MATO2014 - Planejamento de Experimentos II Delineamentos fatoriais fracionários

Rodrigo Citton P. dos Reis rodrigocpdosreis@gmail.com

Universidade Federal do Rio Grande do Sul Instituto de Matemática e Estatística Departamento de Estatística Porto Alegre, 2018

- Benefícios dos delineamentos fatoriais (com múltiplos fatores):
 - 1. Podemos detectar o efeito conjunto (interação).
 - 2. Eficiência.
- Maior o número de fatores:
 - Maior a eficiência;
 - Maior o número de interações que podem ser detectadas;
 - Maior o número de experimentos!

- Uma forma de reduzir o número de execuções do experimento é usar apenas dois níveis para cada fator e executar somente um experimento por célula.
- Na fase inicial de um experimento, em que o objetivo pode ser determinar quais fatores são importantes de uma longa lista de candidadtos, um delineamento fatorial pode necessitar de muitos experimentos para realizar, mesmo com apenas dois níveis e uma repetição por célula.

Table 6.1	Number of Exp	periments Required for 2^k Design
	Number	Number
	of	of
	Factors (k)	Experiments (2^k)
	4	16
	5	32
	6	64_
	7	128
	8	256
	9	512

- Um delineamento com 7 ou mais fatores torna o experimento impraticável.
- Soluções "menos que ótimas":
 - Delineamento "um por vez"
 - Escolha de um subconjunto de fatores.

- Uma solução melhor: usar uma fração de experimentos, ou execuções, necessárias para um experimento fatorial completo.
- Para ser efetiva, a fração de execuções usada deve ser cuidadosamente selecionada, a fim de preservar alguns benefícios dos experimentos fatoriais completos.
 - Uma das propriedades desejáveis de um fatorial 2^k é a ortogonalidade.

- Considere a escolha de metade de um delineamento 2^k .
 - Uma escolha descuidada de metade de $n=2^k$ execuções pode não manter a propriedade da ortogonalidade de um delineamento 2^k .
- Uma maneira de preservar esta propriedade, é escolher as execuções em que os níveis dos fatores codificados para um termo de interação (de preferência, a intereção de ordem mais alta) são constantes.

Table 6.2 Creating a Half-Fraction by Choosing the Runs in a Full Fraction with Constant Values for an Interaction													
Full Factorial					Half-Fraction								
run	X_A	X_B	X_C	X_D	$X_A X_B X_C X_D$)	run	X_A	X_B	X_C	X_D		
1	-	-	-	-	+		1	_	_	_	_		
2	+	_	_	_	_		10	+	-	_	+		
3	-	+	_	_	_		11	_	+	_	+		
4	+	+	-	-	+		4	+	+	_	_		
5	-	-	+	-	-		13	_	_	+	+		
6	+	-	+	-	+		6	+	_	+	-		
7	-	+	+	_	+		7	_	+	+	_		
8	+	+	+	_	_		16	+	+	+	+		
9	_	_	_	+	-								
10	+	_	_	+	+								
11	-	+	_	+	+								
12	+	+	-	+	-								
13	-	-	+	+	+								
14	+	_	+	+	_								
15	_	+	+	+	_								
16	+	+	+	+	+								

- Com 16 execuções em fatorial 2^4 completo, 15 efeitos podem ser estimados, além da média global:
 - 4 efeitos principais
 - 6 efeitos de interação dois-a-dois
 - 4 efeitos de interação de três fatores
 - 1 efeito de interação de ordem quatro
- Em um delineamento fatorial de fração 1/2 temos apenas 8 execuções.
 - Somente 7 efeitos podem ser estimados, além da média global.

- Escolhendo as execuções de um fatorial completo que tem valor constante para uma interação, nós automaticamente perdemos a habilidade de estimar este efeito de interação.
- Além disso, podemos perceber que

$$X_D = X_A imes X_B imes X_C$$
.

• Assim, o efeito que podemos estimar para X_D está completamente confundido pela interação $X_A X_B X_C$.

- Ainda há mais!
 - Cada efeito principal e interação no delineamento é confundido com uma outra interação.
- Esse é o preço que pagamos por executar metade do número total de experimentos.
- No entanto, em experimentos preliminares, em que um grande número de fatores é incluído, a fim de descobrir quais deles são realmente importantes, isso pode não ser um preço sério a ser pago.

- 1. Princípio de dispersão de efeito (Box e Meyer, 1986).
 - Em experimentos preliminares envolvendo um grande número de fatores, geralmente apenas uma pequena proporção dos fatores terá efeitos significativos.
- 2. Princípio hierárquico de ordenação (Wu e Hamada, 2000).
 - Assim como dois planetas se alinharão com a lua no céu noturno com mais frequência do que três planetas, os efeitos principais têm maior probabilidade de serem importantes do que as interações de dois fatores, e as interações de dois fatores têm maior probabilidade de serem importantes que as de três fatores interações e assim por diante.

 Portanto, se as frações de experimentos fatoriais puderem ser planejadas de forma que os efeitos principais sejam confundidos com interações de três fatores e ordem mais alta, a quantidade de informação perdida pelo fracionamento do número de execuções será pequena em comparação com o benefício de um número reduzido de execuções.

- A maneira como um fatorial fracionário de um 2^k é criado na prática é, na verdade, a ordem oposta do que foi mostrado anteriormente.
- Em vez de começar com um fatorial completo e eliminar as execuções para obter a fração desejada, inicie com um fatorial completo contendo o número desejado de execuções e adicione outros fatores ao delineamento.

- Para construir uma fração de metade de um delineamento 2^k , denotado por $\frac{1}{2}2^k$ ou 2^{k-1} , o procedimento é o seguinte:
- 1. Escreva o delineamento de base: um delineamento fatorial completo com k-1 fatores usando os níveis dos fatores codificados (-) e (+).
- 2. Adicione o fator k ao delineamento, tornando seus níveis de fator codificados iguais ao produto dos níveis dos outros fatores (ou seja, a interação de ordem mais alta).
- 3. Use estas colunas para definir o delineamento.
- Exercício: Refaça o delineamento 2^{4-1} , a partir dos fatores A, B e C.

- Uma lista completa de interações confundidas com cada efeito principal e interação em um fatorial de fração 1/2 é chamada de **padrão de confusão** ou **estrutura de aliases** do delineamento.
- Essa lista é fácil de construir com base na atribuição do fator k no item 2 da lista apresentada anteriormente.
- Por exemplo, no delineamento 2^{4-1} , se os níveis do quarto fator são iguais ao produto dos níveis dos três primeiros fatores do delineamento, escrevemos simbolicamente D=ABC.
 - Este é chamado de gerador do delineamento.

Multiplicando nos dois lados do gerador, obtemos:

$$D^2 = ABCD$$

ou

$$I = ABCD$$

em que I representa uma coluna de sinais positivos e é a identidade multiplicativa para produtos elemento a elemento de colunas de níveis codificados de factores.

- A equação, I=ABCD, é chamada de **relação de definição** para o delineamento fatorial fracionário e, multiplicando-se em ambos os lados dessa equação, a interação confundida com qualquer efeito principal ou interação pode ser determinada.
- Por exemplo, multiplicando pelo primeiro fator em ambos os lados da relação de definição, vemos:

$$A(I) = A(ABCD)$$

ou

$$A = BCD$$

• Isso significa que o efeito do primeiro fator A é confundido

com a interação de tres ratores BCD.

- Quando os dados são coletados, o efeito do fator A é estimado como a diferença na resposta média nos níveis alto e baixo do fator A.
 - No entanto, esse efeito realmente estima a soma dos efeitos do fator A e da interação de três fatores.
 - lacktriangle Portanto, escrevemos como A+BCD.

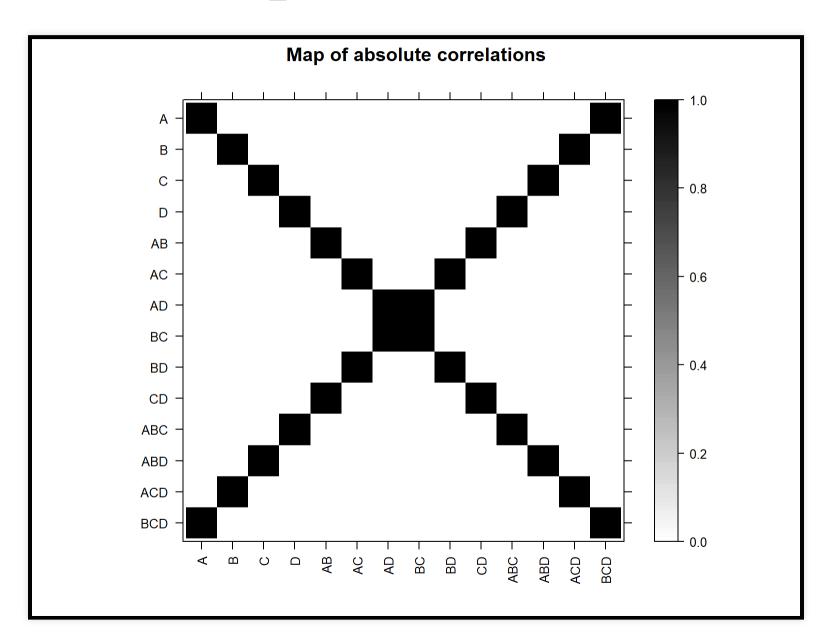
• O padrão de aliases completo para este design 2^{4-1} pode ser determinado multiplicando-se a relação de definição por cada efeito principal e interação resultando em:

$$I + ABCD$$
 $A + BCD$
 $B + ACD$
 $C + ABD$
 $D + ABC$

$$egin{array}{l} AB+CD \ AC+BD \ AD+BC \end{array}$$

• Existem apenas oito resultados únicos neste padrão de aliases, e eles representam os oito efeitos (I sendo a média geral) que podem ser estimados a partir do delineamento fatorial fracionário de 8 execuções.

Frações $\frac{1}{2}$ de delineamentos 2^k



• Criando um delineamento fatorial 2^{k-1} no R com o pacote FrF2 (Fractional Factorial Designs with 2-Level Factors).

```
## A B C D
## 1 -1 -1 -1 -1
## 2 1 -1 -1 1
## 3 -1 1 -1 1
## 4 1 1 -1 -1
## 5 -1 -1 1 1
## 6 1 -1 1 -1
## 7 -1 1 1 -1
## 8 1 1 1 1
## class=design, type= FrF2.generators
```

```
## A B C D
## 1 -1 1 -1 1
## 2 1 1 -1 -1
## 3 -1 -1 -1 -1
## 4 1 -1 1 -1
## 5 1 -1 -1 1
## 6 -1 -1 1 1
## 7 -1 1 1 -1
## 8 1 1 1 1
## class=design, type= FrF2.generators
```

```
design.info(design) # Função do pacote DoE.base
```

```
## $type
## [1] "FrF2.generators"
##
## $nruns
## [1] 8
##
## $nfactors
## [1] 4
##
## $factor.names
## $factor.names$A
## [1] -1 1
##
## $factor.names$B
## [1] -1 1
##
## $factor.names$C
## [1] -1 1
```

```
y \leftarrow runif(8, 0, 1)
aliases( lm( y \sim (.)^3, data = design))
```

```
##
## A = B:C:D
## B = A:C:D
## C = A:B:D
## D = A:B:C
## A:B = C:D
## A:C = B:D
## A:D = B:C
```

Um exemplo: sopa em pó

- Para ilustrar a análise de um delineamento 2^{k-1} , considere um exemplo da mistura de sopa.
- Um fabricante de misturas de sopa seca embalada estava experimentando uma variabilidade excessiva nos pesos da embalagem de uma sopa chamada "intermix".

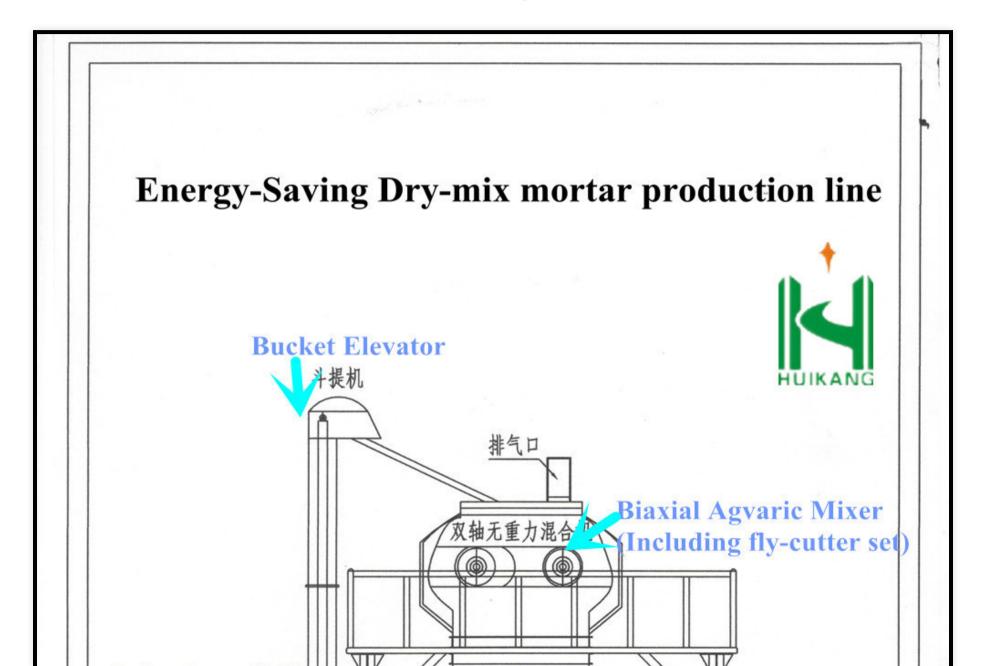


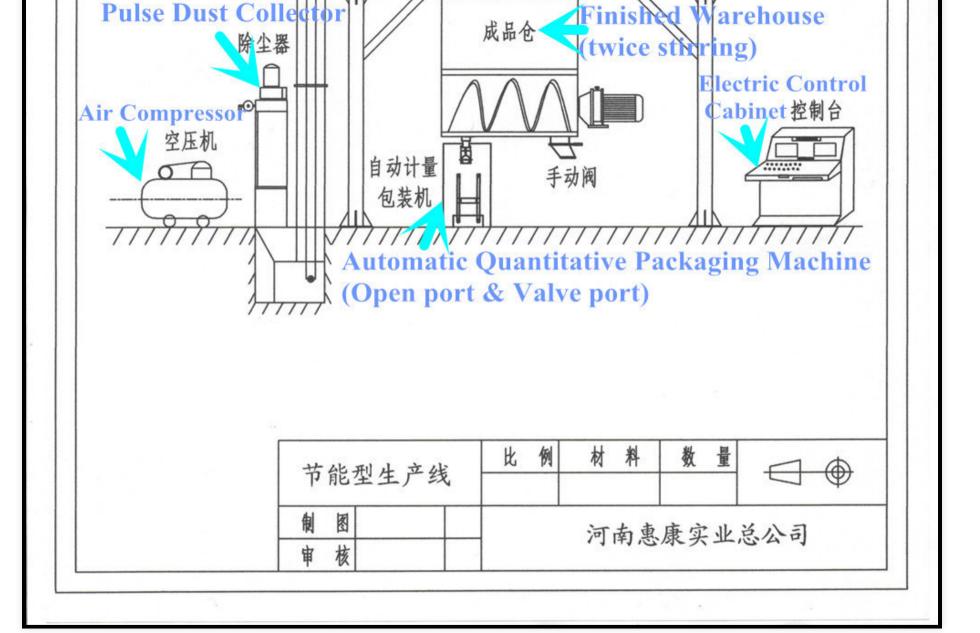


- A intermix é uma mistura de ingredientes saborosos, como óleo vegetal, sal e assim por diante.
 - Muito mistura em um pacote de sopa dá-lhe um sabor muito forte, e pouca mistura dá um sabor muito fraco.
 - Nesse exemplo, a maioria da variabilidade na sopa "intermix" foi encontrada dentro de um lote e não entre lotes.

- Os pesquisadores responsáveis pelo projeto fizeram uma lista de fatores que eles achavam que poderiam influenciar a variabilidade dentro de um lote.
- Estas eram opções que poderiam ser mudadas no misturador onde um lote foi misturado e o intermix foi adicionado através de portas.







- A lista consistiu em:
 - 1. o número de portas onde o intermix foi adicionado
 - 2. a temperatura do misturador (isso poderia ser controlado adicionando água de resfriamento ao revestimento ao redor do misturador)
 - 3. o tempo de mistura
 - 4. peso do lote
 - 5. o tempo de atraso entre a mistura e a embalagem.

 A resposta ou variabilidade em todos os pesos da mistura foi obtida tomando-se cinco amostras da mistura de sopa a cada 15 minutos enquanto o lote estava sendo empacotado.

Table 6.3	Factors and Levels for	r Soup Mix 2 ⁵⁻¹ Ex	xperiment
Factor Label	Name	Low Level	High Level
A	Number of Ports	1	3
В	Temperature	Cooling Water	Ambient
C	Mixing Time	60 sec.	80 sec.
D	Batch Weight	1500 lb	2000 lb
E	Delay Days	7	1

- O tamanho normal de lote era 2000 lb e o tempo de mistura normal era de 60 segundos.
 - Como esse experimento deveria ser executado em uma instalação de produção, as equipes de pesquisa e produção tinham que concordar com o plano.
 - O plano acordado foi o 2^{5-1} .

Table 6.4 2⁵⁻¹ Experiment to Determine Which Factors Are Associated with Fill Variation in Random Order

	(A)	(B)	(C)	(D)	(E)	
Random	Number		Mixing	Batch		
Run	of		Time	Weight	Delay	Response
Order	Ports	Temperature	(sec)	(lb)	(days)	$\hat{\sigma}_p$
12	1	Cool Water	60	1500	1	1.13
13	3	Cool Water	60	1500	7	1.25
5	1	Ambient	60	1500	7	0.97
3	3	Ambient	60	1500	1	1.70
6	1	Cool Water	80	1500	7	1.47
4	3	Cool Water	80	1500	1	1.28
16	1	Ambient	80	1500	1	1.18
14	3	Ambient	80	1500	7	0.98
1	1	Cool Water	60	2000	7	0.78
15	3	Cool Water	60	2000	1	1.36
7	1	Ambient	60	2000	1	1.85
10	3	Ambient	60	2000	7	0.62
11	1	Cool Water	80	2000	1	1.09
2	3	Cool Water	80	2000	7	1.10
9	1	Ambient	80	2000	7	0.76
8	3	Ambient	80	2000	1	2.10

- Todos os níveis de fator poderiam ser alterados entre lotes com relativamente pouco esforço e a aleatorização não era um problema.
- A redução solicitada no peso do lote e o aumento no tempo de mistura para alguns lotes na lista planejada de experimentos não interfeririam seriamente no cronograma de produção se a lista fosse curta.
 - Mas diminuiria a produção se um delineamento de 32 execuções fosse usado.
 - Por essa razão, o fatorial fracionário de 16 execuções foi acordado.

- A unidade experimental para este experimento foi o lote de mistura de sopa seca colocada no misturador.
- A resposta foi uma estimativa do desvio padrão de todo o peso dentro de um lote, calculado a partir de um estudo de amostragem feito durante a embalagem de cada lote.

##		Ports	Temp	MixTime	BatchWt	delay
##	1	1	Cool	60	1500	1
##	2	3	Cool	60	1500	7
##	3	1	Ambient	60	1500	7
##	4	3	Ambient	60	1500	1
##	5	1	Cool	80	1500	7
##	6	3	Cool	80	1500	1
##	7	1	Ambient	80	1500	1
##	8	3	Ambient	80	1500	7
##	9	1	Cool	60	2000	7
##	10	3	Cool	60	2000	1
##	11	1	Ambient	60	2000	1
##	12	3	Ambient	60	2000	7
##	13	1	Cool	80	2000	1
##	14	3	Cool	80	2000	7

## 15	1 Ambient	80	2000	7
## 16	3 Ambient	80	2000	1

```
##
     Ports Temp MixTime BatchWt delay y
## 1
                           1500
          Cool
                      60
                                   1 1.13
## 2
             Cool
                     60
                          1500
                                7 1.25
## 3
                                  7 0.97
                          1500
    1 Ambient
                     60
## 4
    3 Ambient
                                1 1.70
                     60
                          1500
## 5
                          1500
             Cool
                     80
                               7 1.47
## 6
                          1500
        3
             Cool
                     80
                               1 1.28
## 7 1 Ambient
                     80
                          1500
                                  1 1.18
## 8
        3 Ambient
                          1500
                                7 0.98
                     80
## 9
             Cool
                      60
                           2000
                                7 0.78
                           2000
## 10
        3 Cool
                      60
                                   1 1.36
## 11
     1 Ambient
                           2000
                      60
                                  1 1.85
     3 Ambient
## 12
                      60
                           2000
                                  7 0.62
## 13
     1
                           2000
             Cool
                      80
                                1 1.09
## 14
                           2000
        3
             Cool
                      80
                               7 1.10
## 15 1 Ambient
                     80
                        2000
                                  7 0.76
     3 Ambient
## 16
                     80
                           2000
                                  1 2.10
## class=design, type= FrF2 generators
```

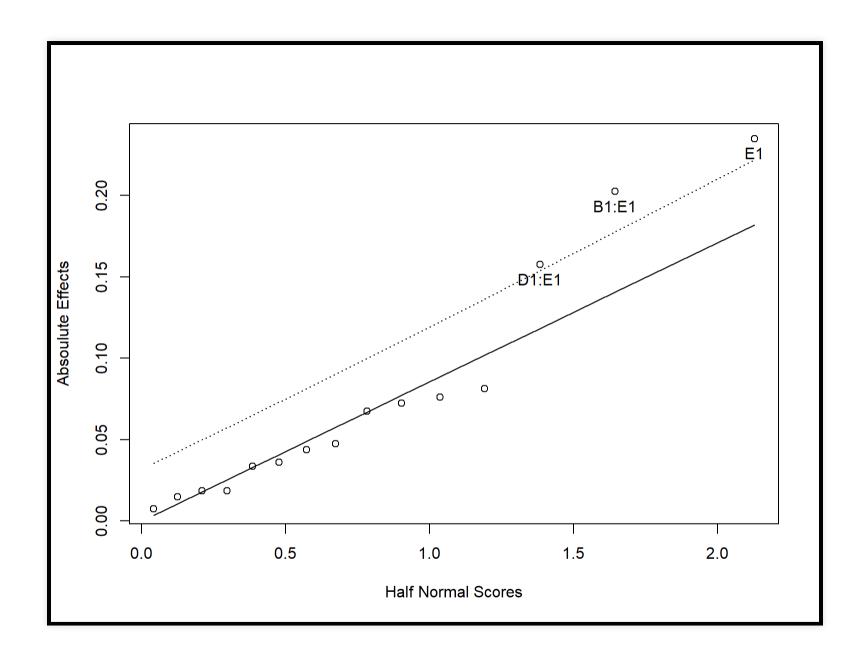
```
aliases(lm(y \sim (.)^4, data = soupc))
```

```
##
##
   A = B:C:D:E
##
   B = A:C:D:E
##
   C = A:B:D:E
##
   D = A:B:C:E
##
   E = A:B:C:D
##
   A:B = C:D:E
##
   A:C = B:D:E
##
   A:D = B:C:E
##
   A:E = B:C:D
##
   B:C = A:D:E
##
   B:D = A:C:E
##
   B:E = A:C:D
##
   C:D = A:B:E
##
   C:E = A:B:D
## D:E = A:B:C
```

```
mod1 <- lm( y ~ (.)^2, data = soup)
summary(mod1)</pre>
```

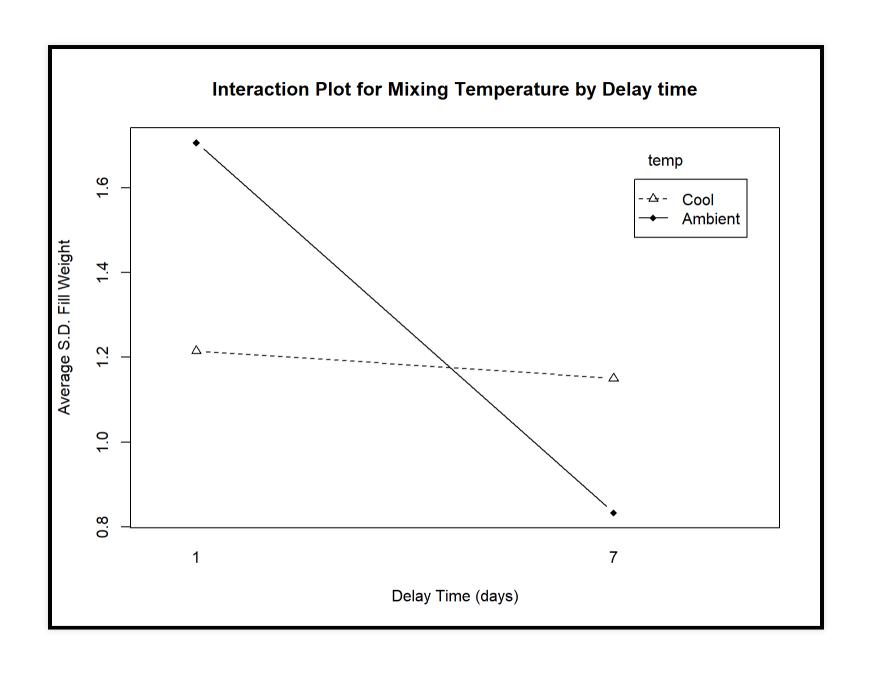
```
##
## Call:
## lm.default(formula = y \sim (.)^2, data = soup)
##
## Residuals:
## ALL 16 residuals are 0: no residual degrees of freedom!
##
## Coefficients:
##
                     Estimate Std. Error t value Pr(>|t|)
                      1.22625
## (Intercept)
                                       NΑ
                                               NA
                                                         NA
## Ports1
                      0.07250
                                       NΑ
                                               NA
                                                         NA
## Temp1
                      0.04375
                                       NΑ
                                               NA
                                                         NA
## MixTime1
                     0.01875
                                       NΑ
                                               NA
                                                         NA
## BatchWt1
                     -0.01875
                                       NA
                                               NΑ
                                                        NA
                     0.23500
## delay1
                                       NA
                                                         NA
                                               NA
## Ports1:Temp1
                   0.00750
                                       NA
                                                         NA
                                               NA
## Ports1:MixTime1 0.04750
                                       NA
                                               NA
                                                         NA
## Ports1:BatchWt1 0.01500
                                       NΑ
                                               NA
                                                         NA
```

```
modc <- lm(y ~ (.)^2, data = soupc)
library(daewr)
LGB(coef(modc)[-1], rpt = FALSE)</pre>
```

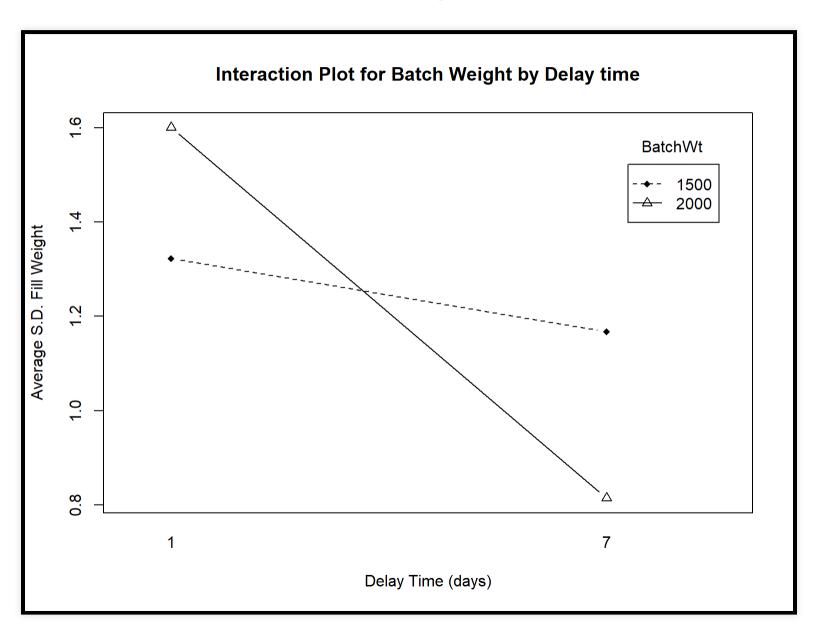


- Efeitos que parecem ser significativos:
 - E (Tempo de Atraso entre a mistura e o empacotamento)
 - BE (a interação entre Temperatura e Tempo de Retardo)
 - DE (a interação entre o Tempo de Atraso e o Tempo de Atraso)
- Se a o princípio hierárquico da ordenação puder ser assumida, esta é a interpretação correta, e as interações de três e quatro fatores podem ser consideradas insignificantes.

- Devido à atribuição não convencional de 7 ao nível baixo do fator E e 1 ao nível alto, ele nos diz que, em média, o tempo de atraso entre a mistura e a embalagem diminui a variabilidade dentro de lote.
- Como as interações BE e DE também parecem significativas, o efeito principal médio do fator E não conta toda a história.

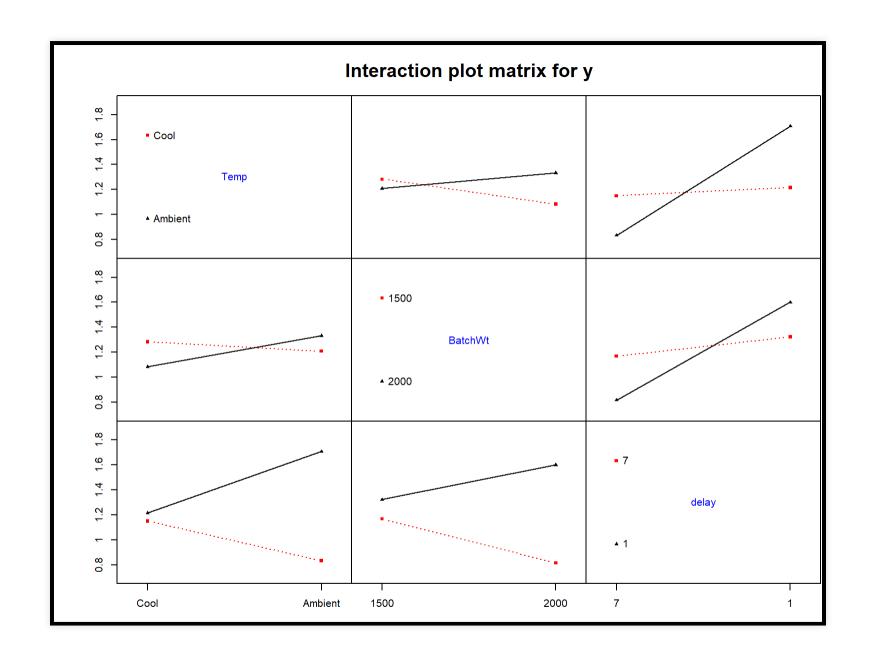


- Pode ser visto que o Tempo de Atraso tem puco efeito na variabilidade dentro do lote quando o misturador é arrefecido com água de resfriamento durante a mistura.
- No entanto, a variabilidade dentro do lote diminui substancialmente aumentando o tempo de atraso entre a mistura e a embalagem quando a mistura foi feita à temperatura ambiente.



 Pode-se observar que aumentar o tempo de atraso entre a mistura e o acondicionamento tem pouco efeito na variabilidade dentro de um lote para lotes pequenos (1500 lb), enquanto aumentar o tempo de atraso entre a mistura e a embalagem diminui a variabilidade dentro do lote para lotes maiores (2000 lb).

```
IAPlot(soup, sel=c(2,4,5), abbrev=7)
```



- Com base nos resultados deste estudo, a variação mínima de mistura dentro de um lote pode ser obtida usando:
 - o tamanho de lote maior (2000 lb),
 - a temperatura ambiente na mistura
 - tempo de 7 dias entre a mistura e a embalagem.
- A equipe de produção implementou as condições ideais.

Conclusões

- É interessante pensar que nenhuma das interações teria sido detectada se tivessem sido realizados experimentos com um fator de cada vez, ao invés do plano fatorial fracionário.
- As conclusões de experimentos do tipo com um fator de cada vez não podem ser reproduzíveis porque não se perceberia que os efeitos (ou ausência de efeitos) vistos poderiam depender de outros fatores que são considerados insignificantes.

Delineamentos fatoriais fracionários 2^{k-2} , 2^{k-p}

Delineamentos fatoriais fracionários 2^{k-p}

- ullet Em delineamentos 2^{k-1} , apenas metade das execuções experimentais são feitas.
 - Cada efeito que pode ser estimado é confundido com uma interação.
- Em um delineamento de fração $\frac{1}{4}=2^{-2},2^{k-2}$, apenas um quarto das execuções experimentais do delineamento completo de 2^k é executado.
 - Cada efeito que pode ser estimado é confundido com três outras interações.
- Em um delineamento de fração $\frac{1}{8}=2^{-3}$, apenas um oitavo das execuções do fatorial completo é feito.
 - Cada efeito estimado será confundido com outras sete interações, e assim por diante.

Delineamentos fatoriais fracionários 2^{k-p}

- Esses delineamentos podem soar confusos no início devido ao grande número de efeitos confundidos com cada efeito estimável.
 - No entanto, eles são usados com bastante frequência na prática, e seguindo o princípio de dispersão de efeito e o princípio de ordenação hierárquica, conclusões úteis podem geralmente ser alcançadas após apenas uma fração das execuções totais necessárias para um fatorial completo.

Construindo um delineamento fatorial $\mathbf{2}^{k-2}$

- 1. Comece com um delineamento de base em 2^{k-2} fatores.
- 2. Acrescente dois fatores adicionais, fazendo com que seus níveis de fator codificado sejam iguais a duas interações entre as primeiras k-2 colunas.

Construindo um delineamento fatorial $\mathbf{2}^{k-2}$

	1	U	X_D	X_E		
X_A	X_B	X_C	$X_A X_B$	$X_A X_C$	$X_B X_C$	$X_A X_B X_C$
_	-	_	+	+	+	_
+	-	_	_	_	+	+
_	+	_	_	+	_	+
+	+	_	+	_	_	_
-	-	+	+	-	_	+
+	-	+	-	+	_	_
-	+	+	-	_	+	_
+	+	+	+	+	+	+
		↓				
X_A	X_B	X_C	X_D	X_E		
-	-	_	+	+		
+	-	_	-	_		
-	+	_	-	+		
+	+	-	+	-		
_	-	+	+	-		
+	-	+	-	+		
_	+	+	-	-		
+	+	+	+	+		

Construindo um delineamento fatorial $\mathbf{2}^{k-2}$

- Existem dois geradores para o delineamento criado, $D = AB \, \mathrm{e} \, E = AC.$
- Destes pode ser visto que I=ABD e I=ACE.
 - Além disso, desde $I^2=I$, uma terceira igualdade, chamada de **interação generalizada**, é I=ABD(ACE) ou I=BCDE.
- Combinando as três equações obtidas dos dois geradores e a interação generalizada resulta na relação de definição para o delineamento I=ABD=ACE=BCDE.

Construindo um delineamento fatorial 2^{k-2}

 O padrão de confusão, ou estrutura de aliases, para o delineamento é encontrado multiplicando-se através da relação de definição por cada efeito que pode ser estimado:

$$A + BD + CE + ABCDE$$
 $B + AD + ABCE + CDE$
 $C + ABCD + AE + BDE$
 $D + AB + ACDE + BCD$
 $E + ABDE + AC + BCD$
 $BC + ACD + ABE + DE$

$$DD + ADD + ADO + OD$$

Construindo delineamentos fatoriais 2^{k-p}

- A construção de frações de um oitavo e superiores é semelhante.
- Para construir uma fração de um oitavo de um delineamento 2^6 ou 2^{6-3} , comece com um projeto de base em 6-3=3 fatores, depois acrescente três fatores adicionais, confundindo-os com interações.

Construindo delineamentos fatoriais $\mathbf{2}^{k-p}$

• Por exemplo, se escolhermos os geradores D=AB, E=AC e F=BC, o código R usando FrF2 pode ser usado para gerar o delineamento.

```
## A B C D E F

## 1 -1 1 -1 -1 1 -1

## 2 1 1 -1 1 -1 -1

## 3 -1 -1 1 1 -1 -1

## 4 1 -1 1 -1 1 -1

## 5 1 -1 -1 -1 1

## 6 -1 1 1 -1 -1

## 7 -1 -1 -1 1 1

## 8 1 1 1 1 1

## class=design, type= FrF2.generators
```

Construindo delineamentos fatoriais 2^{k-p}

- Existem oito execuções nesse delineamento, e sete efeitos além da média geral podem ser estimados.
- Cada efeito será confundido com sete interações.
- Para descobrir quais interações são confundidas com cada efeito que pode ser estimado, primeiro encontre a relação de definição.
 - Dos geradores I = ABD = ACE = BCF.
 - As interações generalizadas de dois fatores são $ABD(ACE) = BCDE, ABD(BCF) = ACDF \ e$ ACE(BCF) = ABEF.
 - A interação generalizada de três fatores é ABD(ACE)(BCF) = DEF. Combinando as equações encontradas entre os geradores e as interações generalizadas, a relação de definição é

inital açots generalizadas, a relação de deninição e

$$I = ABD = ACE = BCF = BCDE = ACDF = ABE$$

Construindo delineamentos fatoriais 2^{k-p}

```
y <- runif(8, 0, 1)
aliases( lm( y~ (.)^6, data = frac))
```

```
##
## A = B:D = C:E = B:E:F = C:D:F = A:B:C:F = A:D:E:F = A:B:C:D:E
## B = C:F = A:E:F = C:D:E = A:B:C:E = B:D:E:F = A:B:C:D:F = A:D
## C = B:F = A:D:F = B:D:E = A:B:C:D = C:D:E:F = A:B:C:E:F = A:E
## D = E:F = A:C:F = B:C:E = A:C:D:E = B:C:D:F = A:B:D:E:F = A:B
## E = D:F = A:B:F = B:C:D = A:B:D:E = B:C:E:F = A:C:D:E:F = A:C
## F = B:C = D:E = A:B:E = A:C:D = A:B:D:F = A:C:E:F = B:C:D:E:F
## A:F = B:E = C:D = A:B:C = A:D:E = B:D:F = C:E:F = A:B:C:D:E:F
```

Critérios para escolha dos geradores de delineamentos $\mathbf{2}^{k-p}$

Critérios para escolha dos geradores de delineamentos 2^{k-p}

- Há mais de uma alternativa ao selecionar os geradores para um delineamento 2^{k-p} .
- Por exemplo, para criar uma fração 1/4 de um delineamento 2^6 , os geradores E=ABC e F=ABD poderiam ser usados, ou os geradores E=AB e F=ACD poderiam ser usados.

Critérios para escolha dos geradores de delineamentos 2^{k-p}

 A primeira seleção resulta na relação de definição e na estrutura de aliases (somente para os principais efeitos) mostrados abaixo:

$$I = ABCE = ABDF = CDEF$$
 $A + BCE + BDF + ACDEF$
 $B + ACE + ADF + BCDEF$
 $C + ABE + ABCDF + DEF$
 $D + ABCDE + ABF + CEF$

$$E+ABC+ABDEF+CDF \ F+ABCEF+ABD+CDE$$

Critérios para escolha dos geradores de delineamentos $\mathbf{2}^{k-p}$

 O segundo conjunto de geradores resulta na relação de definição e estrutura de alias (apenas para os principais efeitos) abaixo:

$$I = ABE = ACDF = BCDEF$$
 $A + BE + CDF + ABCDEF$
 $B + AE + ABCDF + CDEF$
 $C + ABCE + ADF + BDEF$
 $D + ABDE + ABF + CEF$

$$E + ABC + ACF + ABEF$$
 $F + ABEF + ACD + BCDE$

Critérios para escolha dos geradores de delineamentos $\mathbf{2}^{k-p}$

- Ambos os geradores resultam em fatoriais fracionários de 16 execuções,
 - Qual dos dois delineamentos devemos escolher?

Critérios para escolha dos geradores de delineamentos 2^{k-p}

- O primeiro conjunto de geradores pode ser preferível, pois cada efeito principal é confundido com interações de três fatores e interações de ordem mais alta.
- O segundo conjunto de geradores resulta em um delineamento onde os efeitos principais A e B são confundidos com uma **interação de dois fatores** cada.
- O primeiro delineamento tem uma chance menor de confusão, pois o **princípio hierárquico de ordenação** nos diz que as **interações de três fatores** são **menos prováveis** de serem importantes do que as interações de dois fatores.

Critérios

- Três critérios gerais foram propostos para guiar a escolha entre os vários conjuntos possíveis de geradores para qualquer possível delineamento de 2^{k-p} .
 - 1. Critério da resolução.
 - 2. Critério da aberração.
 - 3. Critério dos efeitos claros.

Critério da resolução

- Box e Hunter (1961) propuseram o critério da resolução.
- A resolução do delineamento é definida como o comprimento da palavra mais curta na relação de definição.

$$I = ABCE = ABDF = CDEF$$

- A palavra mais curta tem comprimento 4.
 - Assim, é um delineamento de resolução IV.

$$I = ABE = ACDF = BCDEF$$

- A palavra mais curta para a segunda relação de definição mostrada acima tem comprimento 3.
 - É um delineamento de resolução III.
- Critério: se o número de execuções em dois delineamentos for o mesmo, o delineamento com a resolução mais alta

sei a pi ei ei iuo.

Critério da resolução

- Em um delineamento de resolução R, nenhum efeito envolvendo i fatores é confundido com efeitos de ordem menor que R-i.
- Em delineamentos com resolução V
 - os efeitos principais são confundidos com interações de quatro fatores e interações de ordem mais alta
 - as interações de dois fatores são confundidas com interações de três fatores e interações de ordem mais alta
- Portanto, se todas as interações de três fatores e ordem mais alta puderem ser consideradas insignificantes, todos os efeitos principais e interações de dois fatores podem ser estimados a partir de um delineamento de resolução V.

Critério da resolução

- Nos delineamentos da resolução IV
 - os efeitos principais são confundidos com as interações de três fatores e ordem superior.
- Em um delineamento de resolução III
 - os efeitos principais são confundidos com interações de dois fatores.
- Os delineamentos de resolução III são normalmente usados apenas em experimentos de triagem, onde o objetivo é descobrir quais fatores são importantes o suficiente para serem mais estudados em experimentos de acompanhamento.

Resolução e propriedade da projeção

- A propriedade da projeção de um fatorial fracionário é outro atributo que pode ser determinado a partir de sua resolução.
- ullet Em um delineamento de resolução R, qualquer subconjunto de fatores k=R-1 formará um delineamento completo de 2^k (com possível replicação de algumas execuções).
- ullet Portanto, se um experimento for iniciado com um planejamento fatorial fracionário de resolução R e apenas R-1 dos fatores parecerem signicativos, os dados poderão ser reanalisados, incluindo apenas os fatores significativos R-1.
- Como o delineamento desses fatores é um fatorial completo, as interações de todas as ordens possíveis entre os fatores R-1 podem ser examinadas.

Critério da aberração

- Quando dois ou mais delineamentos, criados com diferentes conjuntos de geradores, têm o mesmo número de execuções e a mesma resolução, Fries e Hunter (1980) propuseram outro critério para decidir qual projeto é preferível.
- Eles chamaram esse critério da aberração mínima.
- Se o número de palavras de comprimento r na relação de definição de um deselineamento é definido como A_r , então um delineamento d_1 é dito ter menos aberração do que um delineamento d_2 se r é o menor inteiro tal que $A_r(d_1) \neq A_r(d_2)$ e $A_r(d_1) < A_r(d_2)$.

Critério da aberração

- Exemplo: se d_1 é um delineamento de resolução IV 2^{7-2} criado com os geradores F=ABCD e G=ABCE, ele tem menos aberração do que o delineamento d_2 , criado com os geradores F=ABC e G=ADE.
 - 1. Relação de definição para d_1 (I=ABCDF=ABCEG=DEFG) tem apenas uma palavra de comprimento 4 ($A_4(d_1)=1$)
 - 2. Relação de definição para d_2 (I=ABCF=ADEG=BCDEFG) tem duas palavras de comprimento 4 ($A_4(d_2)=2$).

Critério da aberração

- Para quaisquer k e p existe sempre um delineamento de aberração mínima de 2^{k-p} que tenha menos aberração do que qualquer outro delineamento de 2^{k-p} .
- Critério: Para dois delineamentos com a mesma resolução, o delineamento de aberração mínima terá menos confusão dos efeitos principais com interações de ordem mais baixa e é geralmente preferido.

Critério dos efeitos claros

- Um critério final que é útil ao selecionar os geradores para um delineamento 2^{k-p} é o número de **efeitos claros**.
- Chen et al. (1993) definem um efeito como claro se nenhum dos seus aliases são efeitos principais ou interações de dois fatores.
- Em alguns casos, um delineamento que não seja o delineamento de aberração mínima pode ter efeitos mais claros do que o delineamento de aberração mínima.
 - Exemplo: Wu e Hamada (2000) explicam que o delineamento 2^{6-2} com relação de definição I=ABCE=ABDF=CDEF tem todos os seis efeitos principais claros, enquanto o delineamento 2^{6-2} com relação de definição

I=ABE=ACDF=BCDEF tem três efeitos

- principais (C, D, eF) claros junto com seis interações de dois fatores BC, BD, BF, CE, DE e EF.
- Nos casos em que se acredita que algumas interações de dois fatores sejam importantes a priori, o segundo delineamento pode ser preferido em relação ao primeiro.

Critério da aberração mínima no R

- A função FrF2 no pacote FrF2 do R pode criar delineamentos de aberração mínima.
- Se o usuário não especificar geradores ao chamar a função FrF2, como no exemplo anterior, a função FrF2 selecionará automaticamente o conjunto de geradores que resultarão em um delineamento de aberração mínima.

Critério da aberração mínima no R

```
library(FrF2)
des1 <- FrF2( 16, 8 )
y <- runif( 16, 0, 1 )
library(DoE.base)
generators(des1)</pre>
```

```
## $generators
## [1] "E=ABC" "F=ABD" "G=ACD" "H=BCD"
```

```
aliases ( lm(y \sim (.)^3, data = des1) )
```

```
##
##
   A = B:C:E = B:D:F = B:G:H = C:D:G = C:F:H = D:E:H = E:F:G
##
   B = A:C:E = A:D:F = A:G:H = C:D:H = C:F:G = D:E:G = E:F:H
##
   C = A:B:E = A:D:G = A:F:H = B:D:H = B:F:G = D:E:F = E:G:H
##
   D = A:B:F = A:C:G = A:E:H = B:C:H = B:E:G = C:E:F = F:G:H
##
   E = A:B:C = A:D:H = A:F:G = B:D:G = B:F:H = C:D:F = C:G:H
##
   F = A:B:D = A:C:H = A:E:G = B:C:G = B:E:H = C:D:E = D:G:H
##
   G = A:B:H = A:C:D = A:E:F = B:C:F = B:D:E = C:E:H = D:F:H
   H = A:B:G = A:C:F = A:D:E = B:C:D = B:E:F = C:E:G = D:F:G
##
##
   A:B = C:E = D:F = G:H
##
   A:C = B:E = D:G = F:H
   A:D = B:F = C:G = E:H
```

```
## A:E = B:C = D:H = F:G

## A:F = B:D = C:H = E:G

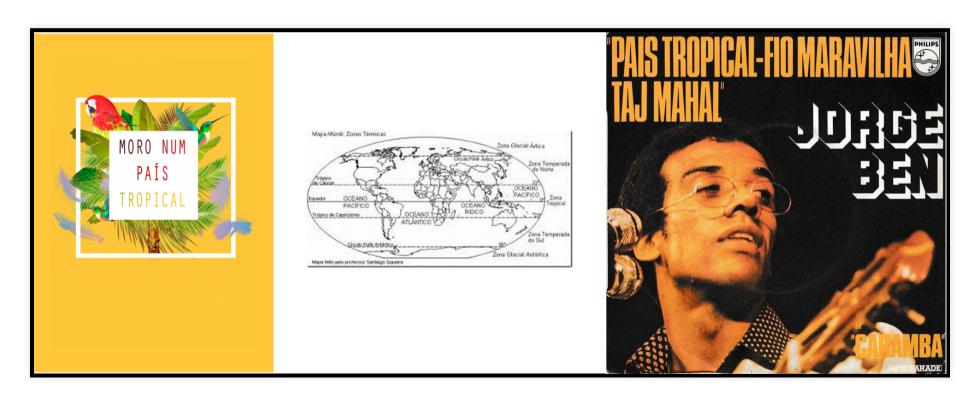
## A:G = B:H = C:D = E:F

## A:H = B:G = C:F = D:E
```

Critério dos efeitos claros no R

- A função FrF2 também pode criar delineamentos com o máximo número de efeitos claros.
- Por exemplo, a chamada FrF2(32,9) produz o delineamento de aberração mínima 2^{9-4} que tem nove efeitos principais claros e oito interações claras de dois fatores.
- No entanto, a chamada FrF2(32,9, MaxC2 = TRUE) produz um delineamento 2^{9-4} que tem todos os nove efeitos principais e 15 interações claras de dois fatores.

- Considere o delineamento e análise de um fatorial fracionário de 2^{8-4} .
- AlmeidaeSilva et al. (1998) realizaram um experimento para encontrar as condições ótimas para a cultura de *Paecilomyces variotii* (um fungo comumente encontrado no ar e em solos de países tropicais) em hidrolisado hemicelulósico de eucalipto com vistas a produzir proteína microbiana.





- Apenas 51,6% da massa seca total de madeira de eucalipto é utilizada pela indústria brasileira, enquanto o restante (galhos, folhas, pequenas árvores, etc.) é deixado nos campos.
- A fração de hemicelulose deste resíduo pode ser facilmente removida por tratamento com ácido, e o hidrolisado resultante é rico em açúcares fermentáveis.
- O fungo *P. variotii* foi selecionado entre 21 espécies de leveduras e fungos lamentosos por seu desempenho no hidrolisado de hemicelulose de eucalipto.
- A biomassa proteica produzida por este fungo durante 72 horas de fermentação tem um perfil de aminoácido que é igual ou superior aos produtos convencionais utilizados para a alimentação animal.

• O objetivo dos experimentos foi estudar a influência de inibidores, nutrientes e tempo de fermentação no crescimento de biomassa produzido por P. variotii.

	Table 6.5 Factors and Levels for Biomass Experiment Levels		
Label	Factors	- Le	reis +
A	Inhibitors (Furfural and Acetic Acid)	$1.25 \mathrm{g/L}$	7.8g/L
В	Rice Bran	$10.0 \mathrm{g/L}$	$30.0 \mathrm{g/L}$
\mathbf{C}	Urea	$0.0 \mathrm{g/L}$	$2.0 \mathrm{g/L}$
D	Magnesium Sulfate	$0.0 \mathrm{g/L}$	$1.5 \mathrm{g/L}$
\mathbf{E}	Ammonium Sulfate	$0.0 \mathrm{g/L}$	$2.0 \mathrm{g/L}$
\mathbf{F}	Potassium Nitrate	$0.0 \mathrm{g/L}$	$2.0 \mathrm{g/L}$
G	Sodium Phosphate	$0.0 \mathrm{g/L}$	$2.0 \mathrm{g/L}$
Н	Fermentation Time	72 hrs	96 hrs

- Um planejamento fatorial fracionário 2^{8-4} foi utilizado com os geradores E=BCD, F=ACD, G=ABC e H=ABD.
- Este é o delineamento de resolução IV de aberração mínima.
 - Os efeitos claros neste delineamento são os oito efeitos principais.
 - Há também sete sequências confundidas de interações de dois fatores que podem ser estimadas.

$$CG + DH + AB + EF$$
 $AC + BG + DF + EH$
 $CF + AD + EG + BH$
 $CH + DG + AE + BF$
 $CD + GH + AF + BE$
 $BC + AG + DE + FH$
 $CE + FG + AH + BD$

Exercício

1. Confira que este é um delineamento de resolução IV. ## Exemplo (*P. variotii*): R

```
library(FrF2)
culture <- FrF2( 16, generators = c("BCD", "ACD", "ABC", "ABD"), randomi:
y1 <- c(5.75, 6.7, 11.12, 10.67, 4.92, 5.35, 2.81, 10.83, 6.08, 7.27, 9.0
culture <- add.response(culture, y1)
culture</pre>
```

```
## A B C D E F G H y1
## 1 -1 -1 -1 -1 -1 -1 -1 -1 5.75
## 2 1 -1 -1 -1 1 1 1 1 6.70
## 3 -1 1 -1 -1 1 1 1 1 1.12
## 4 1 1 -1 -1 1 1 1 1 1 1.12
## 5 -1 -1 1 1 1 1 1 1 4.92
## 6 1 -1 1 -1 1 1 -1 1 5.35
## 7 -1 1 1 1 -1 1 1 1 2.81
## 8 1 1 1 -1 -1 1 1 1 1 0.83
## 9 -1 -1 -1 1 1 1 1 1 1 6.08
## 10 1 -1 -1 1 1 1 1 1 7.27
## 11 -1 1 -1 1 1 1 1 1 9.68
## 12 1 1 -1 1 -1 1 1 4.20
```

```
## 13 -1 -1 1 1 -1 -1 1 3.90

## 14 1 -1 1 1 -1 1 -1 -1 3.78

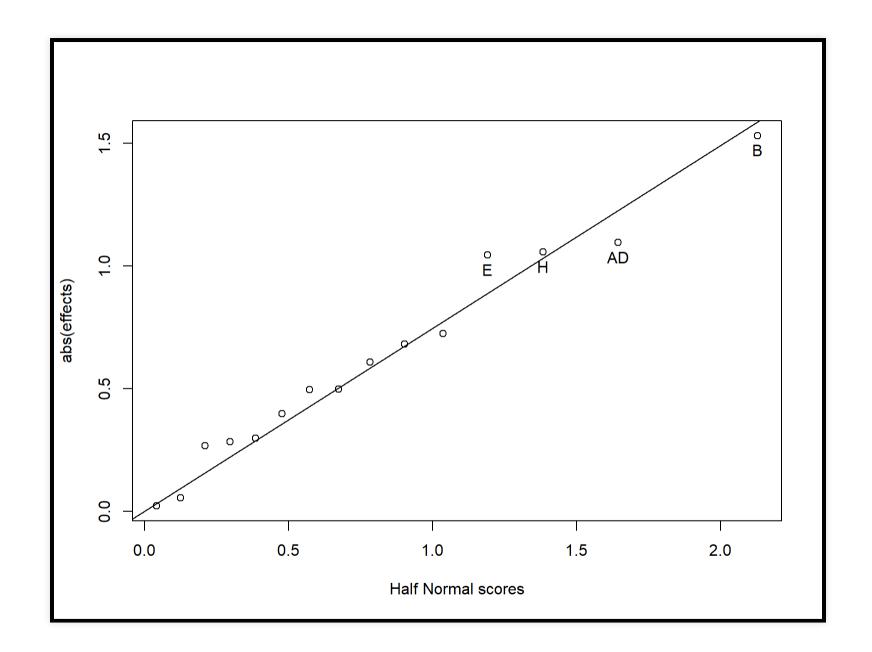
## 15 -1 1 1 1 1 -1 -1 -1 11.57

## 16 1 1 1 1 1 1 1 7.39
```

```
modf <- lm( y1 ~ (.)^2, data = culture)
summary(modf)</pre>
```

```
##
## Call:
## lm.default(formula = v1 \sim (.)^2, data = culture)
##
## Residuals:
## ALL 16 residuals are 0: no residual degrees of freedom!
##
## Coefficients: (21 not defined because of singularities)
##
              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 7.00125
                                       NA
                               NA
                                                NA
## A1
              0.02250
                               NA
                                       NA
                                                NA
## B1
              1.53250
                               NA
                                      NA
                                                NA
## C1
              -0.68250
                                       NA
                                                NA
                               NA
## D1
            -0.26750
                               NA
                                       NA
                                                NA
## E1
             1.04500
                                       NA
                               NA
                                                NA
## F1
             -0.49750
                                                NA
                               NA
                                       NA
## G1
              0.72500
                                       NA
                                                NA
                               NA
## H1
       -1.05750
                               NΑ
                                       NΑ
                                                NΑ
```

```
library(daewr)
cfs <- coef(modf)[2:16]
names <- names(cfs)
halfnorm(cfs, names, alpha = .25, refline = TRUE)</pre>
```



- Os autores do artigo acharam que o experimento lhes dava evidências de que o fator A (o inibidor) tinha pouco efeito, e confirmaram isso citando outros relatórios publicados.
- Eles também sentiram que o experimento mostrou que os efeitos principais D (sulfato de magnésio) e F (nitrito de potássio) eram insignificantes.
- No entanto, devido ao confundimento de interações de dois fatores com o segundo maior efeito (em valor absoluto) e ao fato de que nada se destacou claramente no gráfico half-normal, não foi possível tirar conclusões definitivas.

- Eles decidiram executar outro experimento de acompanhamento 2^{5-1} de resolução V com os fatores B, C, E, G e H com relação de deinição I=BCEGH e os fatores A, D e F mantidos constantes no ponto médio dos níveis utilizados no primeiro experimento.
- Este foi um delineamento de 16 execuções, mas se fosse possível presumir com segurança que os principais efeitos A, D e F eram insignificantes no primeiro conjunto de dezesseis experimentos, então oito dos dezesseis testes para o experimento de acompanhamento proposto já estavam concluídos.
 - A seguir, as execuções 3, 4, 7, 8, 11, 12, 15 e 16 já foram realizadas no experimento anterior.

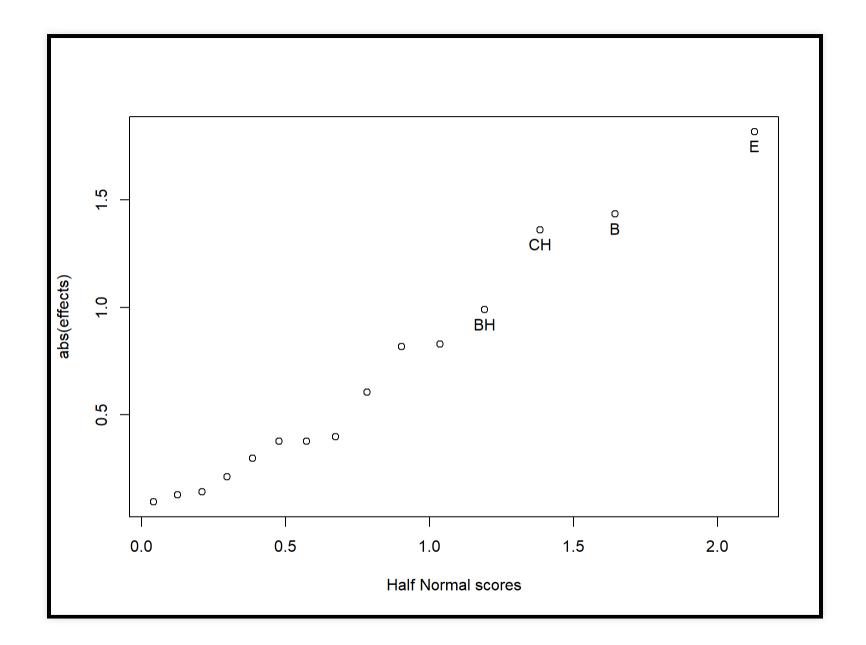
```
culture2 <- FrF2( 16, 5, factor.names = c("B", "C", "E", "G", "H"), rando
y <- c(3.37, 3.55, 3.78, 2.81, 5.53, 10.43, 5.35, 11.57, 2.93, 7.23, 3.9,
culture2 <- add.response( culture2, y )
culture2</pre>
```

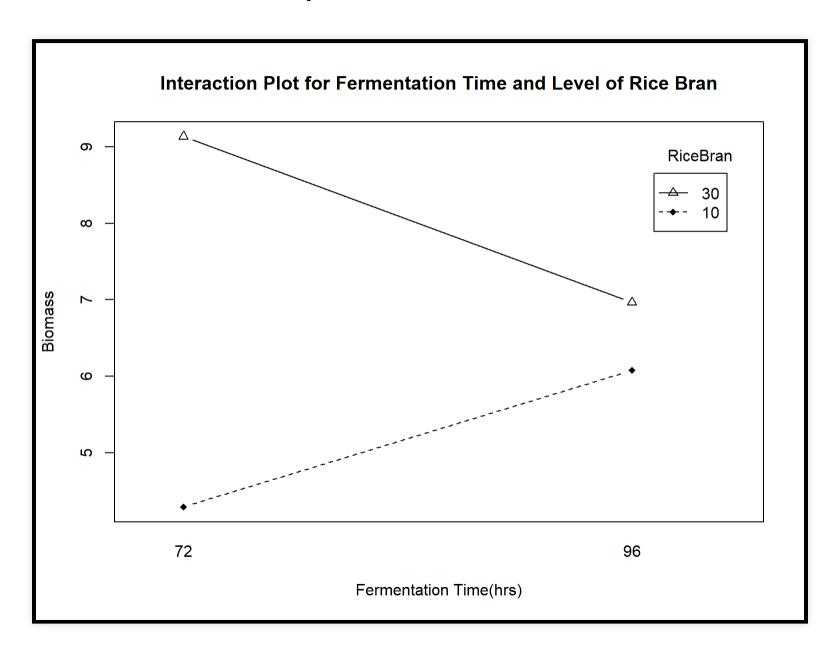
```
##
  B C E G H
## 1 -1 -1 -1 1 3.37
## 2 1 -1 -1 -1 3.55
## 3 -1 1 -1 -1 3.78
## 4 1 1 -1 -1 1 2.81
## 5 -1 -1 1 -1 -1 5.53
## 6 1 -1 1 -1 1 10.43
## 7 -1 1 1 -1 1 5.35
## 8 1 1 1 -1 -1 11.57
## 9 -1 -1 -1 1 -1 2.93
## 10 1 -1 -1 1 1 7.23
## 11 -1 1 -1 1 3.90
## 12 1 1 -1 1 -1 10.83
## 13 -1 -1 1 1 11.69
## 14 1 -1 1 1 -1 10.59
## 15 -1 1 1 1 -1 4.92
## 16 1 1
## class=design, type= FrF2
```

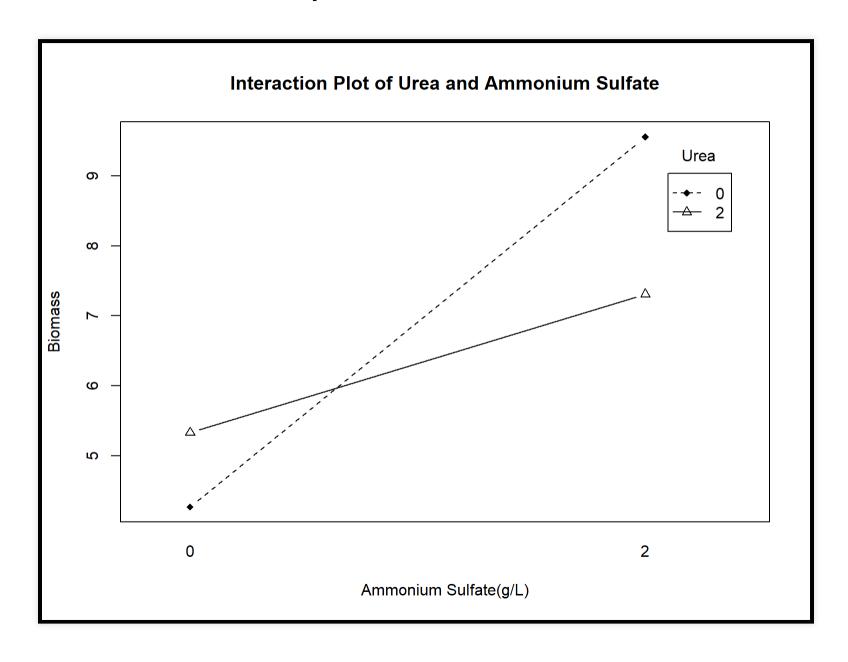
```
moda <- lm(y ~ (.)^2, data = culture2)
summary(moda)</pre>
```

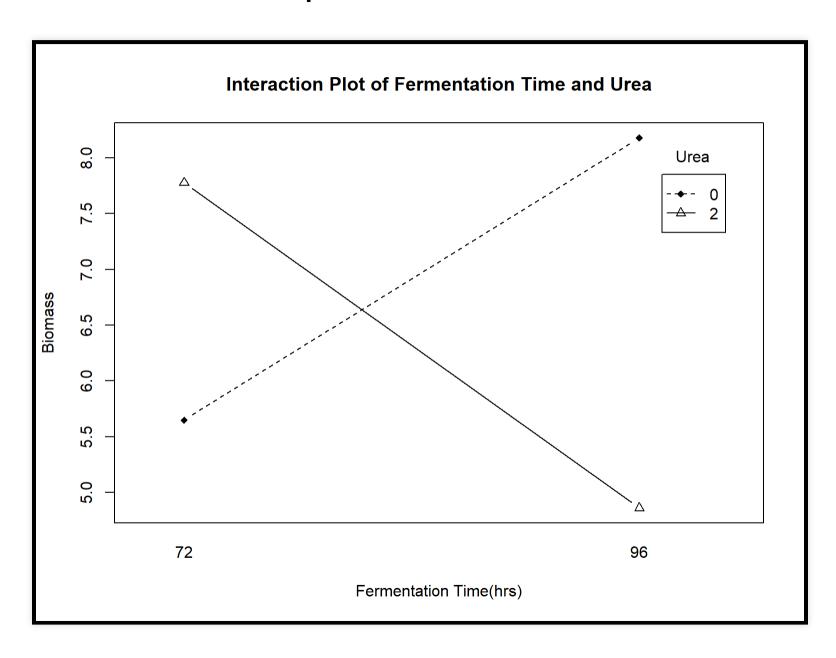
```
##
## Call:
## lm.default(formula = y \sim (.)^2, data = culture2)
##
## Residuals:
## ALL 16 residuals are 0: no residual degrees of freedom!
##
## Coefficients:
##
              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 6.61688
                              NA
                                     NA
                                              NA
## B1
           1.43312
                              NA
                                     NA
                                              NA
## C1
            -0.29812
                              NA
                                     NA
                                              NA
             1.81687
## E1
                              NA
                                     NA
                                              NA
## G1
            0.81813
                                     NA
                              NA
                                              NA
## H1 -0.09562
                                     NA
                              NA
                                              NA
## B1:C1
            0.39813
                                     NA
                                              NA
                              NA
## B1:E1
            0.12813
                                     NA
                                              NA
                              NA
## B1:G1 0.14188
                              NΑ
                                     NΑ
                                              NΑ
```

```
library(daewr)
cfs <- coef(moda)[2:16]
effects <- cfs
names <- names(cfs)
halfnorm(cfs, names, alpha = .2, refline = FALSE)</pre>
```





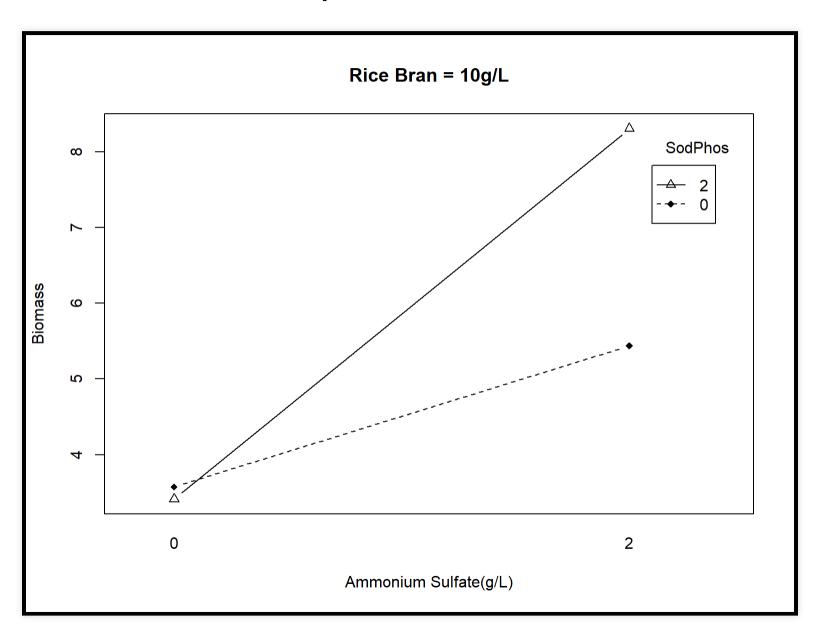


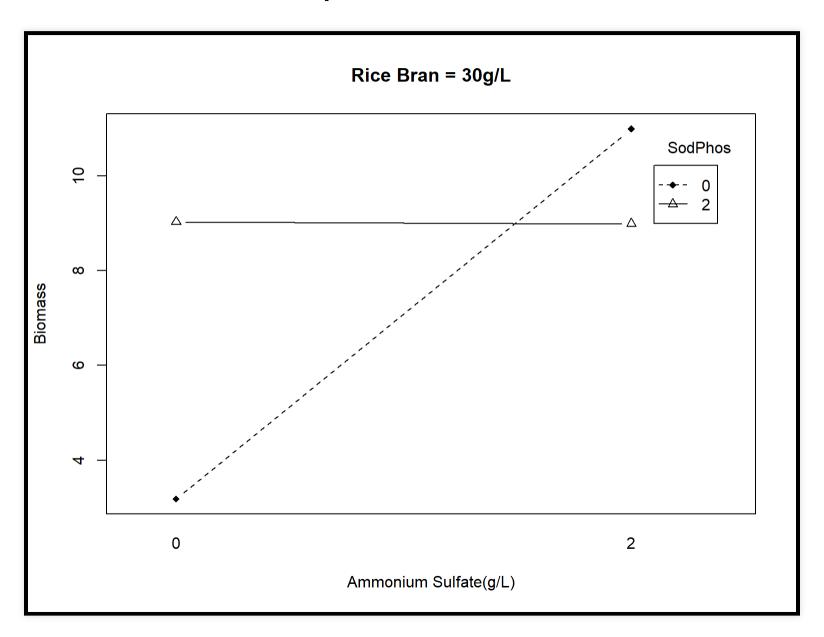


Princípio da hereditariedade do efeito

- Embora seja possível ter uma interação entre dois fatores que não tenham efeitos principais significativos (como o exemplo mostrado anteriormente), isso é raro.
- No estudo de Li et al. (2006) de 113 experimentos fatoriais publicados, isso aconteceu menos de 1% das vezes. -Normalmente, as interações ocorrem entre fatores em que pelo menos um dos dois efeitos principais é significativo.
- Isto foi descrito como o princípio da hereditariedade do efeito por Hamada e Wu (1992).

- Neste experimento, uma vez que a interação de dois fatores entre o tempo de fermentação e a uréia é confundida com a interação de três fatores entre farelo de arroz, sulfato de amônio e fosfato de sódio (CH=BEG), e todos os três últimos fatores têm efeitos principais significantes, é possível que o grande efeito rotulado como CH no gráfico normal seja, na verdade, a interação de três fatores.
- Nesse caso, o princípio da hereditariedade do efeito pode ofuscar o princípio hierárquico de ordenação.





- Pode-se observar que a adição de 2,0 g/L de sulfato de amônio ao meio de crescimento aumenta a biomassa produzida em geral.
- No entanto, este efeito é maior quando há 30 g/L de farelo de arroz e nenhum fosfato de sódio no meio de crescimento.
- O resultado ótimo parece ser com um tempo de fermentação de 72 horas, 30 g/L de farelo de arroz, 2,0 g/L de sulfato de amônio e 0,0 g/L de uréia no meio de crescimento.
 - Nesta combinação, a produção de biomassa foi estimada em 11,57 g/L.

Recapitulando

- Os pesquisadores começaram com oito fatores com dois níveis cada.
 - Seria necessário $2^8=256$ experimentos para completar um experimento fatorial completo.
- O princípio de dispersão de efeito e o princípio hierárquico de ordenação sugerem que provavelmente este experimento não é necessário.
 - Com oito fatores, é duvidoso que todos os efeitos principais e um grande número de interações seriam significativos.

Recapitulando

- Uma série de dois experimentos fatoriais fracionários, empregando um total de 24 execuções, revelou que quatro dos principais efeitos são significativos e no máximo três interações.
- Gráficos de interação ajudaram a fazer uma interpretação dos resultados e a identificação de níveis de fatores para resultados ótimos. = A interpretação das interações foi crítica para a identificação de resultados ótimos e nunca teria sido descoberta se um plano de um fator por vez tivesse sido utilizado.

Próximos capítulos

• Fatoriais fracionados aumentados.

