2}_{SQTrat} + \underbrace{a\sum_{j=1}^{n}\left(\bar{y}_{\bullet j} - \bar{y}_{\bullet\bullet}\right)^2}_{SQBloco} +$$

$$+ \underbrace{\sum_{j=1}^{n}\sum_{i=1}^{a}\left(y_{ij} - \bar{y}_{i\bullet} - \bar{y}_{\bullet j} + \bar{y}_{\bullet\bullet}\right)^2}_{SQResiduo}$$

Iremos decompor a $SQTotal$ e os seus $g.l.$'s como:

$$SQTotal = SQTrat + SQBloco + SQResiduo$$

$(an - 1) = (a-1) + (n - 1) + (a-1)(n-1)$ graus de liberdade.

5. FÓRMULAS DE CÁLCULO:

$$SQTotal = \sum_{j=1}^{n}\sum_{i=1}^{a}\left(y_{ij} - \bar{y}_{\bullet\bullet}\right)^2 = \sum_{j=1}^{n}\sum_{i=1}^{a} y_{ij}^2 - C,$$

onde $C = \frac{1}{an}\, y_{\bullet\bullet}^2$

$$SQTrat = n\sum_{i=1}^{a}(\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet})^2 = \frac{1}{n}\sum_{i=1}^{a} y_{i\bullet}^2 - C$$

$$SQBloco = a\sum_{j=1}^{n}\left(\bar{y}_{\bullet j} - \bar{y}_{\bullet\bullet}\right)^2 = \frac{1}{a}\sum_{j=1}^{n} y_{\bullet j}^2 - C$$

$$SQResiduo = SQTotal - SQTrat - SQBloco$$

Em que:

n = número de blocos e a = número de tratamentos

Forma alternativa de cálculo das $SQ's$ numa calculadora científica, utilizando as médias:

A $SQTotal$ e a $SQTrat$ são obtidas de forma idêntica àquela utilizada no Delineamento Inteiramente Casualizado

$SQBloco$: entrar com as médias dos blocos e os respectivos números de valores usados para calcular a média de cada bloco $\Rightarrow$ calcular o desvio padrão amostral $\Rightarrow$ calcular a variância amostral $\Rightarrow$ multiplicar o resultado por $N-1$, em que $N = an$.

$$SQResiduo = SQTotal - SQTrat - SQBloco$$

Com as $SQ's$ e os respectivos $g.l.$´s podemos montar o seguinte quadro de análise de variância:

6. QUADRO DE ANÁLISE DE VARIÂNCIA

Fonte de variação	g.l.	$\boldsymbol{SQ}$	$\boldsymbol{QM}$	F
Bloco	$(n-1)$	$SQBloco$	$QMBloco$	–
Tratamento	$(a-1)$	$SQTrat$	$QMTrat$	$\dfrac{QMTrat}{QMResiduo}$
Resíduo	$(a-1)(n-1)$	$SQResiduo$	$QMResiduo$	
Total	$an-1$	$SQTotal$	–	–

$\bar{y}_{\bullet\bullet}$ $\qquad s^2 = QMResiduo$ $\qquad CV$

Pode-se provar que:

- $E(QMResiduo) = \sigma^2$, ou seja, o $QMResiduo$ é o melhor estimador da variância σ^2.
- $E(QMTrat) = \sigma^2 + \frac{n}{(a-1)} \sum_{i=1}^{a} t_i^2$, ou seja, $QMTrat$ é um estimador não viesado da variância (σ^2) somente quando a hipótese $H_0: t_1 = \cdots = t_a = 0$ (ou $H_0: \mu_1 = \cdots = \mu_a$) for verdadeira.

Importante:

- Geralmente não testamos se as médias dos *blocos* são iguais (ou os efeitos dos *blocos* são nulos), já que este fator é utilizado exclusivamente para separar em grupos homogêneos as unidades experimentais que diferem entre si por causa de uma fonte de heterogeneidade conhecida.

Para testar

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_a$$
$$H_1: \mu_i \neq \mu_{i^*}, \text{ para algum } i \neq i^*$$

Usamos a estatística $F = QMTrat/QMResiduo$, que sob H_0 tem distribuição F com $(a–1)$ e $(a–1)(n–1)$ graus de liberdade.

Comparamos F_{calc} com o valor de tabela (F_{tab}), ao nível α = 5%:

- Se $F_{calc} > F_{tab}$: rejeitamos H_0 e concluímos que existem pelo menos duas médias de tratamentos diferem entre si.
- Se $F_{calc} \leq F_{tab}$: aceitamos H_0 como verdadeira e concluímos que todas as médias de tratamentos são iguais entre si.

Exemplo 1: Com o objetivo de estudar o efeito da idade da castração no desenvolvimento e produção de suínos, utilizou-se um delineamento em blocos casualizados com quatro tratamentos e quatro repetições. Os blocos foram utilizados para controlar a variabilidade natural existente entre as leitegadas. Os tratamentos consistiram de:

A: suínos castrados aos 56 dias

B: suínos inteiros

C: suínos castrados aos 7 dias

D: suínos castrados aos 21 dias

Com base nos dados de ganhos de peso, em kg, ao final do experimento (252 dias), construa o quadro de ANOVA e teste a hipótese de interesse.

Tratamento	Bloco (leitegada)				Total
	1	2	3	4	
A	93.0	77.9	94.9	97.6	363,4
B	108.6	115.4	96.0	118.7	438,7
C	108.9	100.2	102.1	114.1	425,3
D	102.0	96.5	116.9	117.6	433,0
Total	412,5	390,0	409,9	448,0	1660,4

Cálculo das somas de quadrados:

$SQTotal$ = 1912,07

Para calcular $SQTrat$ vamos usar as médias dos tratamentos:

Trat	A	B	C	D
Média	$90{,}85^{(4)}$	$109{,}675^{(4)}$	$106{,}325^{(4)}$	$108{,}25^{(4)}$

Em que o expoente (4) indica o número de parcelas usadas no cálculo da média correspondente.

$$SQTrat = 913{,}58$$

Para calcular $SQBloco$ usamos as médias dos blocos:

Bloco	1	2	3	4
Média	$103{,}125^{(4)}$	$97{,}5^{(4)}$	$102{,}475^{(4)}$	$112{,}00^{(4)}$

$$SQBloco = 436{,}56$$

Finalmente:

$$SQResiduo = SQTotal - SQTrat - SQBloco = 561{,}93$$

Quadro de Análise de Variância (ANOVA)

Fonte de Variação	g.l.	SQ	QM	F
Bloco	3	436,56	145,52	
Tratamento	3	913,58	304,53	4,88 *
Resíduo	9	561,93	62,44	
Total	15	1912,07		

$\bar{y}_{\bullet\bullet} = 103{,}78\text{kg}$ $s^2 = 62{,}44$ $\text{CV} = 7{,}61\%$

Como $F_{calc} = 4{,}88 > F_{tab}(5\%; 3; 9)$ =3,86, rejeitamos $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$ ao nível $\alpha = 5\%$ e concluímos que pelo menos duas médias de tratamentos apresentam ganhos médios de peso diferentes.

O $\text{CV} < 10\%$ indica que o experimento foi bem conduzido, ou seja, os fatores estranhos ao experimento foram bem controlados.

Teste de Tukey para comparar as médias dos tratamentos:

$$dms = 4{,}42\sqrt{\frac{62{,}44}{4}} = 17{,}46\text{kg}$$

Tratamento	Média	
B	109,68	a
D	108,25	ab
C	106,33	ab
A	90,85	b

Médias seguidas por letras diferentes diferem entre si pelo Teste de Tukey (α = 5%)

Intervalo
[92,22; 109,68]
[90,79; 108,25]

Conclusão: Os suínos inteiros (tratamento B) tiveram um desempenho superior ao dos suínos castrados (tratamento A) aos 56 dias de idade.

EXEMPLO 2. (Apostila do Prof. Joanir, pág. 106) Comparação de quatro rações em 24 bovinos de idade, de pesos iniciais semelhantes, mas de raças diferentes. Utilizou-se o DCB com 6 repetições, usando o fator Raça para definir os blocos. Baseado nos ganhos de peso, em kg, dos animais compare o desempenho das rações em uma ANOVA.

	Tratamento (Ração)				
Bloco (Raça)	1	2	3	4	Total ($y_{\bullet j}$)
1	99	98	107	104	408
2	94	87	102	89	372
3	91	97	95	95	378
4	100	80	92	78	350
5	97	102	94	94	387
6	89	91	98	98	376
Total ($y_{i\bullet}$)	570	555	588	558	2271

Também podemos calcular as SQ's utilizando as médias:

Tratamento	1	2	3	4
Média	$95{,}0^{(6)}$	$92{,}5^{(6)}$	$98{,}0^{(6)}$	$93{,}0^{(6)}$

Bloco	1	2	3	4	5	6
Média	$102{,}0^{(4)}$	$93{,}0^{(4)}$	$94{,}5^{(4)}$	$87{,}5^{(4)}$	$96{,}75^{(4)}$	$94{,}0^{(4)}$

Utilizando as fórmulas adequadas, as somas de quadrados resultaram em (confirme!):

$$SQTotal = 1089{,}63 \qquad SQBloco = 450{,}88 \qquad SQTrat = 112{,}13$$

$$SQResiduo = 526{,}63$$

Quadro de Análise de Variância

Fonte de Variação	g.l.	SQ	QM	F
Bloco	5	450,88	90,18	
Ração	3	112,13	37,38	1,06 n.s.
Resíduo	15	526,63	35,11	
Total	23	1089,63		

$\bar{y}_{\bullet\bullet} = 94{,}63$ $\quad s^2 = 35{,}11$ $\quad$ C.V. = 6,26%

Hipóteses: $H_0\colon \mu_1 = \mu_2 = \mu_3 = \mu_4$

$H_1\colon \mu_i \neq \mu_{i^*}$, para algum $i \neq i^*$

Como $F_{calc} = 1{,}06$ é menor que o valor tabelado, $F(3;15) = 3{,}29$, a hipótese H_0 <u>não deve ser rejeitada</u> ($\alpha = 5\%$).

Concluímos que as médias dos ganhos de peso dos animais que receberam as rações 1, 2, 3 ou 4 são iguais.

- Como não houve diferença entre as médias dos tratamentos, uma boa estimativa do ganho de peso de todos os bovinos é a média geral $\bar{y}_{\bullet\bullet}$ = 94,6 kg.

A seguir vamos apresentar uma solução para o problema de perda de parcela no delineamento casualizado em blocos.

7. ESTIMAÇÃO DE PARCELA PERDIDA

Exemplo 1. Suponhamos que o leitão que recebeu o tratamento C no bloco 2 tenha morrido durante o experimento.

	Bloco (leitegada)				
Trat	1	2	3	4	Total
A	93,0	77,9	94,9	97,6	363,4
B	108,6	115,4	96,0	118,7	438,7
C	108,9	**X**	102,1	114,1	325,1+**X**
D	102,0	96,5	116,9	117,6	433,0
Total	412,5	289,8+**X**	409,9	448,0	1560,2+**X**

Problema: O que fazer quando ocorre a perda de uma parcela em um experimento planejado no delineamento casualizado em blocos?

Uma solução: estimar o valor da parcela perdida usando a fórmula:

$$X = \frac{nB + aT - G}{(n-1)(a-1)}$$

em que n é o número de blocos, a é o número de tratamentos, B é o total das parcelas restantes no bloco em que ocorreu a perda, T é o total das parcelas restantes no tratamento onde ocorreu a perda e G é o total das parcelas disponíveis.

Após a estimação do valor perdido devemos:

1. Colocar o valor calculado – X – no lugar da parcela perdida e realizar a análise, como apresentado anteriormente e calcular as SQ´s da forma usual, como se o experimento fosse balanceado.
2. A estimação de uma parcela perdida provoca a diminuição de um grau de liberdade no resíduo.

3. O $QMResiduo$ fica corretamente estimado, mas o valor da $SQTrat$ fica $ligeiramente$ exagerado e seu valor precisa ser corrigido.
4. Para corrigir a $SQTrat$, basta subtrair dela a quantidade:

$$U = \frac{a-1}{a}\left(X - \frac{B}{a-1}\right)^2$$

No Exemplo 1 tem-se:

$$n = 4;\ a = 4;\ B = 289{,}8;\ T = 325{,}1 \text{ e } G = 1560{,}2$$

A $melhor\ estimativa$ da parcela perdida é igual a:

$$X = \frac{4(289{,}8)+4(325{,}1)-1560{,}2}{(4-1)(4-1)} = 99{,}9$$

	Bloco (leitegada)				
Trat	1	2	3	4	Total
A	93,0	77,9	94,9	97,6	363,4
B	108,6	115,4	96,0	118,7	438,7
C	108,9	99,9	102,1	114,1	425,0
D	102,0	96,5	116,9	117,6	433,0
Total	412,5	389,7	409,9	448,0	1660,1

Com a tabela completa e usando as fórmulas apropriadas obtemos:

$SQTotal = 1914{,}30$ $\quad SQTrat = 912{,}06$ $\quad SQBloco = 440{,}34$

$$\Rightarrow SQResiduo = 561{,}66$$

O fator de correção da $SQTrat$ é: $U = \frac{(4-1)}{4}\left(99{,}9 - \frac{289{,}8}{(4-1)}\right)^2 = 8{,}32$

Logo:

$$SQTrat = 912{,}06 - 8{,}32 = 903{,}74$$

O novo quadro de ANOVA fica:

Fonte de Variação	g.l.	SQ	QM	F
Bloco	3	440,34	146,78	
Tratamento	3	903,74	301,25	4,29*
Resíduo	8	561,66	70,21	
Total	14	1914,30		

$$\bar{y}_{\bullet\bullet} = 104{,}01 \text{ kg} \quad s^2 = 70{,}21 \quad \text{CV} = 8{,}06\%$$

Como $F_{calc} = 4{,}29$ é superior ao $F_{tab}(3; 8; 5\%) = 4{,}07$, rejeitamos H_0 e concluímos que existem pelo menos duas idades de castração que proporcionaram ganhos médios de peso diferentes nos suínos.

O CV = 8,06% ainda é considerado baixo e indica que o experimento foi bem conduzido e que os fatores estranhos a ele foram bem controlados.

Observe que a média geral $\bar{y}_{\bullet\bullet} = 104{,}01$ kg foi calculada somente com as 15 respostas realmente observadas.

A análise pode continuar comparando-se as médias dos tratamentos utilizando o Teste de Tukey (5%), lembrando que o experimento é <u>desbalanceado</u> (o tratamento C tem somente 3 repetições).

- Na aplicação do teste de Tukey a média do tratamento C deve ser calculada com base nas três respostas realmente observadas.