

Análise de variância

Princípios básicos de delineamento experimental

Antes de escolher o delineamento experimental de uma experiência é importante saber claramente qual o objectivo do estudo (quais a(s) hipóteses(s) a testar)

Alguns termos importantes:

Factor: uma variável categórica (qualitativa);

Níveis do factor: as diferentes categorias do factor, ou seja, diferentes situações experimentais a serem testadas e onde se efectuam observações da variável resposta. **Conhecidos como Tratamentos** (diferentes variedades, fertilizantes, regimes alimentares, etc.).

Unidade experimental (indivíduo, parcela de terreno, vaso, placa de *petri*, etc.): as n observações da variável resposta correspondem a n diferentes unidades experimentais.

Princípios gerais a seguir:

Repetição

A repetição de observações independentes é essencial para estimar a variância dos erros aleatórios; permite a diminuição dos erros de amostragem e outros; a precisão de um ensaio aumenta com o número de repetições.

Casualização

ou seja, aleatoriedade na escolha das unidades experimentais e na associação que lhes é feita de um dado nível do factor.

- Condição fundamental para a validade de todo o processo indutivo (para se poder trabalhar com a teoria de probabilidades) e para evitar enviesamentos (mesmo inconscientes).
- É essencial para validar a estimativa a variância dos erros aleatórios.
- Há um paralelismo rigoroso entre o tipo de casualização adoptado e o modelo de análise dos dados.

Controlo da heterogeneidade entre unidades experimentais  Diferentes tipos de casualização/diferentes tipos de delineamento experimental

Delineamento experimentais

Alguns exemplos

Delineamento experimental totalmente casualizado (CRD)

Exemplo:

Exemplo de uma disposição de árvores de 5 proveniências de pinheiro manso (A,B,C,D,E) em 4 repetições, em condições homogêneas.

1 factor em estudo (proveniência), com 5 níveis.

D	B	A	E
C	E	D	B
A	C	E	A
D	B	A	C
B	C	E	D



**Uma casualização possível
(exemplo de um
delineamento totalmente
casualizado (CRD))**

Nota: Repetições e pseudo-repetições

Convém distinguir entre repetições e pseudo-repetições. Por pseudo-repetições entende-se medições que são feitas na mesma unidade experimental. Por exemplo: pesos de vários frutos de uma planta; medições do comprimento das folhas de uma planta; observações em plantas diferentes no mesmo vaso, na mesma parcela de terreno; observações na mesma amostra no laboratório, etc.. Estes exemplos são repetições que não são independentes.

Mas a existência de pseudo-repetições é essencial para a redução da variabilidade entre observações independentes. De facto, substituindo cada grupo de pseudo-repetições por uma única observação média pode-se diminuir a variabilidade entre diferentes observações médias (que são independentes).

A	A	A	A
B	B	B	B
C	C	C	C
D	D	D	D
E	E	E	E



Uma disposição sistemática

A mesma letra repetida são pseudo-repetições (por exemplo, 4 plantas seguidas no campo, 4 plantas seguidas num mesmo vaso, 4 vasos seguidos numa estufa, etc., portanto, existe apenas uma repetição independente por tratamento).

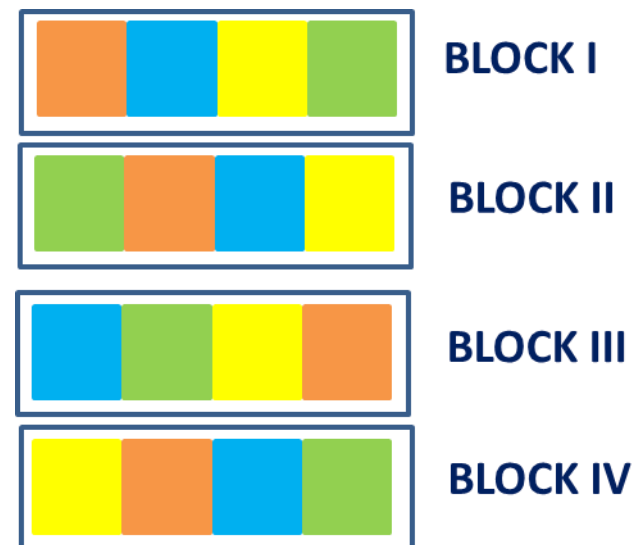
Delineamento em blocos completos casualizados (RCBD) desenvolvido por *Fisher* (1935) é dos delineamentos mais simples para o controlo da heterogeneidade (nomeadamente, da variação espacial em ensaios agrónómicos).



Exemplo: cada linha representa 1 bloco. Existem 4 blocos (I-IV) and 4 tratamentos (diferentes cores)

2 factores em estudo. Por exemplo, factor variedade (ou fertilizante, etc.) (4 níveis) e factor bloco (4 níveis).

- O campo é dividido em unidades homogéneas, de forma a controlar a variação no campo (ou na estufa, laboratório).
- Os níveis do factor a ser estudado (designados por tratamentos), são aleatoriamente atribuídos às unidades experimentais dentro do bloco (cada tratamento aparece uma única vez em cada bloco).
- O número de blocos é o número de repetições.

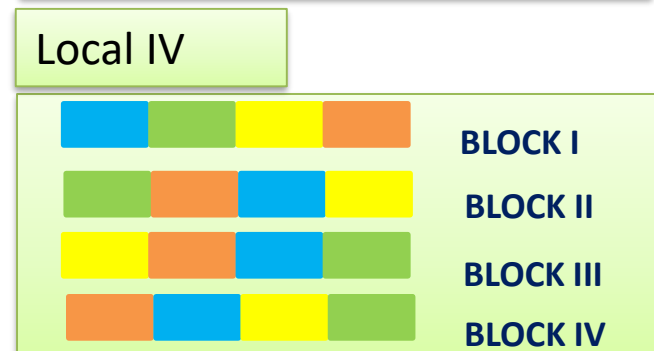
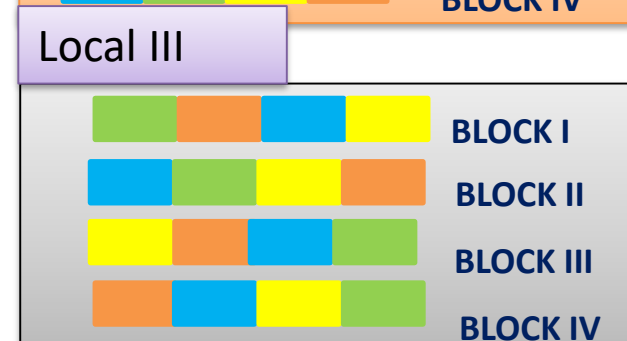
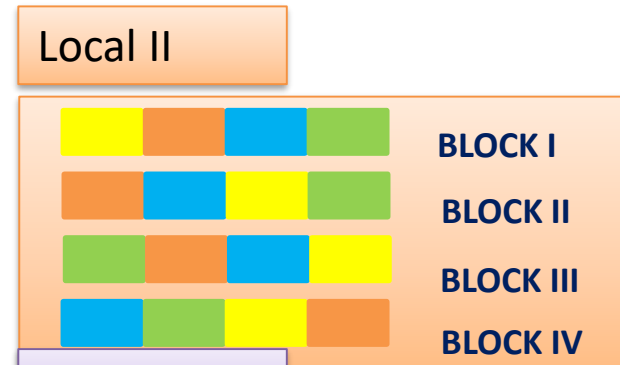
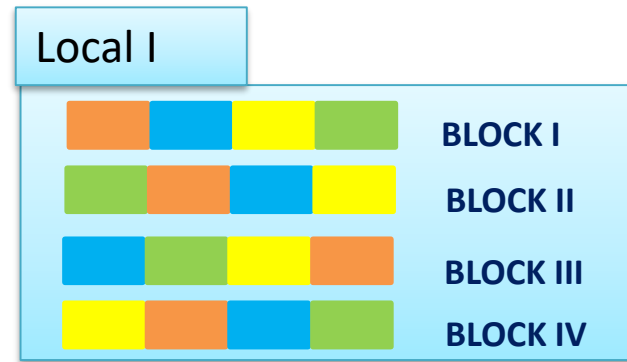


EXEMPLO: RCBD repetido em mais de um local

- Os blocos são dispostos em mais do que um local: factor bloco subordinado ao local; estrutura hierarquizada (nested structure).
- Os tratamentos são atribuídos aleatoriamente a esses blocos.
- Os tratamentos são atribuídos aleatoriamente dentro de blocos, cada tratamento uma vez por bloco.
- O número de blocos é o número de repetições.
- Qualquer tratamento pode ser adjacente a qualquer outro tratamento, mas não ao mesmo tratamento dentro do bloco.

3 factores em estudo: há 4 tratamentos (cores diferentes que podem ser 4 variedades diferentes (ou 4 sistemas de condução, 4 fertilizantes, etc.), 4 locais e 4 blocos por local.

Como em cada local há repetições independentes de cada tratamento pode ser estudada a interação tratamento×local.



Delineamento em Quadrado Latino (LSD) (*Latin Square Design*)

Útil quando padrões de heterogeneidade estão associados a 2 factores com o mesmo número de níveis e configuração quadrada (chamados Linha e Coluna). O número de tratamentos (o número de níveis do factor que representa o objectivo do estudo) tem de ser igual ao número de linhas e de colunas e a alocação dos tratamentos é feita de forma a que cada tratamento apareça exactamente uma única vez em cada linha e uma única vez em cada coluna.

Por exemplo: 4 linhas (L1, L2, L3, L4), 4 colunas (C1, C2, C3, C4) e 4 tratamentos (A,B,C,D)
3 factores em estudo: factor variedade (ou, fertilizante, etc.) (4 níveis), factor Linha (4 níveis) e Factor Coluna (4 níveis).



	C1	C2	C3	C4
L1	A	C	B	D
L2	B	D	C	A
L3	C	A	D	B
L4	D	B	A	C

Mas, quando o número de tratamentos é elevado (nº de níveis de 1 factor), torna-se difícil garantir a homogeneidade dentro do bloco!

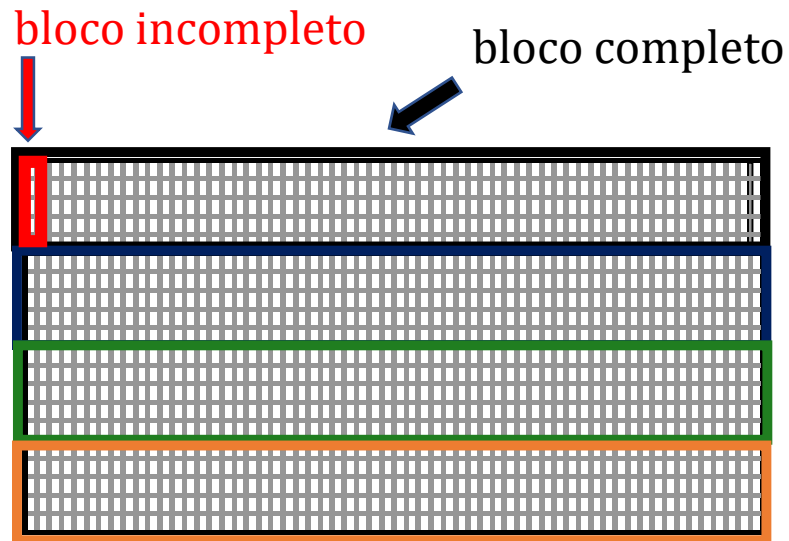
Surgiram depois os delineamentos pertencentes à classe dos

Delineamentos em Blocos Incompletos,

direccionados para o controlo da heterogeneidade (nomeadamente, espacial em ensaios agrícolas) em ensaios com maior número de tratamentos.

- Quando o número de tratamentos (v) é grande, o controlo da heterogeneidade poderá ser conseguido constituindo pequenos conjuntos homogéneos de unidades experimentais, isto é:
uma repetição de tamanho v pode ser dividida em s blocos incompletos, cada um de tamanho k ($k < v$), sendo $v = ks$.

Um delineamento com r repetições (blocos completos) deste tipo é designado por [Delineamento Experimental em Blocos Incompletos Resolúvel](#).

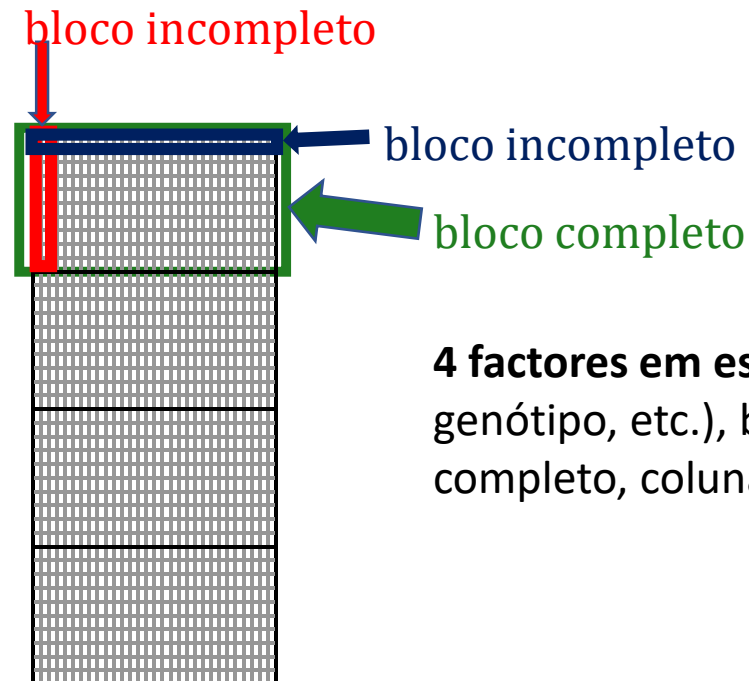


- Os delineamentos em blocos incompletos equilibrados originais de *Yates* (1936; 1940) requerem que v seja o quadrado de k .
- [Delineamentos em blocos incompletos resolúveis, particularmente adequados para ensaios com um grande número de tratamentos, são os delineamentos alfa](#) (Patterson e Williams, 1976): v/k um número inteiro; > eficiência quando o tamanho do bloco, k , for menor que a raiz quadrada do número de tratamentos, v , e então, menor que s (Giesbrecht e Gumpertz, 2004).

3 factores em estudo: factor que se quer estudar (variedade, genótipo, etc.), bloco completo e bloco incompleto subordinado ao bloco completo.

O Delineamento Linha-Coluna Resolúvel (RCD) (Williams e John, 1989) é usado quando se suspeita da existência de variação espacial em duas direções no campo, sendo as **linhas e colunas** utilizadas conjuntamente para introduzir dois fatores de controlo da heterogeneidade espacial (generalização dos quadrados latinos para situações mais complexas).

Neste caso, **têm-se blocos incompletos simultaneamente nas linhas e nas colunas** e o delineamento é resolúvel se, em cada repetição, o número de tratamentos v for igual ao número de linhas l , multiplicado pelo número de colunas c , isto é, $v = lc$.



4 factores em estudo: factor que se quer estudar (variedade, genótipo, etc.), bloco completo, linha subordinada ao bloco completo, coluna subordinada ao bloco completo.

Delineamento em parcelas divididas (*split-plot design em RCB*)

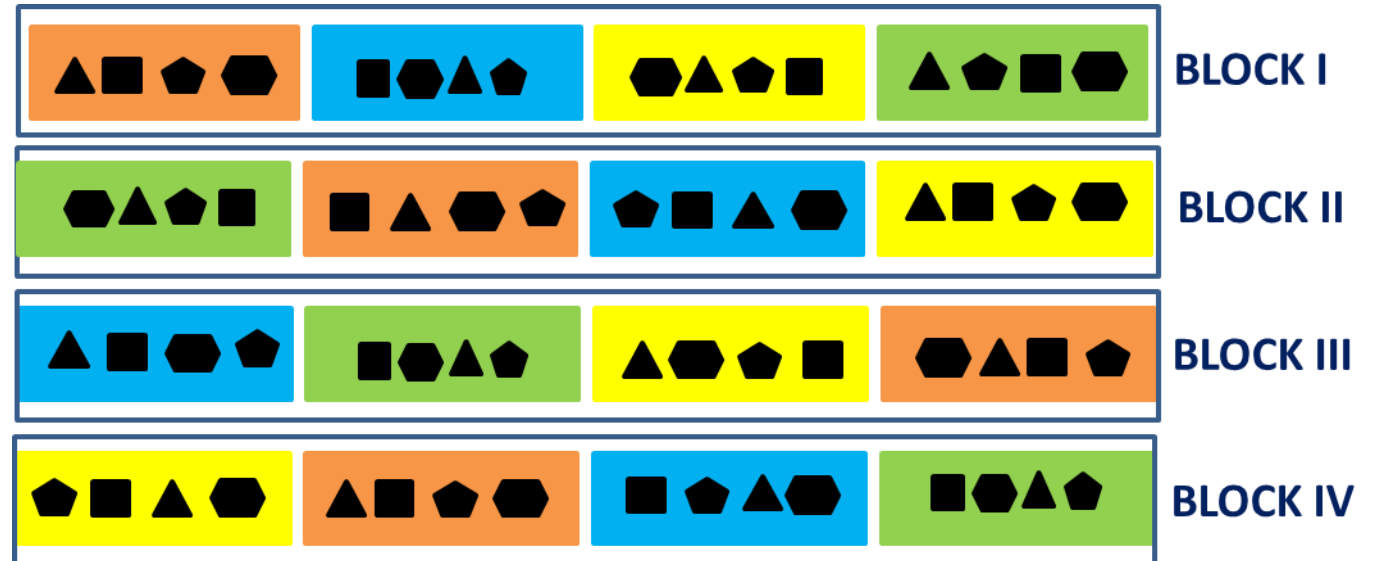
- O campo é dividido em unidades homogéneas, de forma a controlar a variação no campo, estufa, etc.
- O número de blocos é o número de repetições.
- Existem dois outros fatores em estudo (o objectivo do estudo). As unidades experimentais são organizadas em blocos, cada bloco com grandes parcelas (*whole plots*), tantas quantas o número de níveis do factor A (factor principal), e cada grande parcela é depois dividida em pequenas parcelas (*split plots* ou *subplots*), às quais são aleatoriamente atribuídos os níveis do factor B.

Exemplo:

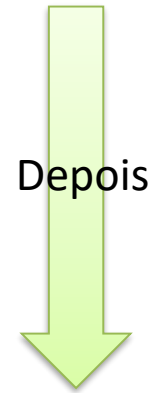
Cada linha representa 1 bloco.

Existem 4 blocos (I-IV) cada um com 4 tratamentos principais (cores, níveis do factor A) divididos em 4 outros tratamentos (*subplot*, símbolos, níveis do factor B).

3 factores em estudo e várias interações



E muitos outros delineamentos experimentais...



Modelos de análise de dados adequados
(devem traduzir o delineamento experimental adotado)