XXXII Ciclo de Atualização em Ciências Agrárias: Estatítica Experimental com Software R

Éder David Borges da Silva
Renato Gonçalves de Oliveira

Página do curso:

http://www.leg.ufpr.br/ragronomia

Porque Planejar o Experimento?

Vamos a um exemplo...

- Um experimento foi realizado para avaliar de que forma se distribuía uma determinada característica;
- A suposição é de que a característica atinge 50% de uma determinada população;
- Para testar o ensaio será utilizado o teste qui-quadrado;

Teste Qui-quadrado

N° de amostras	Observado	Esperado	Desvio	Desvio²	X²
10	6	5	-1	1	0,2
	4	5	1	1	0,2
				Total	0,4

$X^2(5\%) = 3,84$

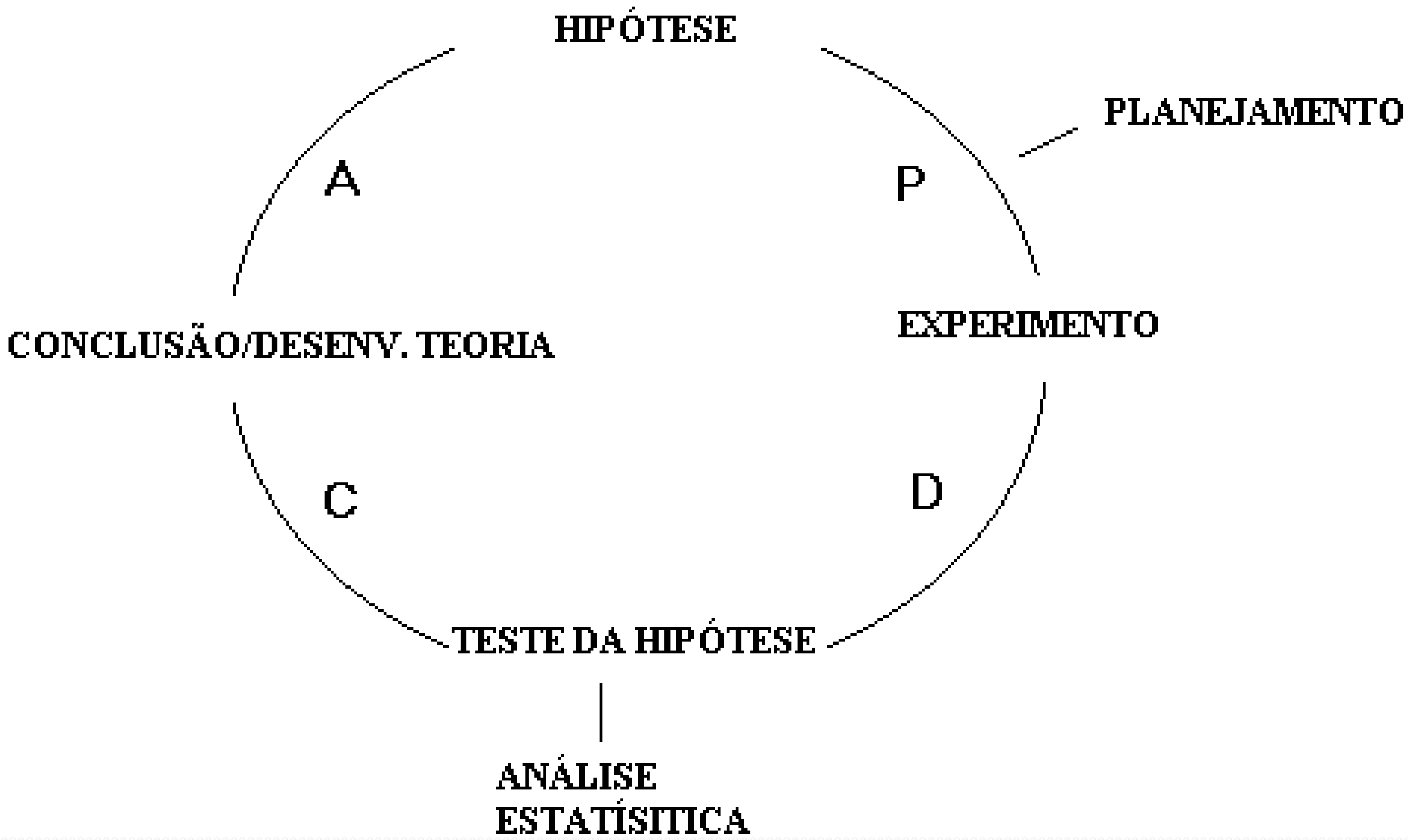
Teste Qui-quadrado

N° de amostras	Observado	Esperado	Desvio	$Desvio^2$	X^2
100	60	50	10	100	2
	40	50	-10	100	2
				Total	4

$X^2(5\%) = 3{,}84$

Planejamento Experimental

Circularidade do Método científico



Princípios básicos de experimentação

Princípios básicos de experimentação

- Casualização
- Repetição
- Controle local

✓ A aleatorização torna os testes estatísticos validos
✓ A repetição torna os teste estatísticos possíveis
✓ O controle local torna o experimento mais eficiente

Delineamentos Experimentais

Um exemplo de aplicação

Dado a demanda para um experimento de adubação de milho (2 hibridos, 2 isolinha BT) e 4 doses de uréia (0,50,100,150 e 200 kg/ha), considerando que se possue 1 ha para o experimento, desenvolva o planejamento com os principais delineamento, mostrando suas pontos positivos e negativos.

Delineamento inteiramente casualizado (DIC)

Modelo estatístico inteiramente casualizado

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad \text{onde} \quad \varepsilon_{ij} \overset{IDD}{\sim} N(0, \sigma^2)$$

$y_{ij} =$ é o valor do i-ésimo tratamento na j-ésima parcela

$\mu =$ é a constante geral do modelo (normalmente a média)

$\tau_i =$ é o efeito do i-ésimo tratamento

$\varepsilon_{ij} =$ é o erro experimental associado ao i-ésimo tratamento na j-ésima parcela

Quadro de anova DIC

Fonte	*GL*	*SQ*	*QM*	*EMEF*	*EMEA*
Tratamento	$t-1$	$SQtrat$	$\dfrac{SQtrat}{t-1}$	$\sigma^2 + r\dfrac{\sum \alpha_i^{\,2}}{t-1}$	$\sigma^2 + r\sigma_\alpha^2$
Erro	$(t-1)(r-1)$	$SQerro$	$\sigma^2 \; \dfrac{SQerro}{(t-1)(r-1)}$	σ^2	σ^2
Total	$tr-1$	$SQtotal$			

Delineamento blocos ao acaso (DBA)

O delineamento em blocos casualizados

- ✓ Um Bloco é uma restrição á casualização.
- ✓ Se não for utilizado considerando esse princípio, provavelmente deve ser um outro fator e deve ser tratado como tal.
- ✓ Portanto, tem-se um experimento fatorial.

Modelo estatístico Blocos ao acaso

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \quad \text{onde} \quad \varepsilon_{ij} \overset{IDD}{\sim} N(0, \sigma^2)$$

$y_{ij} =$ é o valor do i-ésimo tratamento na j-ésima parcela

$\mu =$ é a constante geral do modelo (normalmente a média)

$\tau_i =$ é o efeito do i-ésimo tratamento

$\beta_j =$ é o efeito do j-ésimo bloco

$\varepsilon_{ij} =$ é o erro experimental associado ao i-ésimo tratamento na j-ésima parcela

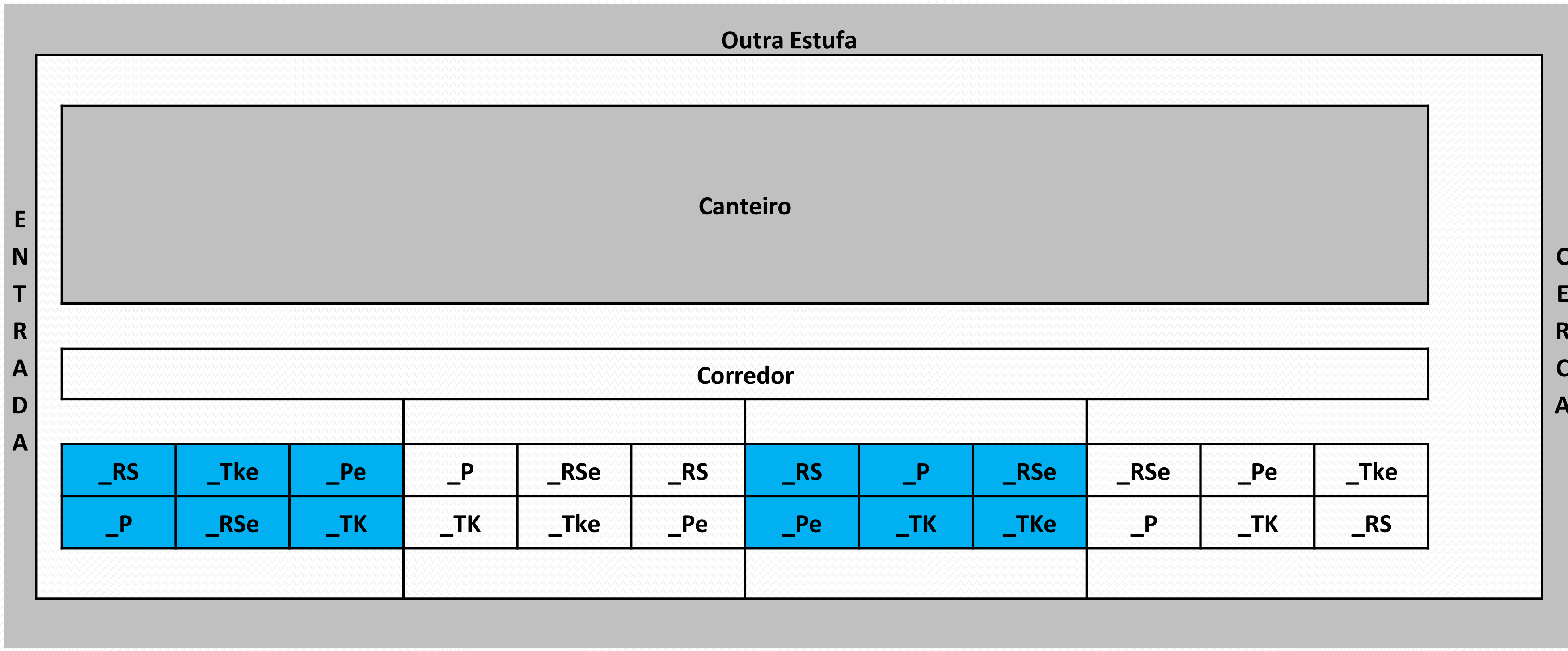
Quadro de anova DBA

Fonte	GL	SQ	QM	EMEF	EMEA
Bloco	$j-1$	$SQbloco$	$QMbloco$	$\sigma^2+\dfrac{t\sum\beta_j^2}{j-1}$	$\sigma^2+t\sigma_\beta^2$
Tratamento	$t-1$	$SQtrat$	$QMtrat$	$\sigma^2+\dfrac{j\sum t_j^2}{t-1}$	$\sigma^2+j\sigma_t^2$
Erro	$(j-1)(t-1)$	$SQerro$	$QMerro$	σ^2	σ^2
Total	$jt-1$	$SQtotal$	$QMtotal$		

DBA

I	**a**	**c**	**b**
II	c	b	a
III	c	a	b

DBA


Outra Estufa
Canteiro
Corredor
ENTRADA
CERCA
_RS _Tke _Pe _P _RSe _RS _RS _P _RSe _RSe _Pe _Tke
_P _RSe _TK _TK _Tke _Pe _Pe _TK _TKe _P _TK _RS

Delineamento Quadrado latino (DQL)

Modelo estatístico Quadrado latino

$$y_{ijk} = \mu + \tau_i + l_j + c_k + \varepsilon_{ijk} \text{ onde } \varepsilon_{ijk} \overset{IDD}{\sim} N(0, \sigma^2)$$

$y_{ij} =$ é o valor do i-ésimo tratamento na j-ésima parcela

$\mu =$ é a constante geral do modelo (normalmente a média)

$\tau_i =$ é o efeito do i-ésimo tratamento

$l_j =$ é o efeito da j-ésima linha

$c_k =$ é o efeito da k-ésima linha

$\varepsilon_{ij} =$ é o erro experimental associado ao i-ésimo tratamento na j-ésima parcela

Delineamento Quadrado latino

- **Casualização:**

Somente uma repetição de cada tratamento apareça em cada bloco (Linhas e Colunas).

- **Limitação:**

O número de tratamentos deve ser igual ao número de repetições. Muitas vezes, não há material suficiente para completar o delineamento.

- **Desvantagem:**

O número de repetições aumenta á medida que o número de tratamentos também aumenta.

Delineamento Quadrado latino

Linhas	Colunas				
	I	II	III	IV	V
I	D	A	B	C	E
II	C	E	A	B	D
III	E	B	C	D	A
IV	B	D	E	A	C
V	A	C	D	E	B

Esquema Fatorial

Esquema fatorial

Esquema fatorial não é um delineamento apenas uma arranjo entre os tratamentos .

Experimento fatorial podem ser conduzido:

- ✓ Delineamento completamente casualizado
- ✓ Blocos casualizados
- ✓ Quadrado latino
- ✓ Outros

Esquema fatorial

Definições

- ✓ Fator: uma causa de variação conhecida e de interesse do pesquisador (um tipo de tratamento);
- ✓ Nível: subdivisão do fator;
- ✓ Efeito principal: pode-se estudar isoladamente o efeito de cada fator no experimento;
- ✓ Efeito da interação: quando existir, estudar o comportamento de cada fator, na presença ou ausência de níveis dos demais fatores

Modelo estatístico Esquema Fatorial

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \text{ onde } \varepsilon_{ij} \overset{IDD}{\sim} N(0, \sigma^2)$$

$y_{ijk} =$ é o valor do i-ésimo tratamento na j-ésima parcela e na k-ésima repetição

$\mu =$ é a constante geral do modelo (normalmente a média)

$\alpha_i =$ é o efeito do i-ésimo nível do fator A

$\beta_j =$ é o efeito da j-ésimo nível do fator B

$(\alpha\beta)_{ij} =$ é o efeito da interação entre i-ésimo nível do fator A e j-ésimo nível do fator B

$\varepsilon_{ijk} =$ é o erro experimental entre i-ésimo nível do fator A e j-ésimo nível do fator B na k-ésima repetição.

Quadro de anova fatorial

Fonte	GL	SQ	QM	EMEF	EMEA
$Tratamento A$	$a-1$	SQA	QMA	$\sigma^2 + br\dfrac{\sum \alpha_i^2}{a-1}$	$\sigma^2 + r\sigma_{\alpha\beta}^2 + br\sigma_\alpha^2$
$Tratamento B$	$b-1$	SQB	QMB	$\sigma^2 + ar\dfrac{\sum \beta_j^2}{t-1}$	$\sigma^2 + r\sigma_{\alpha\beta}^2 + ar\sigma_\beta^2$
AB	$(a-1)(b-1)$	$SQAB$	$QMAB$	$\sigma^2 + r\dfrac{\sum (\alpha\beta)_{ij}^2}{(a-1)(b-1)}$	$\sigma^2 + r\sigma_{\alpha\beta}^2$
$Erro$	$ab(r-1)$	$SQErro$	$QMErro$	σ^2	σ^2

Delineamento em parcela subdividida

Modelo estatístico Delineamento em parcela subdividida

$$y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} + \gamma_k + \delta_{ik} + \varepsilon_{ijk} \text{ onde } \varepsilon_{ijk} \overset{IDD}{\sim} N(0, \sigma^2)$$

$y_{ijk} =$ é o valor do i-ésimo tratamento na j-ésima parcela e na k-ésima repetição

$\mu =$ é a constante geral do modelo (normalmente a média)

$\alpha_i =$ é o efeito do i-ésimo nível do fator A

$\beta_j =$ é o efeito da j-ésimo bloco

$\varepsilon_{ij} =$ é o erro experimental entre i-ésimo nível do fator A e j-ésimo bloco

$\gamma_k =$ é o efeito da k-ésimo nível do fator B

$\delta_{ik} =$ é o efeito da interação entre o i-ésimo nível do fator A e da k-ésimo nível do fator B

$\varepsilon_{ijk} =$ é o erro experimental entre i-ésimo nível do fator A e j-ésimo bloco e k-ésimo nível do fator B

Anova de split-plot

$Fonte$	GL	SQ	QM	$EMEF$	$EMEA$
$Bloco$	$r-1$	$SQBloco$	$QMBloco$	$\sigma^2+b\sigma_\gamma^2+ab\sigma_p^2$	$\sigma^2+b\sigma_\gamma^2+ab\sigma_p^2$
A	$a-1$	SQA	QMA	$\sigma^2+b\sigma_\gamma^2+rb\dfrac{\sum\alpha_i^2}{a-1}$	$\sigma^2+b\sigma_\gamma^2+rb\sigma_\alpha^2$
$ErroA$	$(r-1)(a-1)$	$SQerroa$	$QMerroa$	$\sigma^2+b\sigma_\gamma^2$	$\sigma^2+b\sigma_\gamma^2$
B	$b-1$	SQB	QMB	$\sigma^2+ra\dfrac{\sum\beta_j^2}{b-1}$	$\sigma^2+ra\sigma_\beta^2$
AB	$(a-1)(b-1)$	$SQAB$	$QMAB$	$\sigma^2+r\dfrac{\sum(\alpha\beta)_{ij}^2}{(a-1)(b-1)}$	$\sigma^2+r\sigma_{\alpha\beta}^2$
$ErroB$	$a(b-1)(r-1)$	$SQerrob$	$QMerroB$	σ^2	σ^2
$Total$	$abr-1$	$SQtotal$			

Experimento Split-plot

Quando instalar:

- ✓Em experimentos fatoriais com dois ou mais fatores;
- ✓Quando há alguma limitação para instalar o experimento;
- ✓ Facilidade para instalação.
- ✓ Em alguns casos, é a única forma de aplicação dos tratamentos às unidades experimentais.

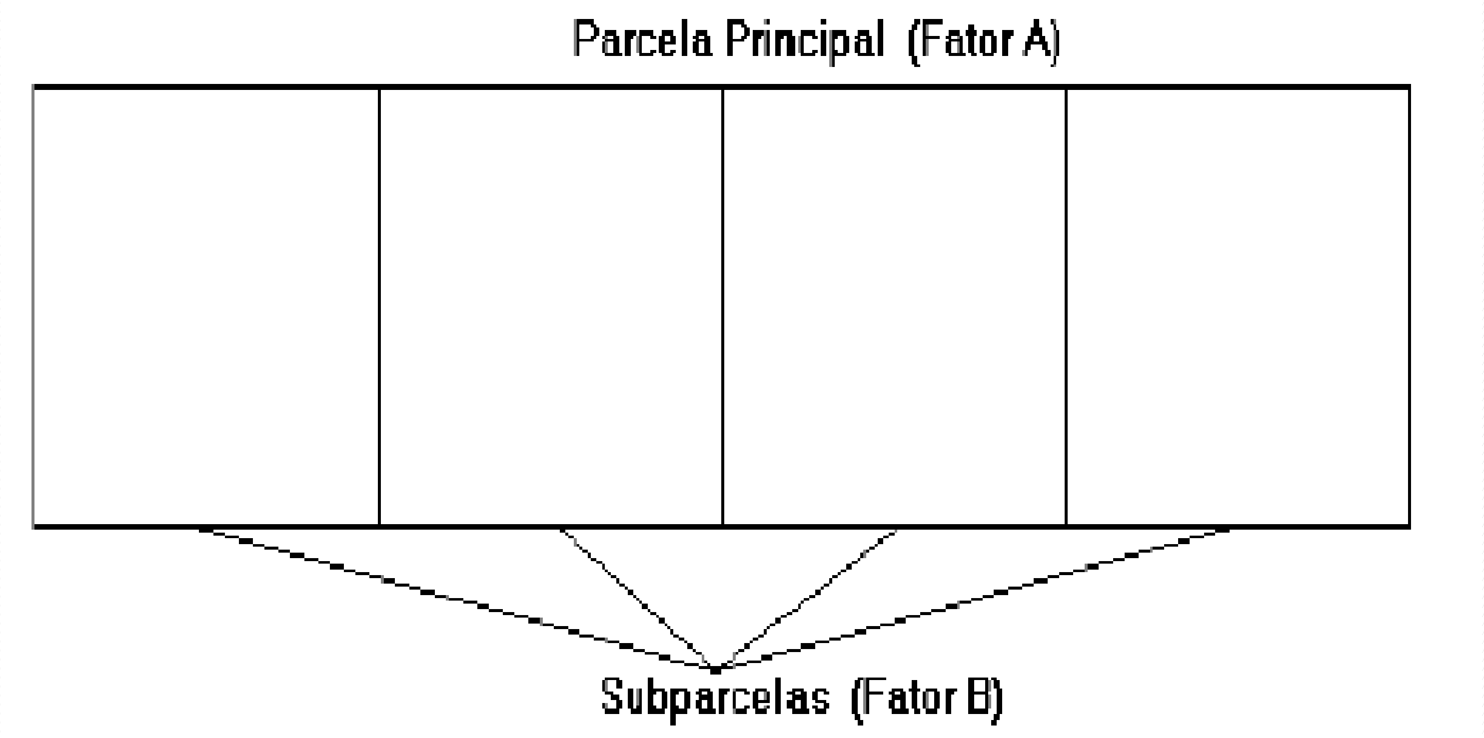
Experimento Split-plot

Definição:

- ✓ Este tipo de experimento aloca o fator A em parcelas principais (ou primária) e o fator B nas sub-parcelas (secundárias).
- ✓ Cada parcela funciona como um “bloco”para as sub-parcelas.
- ✓ Obs: *Se existirem mais de dois fatores, o experimento é chamado de parcelas sub-subdivididas e assim por diante.*

Experimento Split-plot

- Croqui de uma parcela principal de um experimento em Parcelas subdivididas

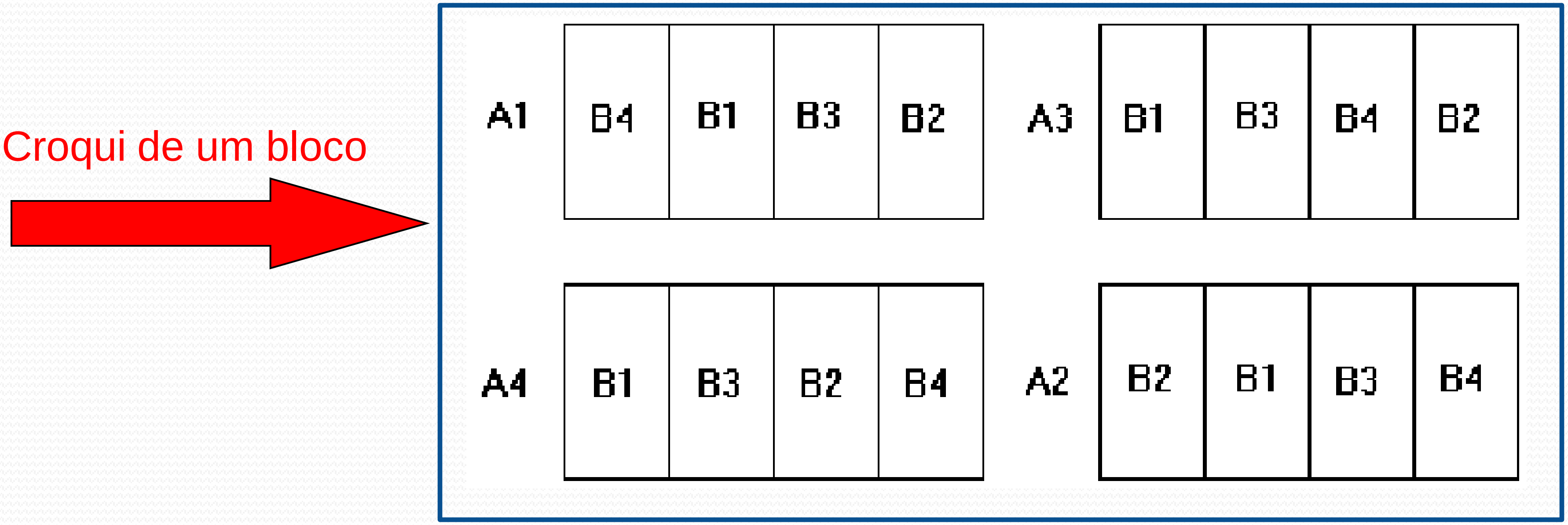


Experimento Split-plot

- Por exemplo: Experimento com 2 fatores (A e B), cada um com 4 níveis, dispostos em 3 blocos:
- A = A1; A2; A3; A4
- B = B1; B2; B3; B4
- BLOCO = I; II; IV

Croqui de um bloco

Bloco I



Experimento Split-plot

Instalação:

- ✓Primeiro deve-se casualizar os níveis do Fator A (Parcela Principal);
- ✓ Segundo, deve-se casualizar os níveis do fator B (Sub-parcelas) dentro do bloco.

Experimento Split-plot

- O fator de maior interesse é colocado nas sub-parcelas, quando possível.

Caso contrário

- Aplicação dos tratamentos às parcelas principais ou sub-parcelas, dependerá da facilidade de instalação do experimento.

Delineamento em faixas

Experimento Em Faixas ou Split Block

- É uma variação dos experimentos em parcelas subdivididas.
- Os fatores A e B, são dispostos em faixas, como se fossem parcelas principais
- Os dois fatores são alocados em parcelas principais.

Modelo estatístico Delineamento em faixas

$$y_{ijk} = \mu + \beta_j + \alpha_i + \varepsilon_{ij} + \gamma_k + \varepsilon_{jk} + (\alpha\gamma)_{ij} + \varepsilon_{ijk} \text{ onde } \varepsilon_{ijk} \overset{IDD}{\sim} N(0, \sigma^2)$$

$y_{ijk} =$ é o valor do i-ésimo tratamento A e k-ésimo tratamento B no j-ésima bloco

$\mu =$ é a constante geral do modelo (normalmente a média)

$\beta_j =$ é o efeito do j-ésimo bloco

$\alpha_i =$ é o efeito da i-ésimo tratamento A

$\varepsilon_{ij} =$ é o erro experimental entre i-ésimo nível do fator A e j-ésimo bloco

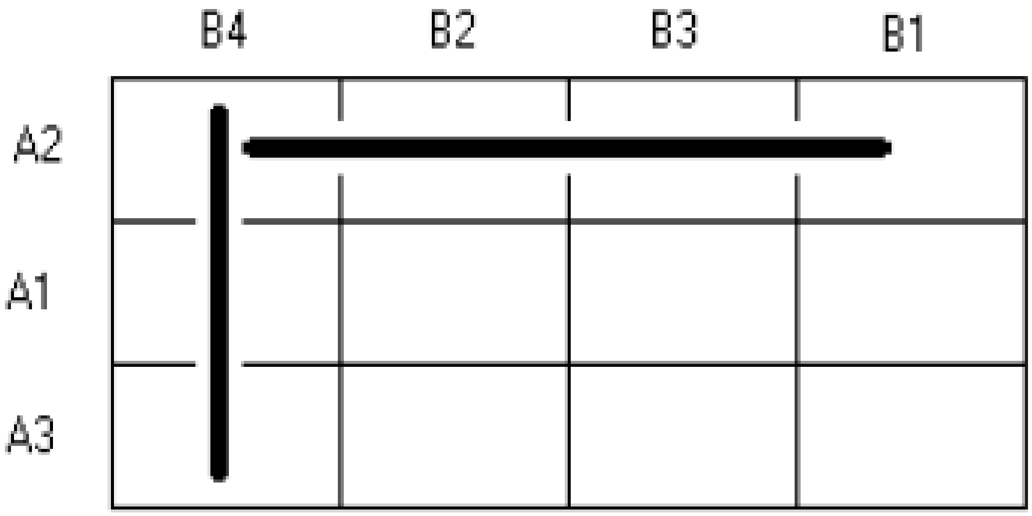
$\gamma_k =$ é o efeito da k-ésimo nível do fator B

$\varepsilon_{jk} =$ é o erro experimental entre k-ésimo nível do fator B e j-ésimo bloco

$(\alpha\gamma)_{ij} =$ é o efeito da interação entre o i-ésimo nível do fator A e da k-ésimo nível do fator B

$\varepsilon_{ijk} =$ é o erro experimental entre i-ésimo nível do fator A e k-ésimo nivel de B no j-ésimo bloco

Exemplo do Delineamento em faixas



Procedimentos de comparações múltiplas

Procedimentos para comparações múltiplas

- Teste de Tukey ou DHS (HSD):

$$\Delta = q_{(p,v,\alpha)}\sqrt{\frac{QME}{r}}$$

- Teste de Duncan:

$$\Delta = q_{(p,v,\alpha_p)}\sqrt{\frac{QME}{r}}$$

Procedimentos para comparações múltiplas

- Teste Scott-Knott:

$$B_0 = \frac{T_1^2}{k_1} + \frac{T_2^2}{k_2} - \frac{(T_1 + T_2)^2}{k_1 + k_2}$$

$$\hat{\sigma}_0^2 = \frac{1}{g+v}\left[\sum_{i=1}^{g}(\bar{Y}_{(i)} - \bar{Y})^2 + v_{s_{\bar{Y}}^2}\right]$$

$$\lambda = \frac{\pi}{2(\pi - 2)} * \frac{\beta_0^2}{\hat{\sigma}_0^2}$$

Procedimentos para comparações múltiplas

- Um teste pode ter dois parâmetros;

  - Poder do teste:

  Capacidade do teste em detectar diferenças reais entre os tratamentos.

  - Rigorosidade:

  Confiança no resultado obtido.

Procedimentos para comparações múltiplas

- Erro tipo I por comparação:

  Probabilidade de se rejeitar uma hipótese verdadeira nas comparações dos tratamentos tomados dois a dois;

- Erro tipo I por experimento:

  Probabilidade de se realizar pelo menos uma inferência errada por experimento;

- Erro tipo III:

  Probabilidade de se classificar um tratamento superior ao outro quando o segundo supera o primeiro;

Procedimentos para comparações múltiplas

Linhagens	Médias	
1	14,65	a
2	12,34	ab
3	10,42	b

Ambigüidade dos resultados

Procedimentos para comparações múltiplas

	Valores médios		Procedimentos de comparações múltiplas				
Trat.	Par.	Est.	Tukey	SNK	LSD	LSDB	SK
3	85	80,461	D	C	E	D	B
1	85	83,498	D	C	DE	CD	B
4	85	86,488	CD	C	DE	CD	B
2	85	90,742	CD	BC	DE	CD	B
6	95	95,986	CD	BC	CD	BCD	B
5	95	96,511	BCD	BC	CD	BCD	B
7	105	107,689	ABC	AB	BC	ABC	A
8	115	120,436	AB	A	AB	AB	A
9	125	123,492	A	A	A	A	A
10	125	123,942	A	A	A	A	A

Procedimentos para comparações múltiplas

- Comparação do teste de Scott-Knott com os demais:
  - Comparações realizadas através de experiemntos com dados simulados;
  - Taxas de erro tipo I sempre abaixo do nível de significância (menores que α);
  - O poder do teste foi duas vezes maiores que os do teste de Duncan, t e SNK, oito vezes mais poderoso que o teste Tukey e semelhante ao t-Bayesiano;

Novas ferramentas para análise

- GLM:
  - Modelos Lineares Generalizados;
- REML:
  - Máxima verossimilhança restrita;
- BLUP:
  - Melhor predição linear não viesada;
- Selegen:
  - Sistemas REML/BLUP;

Por enquanto, é isso...