

# MAT02014 - Planejamento de Experimentos II

Delineamentos em Blocos Incompletos

Rodrigo Citton P. dos Reis

Departamento de Estatística



# Introdução

Delineamentos em **blocos completos casualizados**:

- Agrupam unidades experimentais heterogêneas em subgrupos (**blocos**) de unidades experimentais homogêneas.
  - **Redução da variância**,  $\sigma_{\epsilon}^2$ , do erro experimental.
    - **Aumenta o poder** para detectar efeitos dos fatores de tratamento.

# Introdução

Uma restrição dos delineamentos BCC é que o número de unidades experimentais em cada bloco **deve ser igual ou maior** ao número de níveis do fator de tratamento.

- Delineamento BCC:  $t$  níveis do fator de tratamento.
  - $t$  unidades experimentais por bloco.
- Delineamento BCG:  $t$  níveis do fator de tratamento e  $r$  replicações.
  - $tr$  unidades experimentais por bloco.

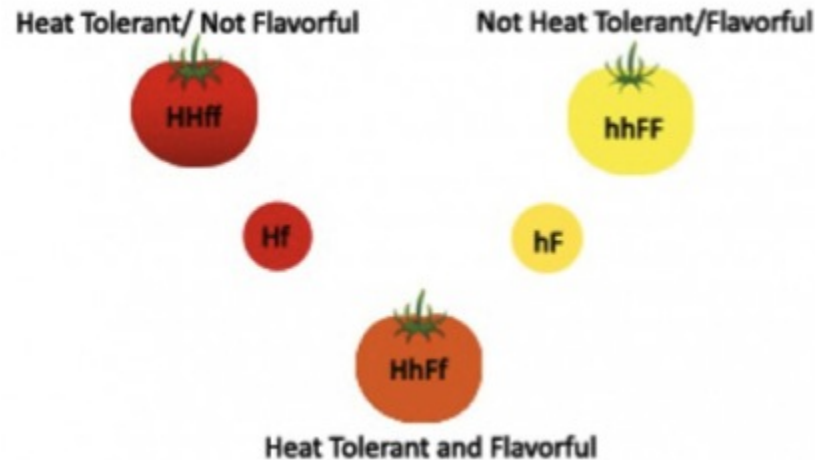
# Introdução

Número muito grande de níveis de tratamento.

- Blocos de tamanho muito grande.
  - Possível perda da homogeneidade das unidades experimentais de cada bloco.
    - (Aumento da variância do erro experimental; Baixo poder para detectar diferenças entre os níveis do fator de tratamento)

# Exemplo: experimentação agrícola

- Novas culturas híbridas
  - Número de culturas (níveis de tratamento) pode chegar a 50, 100.



- Como as unidades experimentais são lotes de terra, quanto maior o número de parcelas incluídas em um bloco, mais diversificadas serão as parcelas, uma vez que elas cobrem uma área física mais ampla.

# Introdução

Uma solução para o problema é construir um delineamento em blocos em que cada bloco contém apenas um subconjunto dos possíveis níveis do fator de tratamento.

- O tamanho do bloco pode se manter pequeno.

Estes delineamentos são chamados de **delineamentos em blocos incompletos**.

- **Blocos incompletos balanceados (BIB)**
- **Blocos incompletos parcialmente balanceados (BIPB)**

Delineamentos em blocos incompletos são comumente usados em experimentação agrícola, ciência animal, testes educacionais e ciência de alimentos.

# Exemplo

- Na ciência dos alimentos, painéis de sabor são frequentemente usados para testar a palatabilidade de novas receitas para novos produtos alimentícios.

**LIMÃO, COM UM DELICIOSO SABOR DE TAMARINDO!**

**APROVEITE NOSSAS  
PROMOÇÕES!**

**300ml - 50 cruzeiros**

**500ml - Vintão**

**E depende da  
cara do freguês...**

**É UMA  
DELÍCIA!**

**ADQUIRA O SEU EM QUALQUER TIENDA DEL CHAVO!**



# Exemplo

- Nestes painéis, os sujeitos são convidados a provar e avaliar diferentes receitas.
- As diferentes receitas representam os níveis do fator de tratamento.
- Uma amostra aleatória de indivíduos (as unidades experimentais) é incluída para representar o mercado potencial.
- Uma vez que existe uma grande variabilidade no gosto entre os participantes, é melhor usar um design bloqueado, onde cada sujeito saboreia e classifica cada receita.



# Exemplo

- No entanto, se houver muitas receitas a serem testadas, os participantes perderão sua capacidade de discriminar e será difícil detectar diferenças.



- Uma solução é fazer com que cada sujeito (bloco) prove apenas um subconjunto das receitas.

# Blocos incompletos balanceados

- Deve-se ter atenção na escolha do subconjunto das combinações de tratamentos testadas em cada bloco.
  - Se um nível de tratamento fica de fora de **todos** os blocos, este não pode ser comparado com os demais níveis de tratamento.
  - Se diferentes níveis de tratamento são representados de forma desigual, algumas comparações de níveis do fator terão uma precisão maior que outras.

# Blocos incompletos balanceados

A maneira ideal de criar um delineamento em blocos incompletos é ter cada nível de tratamento **igualmente replicado** e aparecendo dentro de um bloco, com todos os outros níveis de tratamento um número igual de vezes.

- Isso é chamado de delineamento em *blocos incompletos balanceados* (**BIB**).
- Fazendo isto, todas as diferenças de pares de médias de tratamentos de mínimos quadrados terão **o mesmo erro padrão**, e o poder para detectar uma diferença em quaisquer duas médias será o mesmo.

# Blocos incompletos balanceados

A maneira mais simples de construir um BIB para o caso em que há  $t$  níveis do fator de tratamento e  $k < t$  unidades experimentais por bloco é formar todos os subconjuntos possíveis de  $k$  níveis de tratamento escolhidos a partir de  $t$ .

- **Por exemplo**, no painel de sabor descrito anteriormente, se houvesse  $t = 6$  receitas a serem testadas, e foi determinado que cada indivíduo poderia saborear no máximo  $k = 3$  receitas sem perder poder discriminatório,  $\binom{6}{3} = 20$  indivíduos seriam necessários.
- Todos os subconjuntos possíveis estão listados abaixo.

(1 2 3), (1 2 4), (1 2 5), (1 2 6), (1 3 4)  
(1 3 5), (1 3 6), (1 4 5), (1 4 6), (1 5 6)  
(2 3 4), (2 3 5), (2 3 6), (2 4 5), (2 4 6)  
(2 5 6), (3 4 5), (3 4 6), (3 5 6), (4 5 6)

- O sujeito um experimentaria as receitas 1, 2 e 3 em uma ordem aleatória.
- O sujeito 2 saborearia receitas 1, 2 e 4, e assim por diante.
- Este plano é completamente balanceado na medida em que cada receita (**níveis do tratamento**) é replicado  $r = 10$  vezes (**provado por 10 indivíduos**) e cada **par de níveis** de tratamento ocorre junto no mesmo bloco  $\lambda = 4$  vezes.
  - Os níveis de tratamento 1 e 2 ocorrem juntos apenas nos primeiros quatro blocos.

# Blocos incompletos balanceados

A forma mais simples de construir um BIB pode requerer muitos blocos.

Se a precisão do delineamento não exigir  $\binom{t}{k}$  blocos, não haverá necessidade de se usar tantos blocos.

- Por exemplo, se um tamanho de diferença prática em receitas (médias dos escores) pudesse ser detectada com menos de 20 indivíduos e 10 repetições de cada nível de tratamento, no painel de teste de sabor, talvez um BIB pudesse ser encontrado com menos de 20 blocos.

# Blocos incompletos balanceados

Se

- $r$  é o número de vezes que um nível de tratamento é replicado em um delineamento em blocos incompletos;
- $\lambda$  é o número de vezes que cada nível de tratamento ocorre com todo o outro nível de tratamento no mesmo bloco;
- $t$  é o número de níveis do fator de tratamento;
- $k$  é o número de unidades experimentais em um bloco;
- $b$  é o número de blocos

os seguintes requisitos devem ser atendidos para que esse delineamento seja um delineamento BIB.

$$b \geq t$$

$$tr = bk$$

$$\lambda(t - 1) = r(k - 1)$$



# Blocos incompletos balanceados

As relações do slide anterior podem ser utilizadas para se encontrar o **número mínimo de blocos** necessários para um delineamento BIB.

- Da última equação, temos que  $\lambda(t - 1)$  deve ser divisível por  $k - 1$ .
  - No exemplo do painel de sabor, temos  $t = 6$  e  $k = 3$ ,  $5\lambda$  deve ser divisível por 2.
    - O menor inteiro  $\lambda$  para o qual este é satisfeito é  $\lambda = 2$ .
  - Assim, deverá ser possível encontrar um BIB com  $\lambda = 2$ ,  $r = 10/2 = 5$  e  $b = (6 \times 5)/3 = 10$ .

# Blocos incompletos balanceados

- Dado um número de níveis de tratamento  $t$  e o tamanho do bloco  $k$ , a função `BIBsize` do pacote `daewr` do R encontra valores de  $\lambda$  e  $r$  que satisfazem as equações listadas anteriormente.
  - Por exemplo,  $t = 6$  e  $k = 3$

```
library(daewr)  
BIBsize(6, 3)
```

```
## Possible BIB design with b= 10  and r= 5  lambda= 2
```

# Blocos incompletos balanceados

- A satisfação das equações listadas anteriormente é uma condição **necessária**, mas não é **suficiente**.
  - Assim, pode ser o caso que não consigamos encontrar um BIB.
- Uma possível solução:
  1. Avaliar se a condição necessária é satisfeita (BIBsize).
  2. Encontre um BIB com a função `optBlock` do pacote `AlgDesign`.
    - A opção `blocksizes = rep(3,10)` especifica 10 blocos de tamanho 3;
    - A opção `withinData = factor(1:6)` especifica 6 níveis de um fator.

```
library(AlgDesign)
BIB <- optBlock( ~ ., withinData = factor(1:6),
                blocksizes = rep(3, 10))
```

```
class(BIB)
```

```
## [1] "list"
```

```
names(BIB)
```

```
## [1] "D"           "diagonality" "Blocks"      "design"      "rows"
```

# Blocos incompletos balanceados

```
des <- BIB$rows
dim(des) <- NULL
des <- matrix(des, nrow = 10, ncol = 3, byrow = TRUE,
              dimnames = list(c( "Block1", "Block2", "Block3", "Block4",
                                "Block5", "Block6", "Block7", "Block8", "Block9",
                                "Block10"),
                              c("unit1", "unit2", "unit3"))))
des
```

##	unit1	unit2	unit3
## Block1	1	2	5
## Block2	2	3	4
## Block3	2	3	5
## Block4	3	5	6
## Block5	2	4	6
## Block6	1	3	6
## Block7	1	3	4
## Block8	4	5	6
## Block9	1	4	5
## Block10	1	2	6

- Note que  $r = 5$  e  $\lambda = 2$ .
- Após o delineamento BIB ter sido encontrado, os níveis de tratamento dentro de cada bloco deve ser aleatorizado (como no BCC).

# KRAFTFULL DAS MODEL

BEREICHUNG VERBUNDEN





TELE

# Análise de delineamentos BIB

O modelo para a análise de um delineamento BIB é

$$y_{ij} = \mu + b_i + \tau_j + \epsilon_{ij}.$$

- Este modelo é idêntico ao modelo para análise de um BCC.
  - No entanto, a análise é um pouco diferente devido aos dados ausentes.

## Exemplo: painel de sabores



# Exemplo: painel de sabores

Considere os dados do experimento de painéis de sabores.

- $t = 4$  níveis de tratamento (receitas).
- Tamanho do bloco  $k = 2$ .
  - Cada painelista (indivíduo) prova (saboreia) apenas duas receitas em uma ordem aleatória e atribui uma pontuação (escala de 1 a 10, em que o 10 representa a preferência máxima pela receita).

## Exemplo: painel de sabores

- Apenas  $\binom{4}{2} = 6$  blocos ou painelistas seriam necessários para um delineamento BIB, mas neste experimento este número foi dobrado para aumentar o poder de detectar diferenças.
  - Os seis primeiros painelistas e os últimos seis painelistas são uma repetição do mesmo delineamento BIB.
  - Os indivíduos foram aleatoriamente atribuídos a números de painel e a ordem das duas receitas a serem saboreadas foi sorteada para cada painalista.

## Exemplo: painel de sabores

Panelist	Recipe			
	A	B	C	D
1	5	5	-	-
2	7	-	6	-
3	5	-	-	4
4	-	6	7	-
5	-	6	-	4
6	-	-	8	6
7	6	7	-	-
8	5	-	8	-
9	4	-	-	5
10	-	7	7	-
11	-	6	-	5
12	-	-	7	4

## Exemplo: painel de sabores

- Ao analisar os dados de um BIB, as médias de tratamento marginal **não são estimadores não-viesados** dos efeitos estimáveis  $\mu + \tau_i$ .
  - Por exemplo, na Tabela acima, a média marginal da receita A pode ser influenciada pelo fato de que não foi provada pelos membros dos painéis 4, 5, 6, 10, 11 e 12 que parecem atribuir notas mais altas para as receitas provadas.
- Da mesma forma, o fator de não centralidade para as somas sequenciais de quadrados para tratamentos (ou receitas) pode conter efeitos de bloqueio, bem como efeitos de tratamento.
- As médias por **mínimos quadrados** são usadas em vez das médias marginais, e as somas de quadrados ajustadas (**tipo III**) para tratamentos devem ser usadas.
  - No R, a função Anova do pacote car pode ser usada para se obter as somas de quadrados ajustadas para receitas.
  - A função lm também pode ser utilizada, mas a ordem dos fatores na fórmula importa (receitas deve ser o último termo).
    - Esta função não irá produzir somas de quadrados ajustadas para blocos (painelistas), mas as diferenças entre os painelistas não são do interesse do experimento.

# Exemplo: painel de sabores

- Ordem correta:

```
library(daewr)
mod1 <- aov(score ~ panelist + recipe, data = taste)
summary(mod1)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## panelist     11 19.333   1.7576    2.301 0.1106
## recipe        3  9.125   3.0417    3.982 0.0465 *
## Residuals     9  6.875   0.7639
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Ordem incorreta:

```
mod2 <- aov(score ~ recipe + panelist, data = taste)
summary(mod2)
```

```
##              Df Sum Sq Mean Sq F value  Pr(>F)
## recipe        3 21.000   7.000    9.164 0.00425 **
## panelist     11  7.458   0.678    0.888 0.58110
## Residuals     9  6.875   0.764
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

# Exemplo: painel de sabores

```
library(car)
Anova(mod1, type = "III")
```

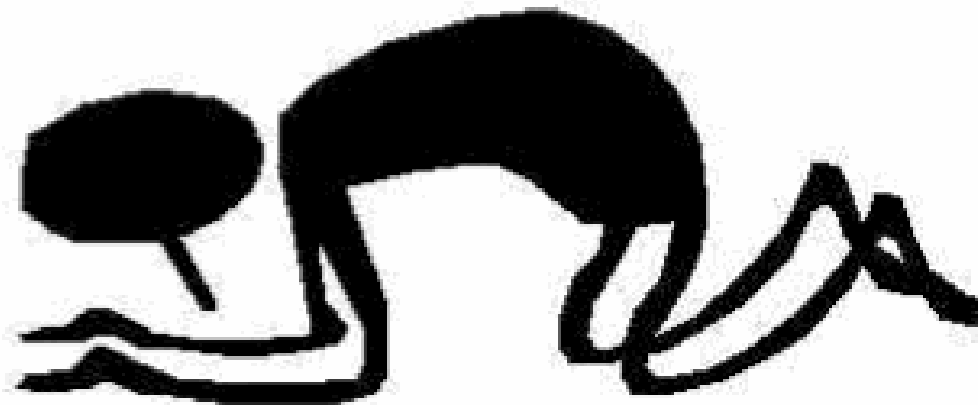
```
## Anova Table (Type III tests)
##
## Response: score
##           Sum Sq Df F value    Pr(>F)
## (Intercept) 34.225  1 44.8036 8.919e-05 ***
## panelist     7.458 11  0.8876  0.58110
## recipe       9.125  3  3.9818  0.04649 *
## Residuals    6.875  9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Anova(mod2, type = "III")
```

```
## Anova Table (Type III tests)
##
## Response: score
##           Sum Sq Df F value    Pr(>F)
## (Intercept) 34.225  1 44.8036 8.919e-05 ***
## recipe       9.125  3  3.9818  0.04649 *
## panelist     7.458 11  0.8876  0.58110
## Residuals    6.875  9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



**$p < 0.05$**



# Exemplo: painel de sabores

```
library(lsmeans)
lsmeans(mod1, pairwise ~ recipe, adjust = ("tukey"))
```

```
## $lsmeans
##   recipe    lsmean      SE df lower.CL upper.CL
##   A      5.458333 0.4183992  9 4.511849 6.404818
##   B      6.208333 0.4183992  9 5.261849 7.154818
##   C      6.833333 0.4183992  9 5.886849 7.779818
##   D      4.833333 0.4183992  9 3.886849 5.779818
##
## Results are averaged over the levels of: panelist
## Confidence level used: 0.95
##
## $contrasts
##   contrast estimate      SE df t.ratio p.value
##   A - B      -0.750 0.6180165  9  -1.214  0.6342
##   A - C      -1.375 0.6180165  9  -2.225  0.1882
##   A - D       0.625 0.6180165  9   1.011  0.7472
##   B - C      -0.625 0.6180165  9  -1.011  0.7472
##   B - D       1.375 0.6180165  9   2.225  0.1882
##   C - D       2.000 0.6180165  9   3.236  0.0421
##
## Results are averaged over the levels of: panelist
## P value adjustment: tukey method for comparing a family of 4 estimates
```



# Exemplo: painel de sabores

## Conclusões

- Os resultados mostram que a pontuação média da receita C é significativamente maior que a pontuação da receita D ao nível  $\alpha = 0,05$ , mas que nenhuma outra diferença é significativa ao nível  $\alpha = 0,05$ .
- Também deve ser notado na saída o fato de que o erro padrão das diferenças de médias é o mesmo para todos os pares de médias.
  - Isso é resultado do fato de que um delineamento BIB foi usado.

