

Delineamentos em Blocos Completos Casualizados

Profa. Cibele Russo

(Referências: Montgomery (2012), Notas de aula de Roseli Leandro; Clarice Demétrio; Marinho Andrade)

Compare os dois croquis a seguir:

1. Delineamento inteiramente casualizado (DIC)
2. Delineamento em blocos completos casualizados (DBCC)

Quais as diferenças?

- Quais princípios básicos da experimentação estão sendo contemplados em cada um deles?

Princípios básicos da experimentação:

1. Repetição
2. Casualização
3. Controle local

Delineamento inteiramente casualizado

Croqui DIC

4 tratamentos com 4 repetições

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

Tratamentos A, B, C, D

Figure 1: Delineamento inteiramente casualizado

Delineamento em blocos completos casualizados

Croqui DBC

4 blocos (coluna) 4 tratamentos

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

Tratamentos A, B, C, D

Figure 2: Delineamento blocos completos casualizado

Delineamento inteiramente casualizado

- Dois princípios são contemplados:
 - repetição
 - casualização

Delineamento em Blocos Completos Casualizados (DBCC)

- No DBCC todos os três princípios são contemplados
- O **controle local** é representado pelos blocos, cada um dos quais inclui todos os tratamentos casualizados.
- Para o delineamento ser eficiente o bloco deve ser o mais homogêneo possível.

Automação de máquinas: Bloco são as diferentes máquinas



Automação de máquinas: Bloco são as diferentes máquinas



Um exemplo na agricultura: o que são blocos?

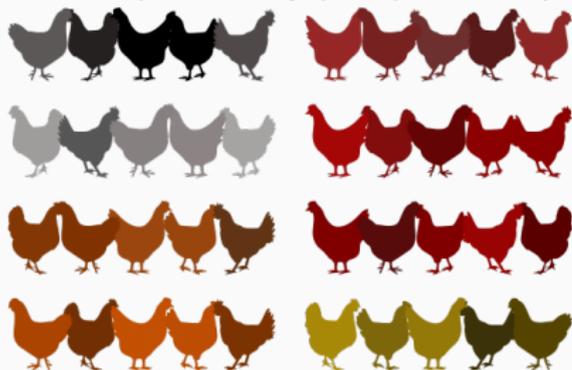


Outro exemplo

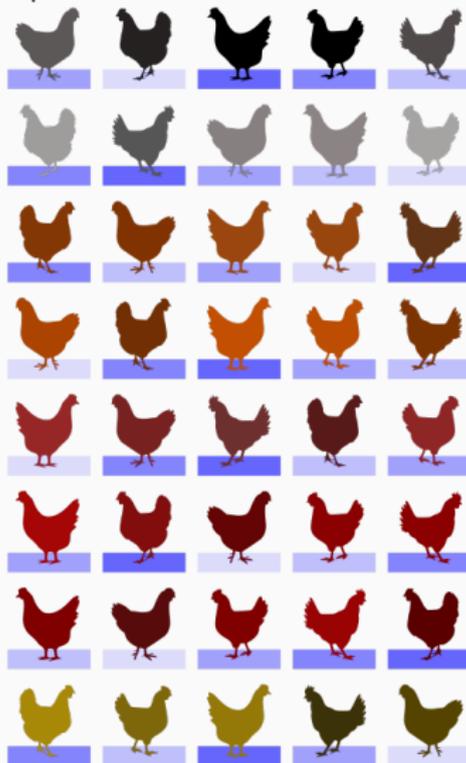
Unidades experimentais disponíveis



Unidades experimentais agrupadas por semelhança



Casualização dos tratamentos às unidades experimentais em cada bloco



- Em qualquer experimento, a variabilidade decorrente de um fator incômodo (nuisance factor) pode afetar os resultados.
- Geralmente, define-se um fator incômodo como um fator de delineamento, que provavelmente afeta a resposta, mas no qual não se tem interesse.
- A utilização de blocos é uma técnica extremamente importante usada extensivamente na experimentação (industrial, agrícola, etc)

Exemplo

Considere um experimento envolvendo a determinação de dureza em peças metálicas usando duas ponteiros diferentes. Inicialmente, propõe-se um delineamento inteiramente casualizado, onde metade das peças é testada com uma ponteira e a outra metade com a outra, permitindo a comparação das médias de dureza usando o teste t.

No entanto, esse método pode ser prejudicado se as peças não forem homogêneas, aumentando a variabilidade e dificultando a detecção de diferenças reais.

Para evitar isso, sugere-se um delineamento alternativo, dividindo cada peça em duas partes e testando cada parte com uma ponta diferente, reduzindo assim o efeito da falta de homogeneidade.

$$y_{ij} = \mu_i + \beta_j + \epsilon_{ij} \begin{cases} i = 1, 2 \\ j = 1, 2, \dots, 10 \end{cases} \quad (2.39)$$

■ **TABLE 2.6**
Data for the Hardness Testing Experiment

Specimen	Tip 1	Tip 2
1	7	6
2	3	3
3	3	5
4	4	3
5	8	8
6	3	2
7	2	4
8	9	9
9	5	4
10	4	5

Suponha agora que o objetivo é determinar se quatro ponteiros diferentes produzem ou não leituras diferentes em uma máquina de teste de dureza.

Teste



Equipamento para ensaio de dureza Rockwell. Suponha que o equipamento está sendo manuseado por um único operador e que estão sendo testadas quatro ponteiros diferentes. Nesse caso, o tipo de ponteira pode influenciar a leitura da dureza (suponha que o material submetido ao teste são diferentes, o que caracteriza blocos.)

$$y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

em que

- y_{ij} é o valor observado do tratamento i , no bloco j
- μ é uma constante inerente aos dados;
- τ_i é o efeito do tratamento i , $i = 1, \dots, a$
- β_j é o efeito do bloco j , $j = 1, \dots, b$
- e_{ij} são erros aleatórios independentes e identicamente distribuídos, $e_{ij} \sim N(0, \sigma^2)$
- $N = ab$ é o número de unidades experimentais

$$Y = X\theta + \epsilon$$

- $X_{N \times p}$ matriz do delineamento
- $\theta_{(1+a+b+1) \times 1}$ vetor de parâmetros

$$\theta' = [\mu, \beta_1, \beta_2, \dots, \beta_b, \tau_1, \tau_2, \dots, \tau_a, \sigma^2]$$

Exemplo matricial

- Considerare $a = 2$ e $b = 3 \implies N = ab = 6$

$$\begin{bmatrix} y_{11} \\ y_{21} \\ \hline y_{12} \\ y_{22} \\ \hline y_{13} \\ y_{23} \end{bmatrix} = \begin{bmatrix} \mu & \beta_1 & \beta_2 & \beta_3 & \tau_1 & \tau_2 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ \hline 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ \hline 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \tau_1 \\ \tau_2 \end{bmatrix} + \begin{bmatrix} \epsilon_{11} \\ \epsilon_{21} \\ \hline \epsilon_{12} \\ \epsilon_{22} \\ \hline \epsilon_{13} \\ \epsilon_{23} \end{bmatrix}$$

Dados

Block 1	Block 2	...	Block b
y_{11}	y_{12}		y_{1b}
y_{21}	y_{22}		y_{2b}
y_{31}	y_{32}		y_{3b}
.	.		.
.	.		.
.	.		.
y_{a1}	y_{a2}		y_{ab}

$$y_{i.} = \sum_{j=1}^b y_{ij} \quad i = 1, 2, \dots, a$$

$$y_{.j} = \sum_{i=1}^a y_{ij} \quad j = 1, 2, \dots, b$$

$$y_{..} = \sum_{i=1}^a \sum_{j=1}^b y_{ij} = \sum_{i=1}^a y_{i.} = \sum_{j=1}^b y_{.j}$$

$$\bar{y}_{i.} = y_{i.}/b \quad \bar{y}_{.j} = y_{.j}/a \quad \bar{y}_{..} = y_{..}/N$$

em que

- $y_{.j}$ é a soma dos valores observados no bloco $j = 1, \dots, b$
- $y_{i.}$ é a soma dos valores observados no tratamento $i = 1, \dots, a$
- $y_{..}$ é a soma total

Decomposição da soma de quadrados total corrigida pela média geral

$$\begin{aligned}\sum_{i=1}^a \sum_{j=1}^b (y_{ij} - \bar{y}_{..})^2 &= \sum_{i=1}^a \sum_{j=1}^b [(\bar{y}_i - \bar{y}_{..}) \\ &\quad + (\bar{y}_j - \bar{y}_{..}) + (y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}_{..})]^2 \\ \sum_{i=1}^a \sum_{j=1}^b (y_{ij} - \bar{y}_{..})^2 &= b \sum_{i=1}^a (\bar{y}_i - \bar{y}_{..})^2 + a \sum_{j=1}^b (\bar{y}_j - \bar{y}_{..})^2 \\ &\quad + \sum_{i=1}^a \sum_{j=1}^b (y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}_{..})^2 + 2 \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_i - \bar{y}_{..})(\bar{y}_j - \bar{y}_{..}) \\ &\quad + 2 \sum_{i=1}^a \sum_{j=1}^b (\bar{y}_j - \bar{y}_{..})(y_{ij} - \bar{y}_i - \bar{y}_j + \bar{y}_{..})\end{aligned}$$

Decomposição da soma de quadrados total corrigida pela média geral

$$\begin{aligned} \sum_{i=1}^a \sum_{j=1}^b (y_{ij} - \bar{y}_{..})^2 &= b \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2 + a \sum_{j=1}^b (\bar{y}_{.j} - \bar{y}_{..})^2 \\ &\quad + \sum_{i=1}^a \sum_{j=1}^b (y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y}_{..})^2 \end{aligned} \quad (4.7)$$

$$SS_T = SS_{\text{Treatments}} + SS_{\text{Blocks}} + SS_E \quad (4.8)$$

Esperança dos quadrados médios e teste F

$$E(MS_{\text{Treatments}}) = \sigma^2 + \frac{b \sum_{i=1}^a \tau_i^2}{a - 1}$$

$$E(MS_{\text{Blocks}}) = \sigma^2 + \frac{a \sum_{j=1}^b \beta_j^2}{b - 1}$$

$$E(MS_E) = \sigma^2$$

$$F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$$

■ TABLE 4.2

Analysis of Variance for a Randomized Complete Block Design

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Treatments	$SS_{\text{Treatments}}$	$a - 1$	$\frac{SS_{\text{Treatments}}}{a - 1}$	$\frac{MS_{\text{Treatments}}}{MS_E}$
Blocks	SS_{Blocks}	$b - 1$	$\frac{SS_{\text{Blocks}}}{b - 1}$	
Error	SS_E	$(a - 1)(b - 1)$	$\frac{SS_E}{(a - 1)(b - 1)}$	
Total	SS_T	$N - 1$		

Os cálculos para obtenção de soma de quadrados podem ser simplificados

$$SS_T = \sum_{i=1}^a \sum_{j=1}^b y_{ij}^2 - \frac{y_{..}^2}{N} \quad (4.9)$$

$$SS_{\text{Treatments}} = \frac{1}{b} \sum_{i=1}^a y_i^2 - \frac{y_{..}^2}{N} \quad (4.10)$$

$$SS_{\text{Blocks}} = \frac{1}{a} \sum_{j=1}^b y_j^2 - \frac{y_{..}^2}{N} \quad (4.11)$$

$$SS_E = SS_T - SS_{\text{Treatments}} - SS_{\text{Blocks}} \quad (4.12)$$

Restrição para estimação dos parâmetros

- Você lembra-se quais foram as formas utilizadas no DIC?
- Como você aplicaria em DBCC?
- Exercício
 - Uma delas: restrição soma zero

$$\sum_{i=1}^a \tau_i = 0 \quad \text{and} \quad \sum_{j=1}^b \beta_j = 0$$

- Obtenha a tabela ANOVA para o exemplo 4.1, Montgomery Ed.8 (página 144)
- Verifique se os pressupostos básicos estão satisfeitos
- Interprete o quadro da ANOVA
- Se o Teste F for significativo, realize um TCM:Tukey

Exemplo 4.1 Montgomery

Um fabricante de dispositivos médicos produz enxertos vasculares (veias artificiais). Esses enxertos são produzidos por extrusão de tarugos de resina de politetrafluoretileno (PTFE) combinados com um lubrificante em tubos. Frequentemente, alguns dos tubos em uma produção contém saliências pequenas e duras na Superfície externa. Esses defeitos são fontes de descarte.

O desenvolvedor do produto responsável pela área de enxertos vasculares suspeita que a pressão de extrusão afeta a ocorrência de filmes e por isso pretende realizar uma experiência para investigar esta hipótese.

Exemplo 4.1 Montgomery

A resina é fabricada por um fornecedor externo e entregue em lotes. O engenheiro também suspeita que pode haver variações significativas de lote para lote na composição, porque embora o material deva ser consistente com relação a parâmetros como peso molecular, proporção média do tamanho do artigo, retenção e proporção de altura de pico, provavelmente não é devido à variação de fabricação no fornecedor de resina e na variação natural no material.

Exemplo 4.1 Montgomery

O desenvolvedor do produto decide investigar o efeito de quatro níveis diferentes da pressão de extrusão em movimentos usando um teste aleatório de projeto completo de blocos considerando lotes de resina como blocos.

Observe que existem quatro níveis de pressão de extrusão (tratamentos) e seis lotes de resina (blocos). Lembre-se que a ordem em que a pressão de extrusão é testada dentro de cada bloco é aleatória. A variável de resposta é o rendimento, ou a porcentagem de tubos na execução de produção que não continha nenhum filme.

Exemplo 4.1 Montgomery

■ TABLE 4.3
Randomized Complete Block Design for the Vascular Graft Experiment

Extrusion Pressure (PSI)	Batch of Resin (Block)						Treatment Total
	1	2	3	4	5	6	
8500	90.3	89.2	98.2	93.9	87.4	97.9	556.9
8700	92.5	89.5	90.6	94.7	87.0	95.8	550.1
8900	85.5	90.8	89.6	86.2	88.0	93.4	533.5
9100	82.5	89.5	85.6	87.4	78.9	90.7	514.6
Block Totals	350.8	359.0	364.0	362.2	341.3	377.8	$y_{..} = 2155.1$

Exemplo 4.1 Montgomery

To perform the analysis of variance, we need the following sums of squares:

$$\begin{aligned}SS_T &= \sum_{i=1}^4 \sum_{j=1}^6 y_{ij}^2 - \frac{y_{..}^2}{N} \\&= 193,999.31 - \frac{(2155.1)^2}{24} = 480.31 \\SS_{\text{Treatments}} &= \frac{1}{b} \sum_{i=1}^4 y_i^2 - \frac{y_{..}^2}{N} \\&= \frac{1}{6} [(556.9)^2 + (550.1)^2 + (533.5)^2 \\&\quad + (514.6)^2] - \frac{(2155.1)^2}{24} = 178.17\end{aligned}$$

$$\begin{aligned}SS_{\text{Blocks}} &= \frac{1}{a} \sum_{j=1}^6 y_j^2 - \frac{y_{..}^2}{N} \\&= \frac{1}{4} [(350.8)^2 + (359.0)^2 + \cdots + (377.8)^2] \\&\quad - \frac{(2155.1)^2}{24} = 192.25 \\SS_E &= SS_T - SS_{\text{Treatments}} - SS_{\text{Blocks}} \\&= 480.31 - 178.17 - 192.25 = 109.89\end{aligned}$$

The ANOVA is shown in Table 4.4. Using $\alpha = 0.05$, the critical value of F is $F_{0.05, 3, 15} = 3.29$. Because $8.11 > 3.29$, we conclude that extrusion pressure affects the mean yield. The P -value for the test is also quite small. Also, the resin batches (blocks) seem to differ significantly, because the mean square for blocks is large relative to error.

Exemplo 4.1 Montgomery

■ TABLE 4.4

Analysis of Variance for the Vascular Graft Experiment

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P -Value
Treatments (extrusion pressure)	178.17	3	59.39	8.11	0.0019
Blocks (batches)	192.25	5	38.45		
Error	109.89	15	7.33		
Total	480.31	23			

Exercício (Experimento agrícola)

Adaptado de Pimentel Gomes (1990) página 57

Num experimento de competição de variedades de batatas, em blocos completos casualizados (DBCC), as produções obtidas, em *t/ha*, foram as seguintes.

Variedades	Bloco 1	Bloco 2	Bloco 3	Bloco 4
Kennebec	9,2	13,4	11,0	9,2
Huinkul	21,1	27,0	26,4	25,7
S. Rafaela	22,6	29,9	24,2	25,1
Buena Vista	15,4	11,9	10,1	12,3
B 25-50 E	12,7	18,0	18,2	17,1
B 1-52	20,0	21,1	20,0	28,0
B 116-51	23,1	24,2	26,4	16,3
B 72-53 A	18,0	24,6	24,0	24,6

Pede-se:

1. verificar se existe diferenças entre médias de tratamentos;
2. Se houver
 - a. Aplique o teste de Tukey. Interprete.
 - b. Conclua

Exercício (Área médica)

Adaptado de Pimentel Gomes (1990) página 61

Numa pesquisa sobre o efeito do óleo de milho no teor de colesterol do sangue, foram obtidos os seguintes dados: Antes da dieta de óleo de milho (X) e depois da dieta de óleo de milho (Y):

Teor de colesterol no sangue, em *mg*, por 100*g*

Paciente	Antes (X)	Depois (Y)
1	270	175
2	410	308
3	350	248
4	360	231
5	350	196
6	430	190
7	268	252

Pergunta-se:

- Há efeito significativo da dieta sobre o teor de colesterol no sangue?

amostras não pareadas

- Teste t para amostras independentes
- ANOVA (DIC)

amostras pareadas

- Teste t amostras dependentes
- ANOVA (DBCC)

Explore o R para fazer as análises de dados instalados em DBCC

```
bloco <- gl(4,8) #Quatro blocos 8 tratamentos
trat  <- gl(8,1,32)
prod  <- c( 9.2,21.1,22.6,15.4,12.7,20.0,23.1,18.0,
           13.4,27.0,29.9,11.9,18.0,21.1,24.2,24.6,
           11.0,26.4,24.2,10.1,18.2,20.0,26.4,24.0,
           9.2,25.7,25.1,12.3,17.1,28.0,16.3,24.6)

dad <- data.frame(bloco=bloco,trat=trat,prod=prod)
#trat = variedades
View(dad)
# Número de unidades experimentais
(n<-length(dad$bloco))
```

Estatísticas descritivas

Quantos blocos? E quantas observações por bloco?

```
tapply(dad$prod,dad$bloco,length)
```

4 blocos e 8 observações por bloco

```
b <- 4
```

Quantos tratamentos? E quantas observações dos tratamentos?

```
tapply(dad$prod,dad$trat,length)
```

8 tratamentos repetidos nos 4 blocos

```
a <- 8
```

Soma por bloco

```
(soma.bloco <- tapply(dad$prod,dad$bloco,sum))
```

Soma por tratamento

```
(soma.trat <- tapply(dad$prod,dad$trat,sum))
```

Soma total

```
(soma.total <- sum(prod))
```

```
#Média por tratamento
(media <- tapply(dad$prod,dad$trat,mean))
#Média geral
(media.geral <- mean(dad$prod))
boxplot(dad$prod ~ dad$trat )
#Aparentemente existem diferenças entre médias de tratamen

# Verificação das pressuposições básicas
mod1 <- aov(dad$prod ~ dad$bloco + dad$trat)
names(mod1)
res <- rstudent(mod1)
```

```
#Interprete
```

```
par(mfrow=c(1,2))
```

```
qqnorm(res)
```

```
qqline(res)
```

```
# Exiba os cinco passos do teste hipóteses
```

```
shapiro.test(res)
```

```
boxplot(res)
```

```
layout(1)
```

```
# Homogeneidade de variâncias  
# Interprete  
plot(mod1$fit,res)  
abline(h=0)  
require(lattice)  
dotplot(dad$prod ~ dad$trat)
```

```
# Exiba os 5 passos do teste de hipóteses
```

```
bartlett.test(dad$prod, dad$trat)
```

```
require(car)
```

```
##?leveneTest
```

```
leveneTest(dad$prod, dad$trat)
```

```
# Independência
```

```
##?durbinWatsonTest
```

```
durbinWatsonTest(lm(dad$prod ~ dad$bloco + dad$trat))
```

```
# Conclua sobre a verificação das pressuposições básicas
```

```
# Quadro da ANOVA
```

```
anova(mod1)
```

```
summary(mod1, intercept = T)
```

```
#Conclusão: Existe pelo menos um contraste
```

```
#de duas médias diferente de zero.
```

```
## Realizando os cálculos
```

```
# Compare com o quadro da ANOVA
```

```
(correcao <- round(soma.total^2/n,0))
```

```
correcao <- soma.total^2/n
```

```
(SQBloco<- round(sum(soma.bloco^2)/8-correcao,0))
```

```
(SQTrat<- round(sum(soma.trat^2)/4-correcao,0))
```

```
(SQTotal <- sum(t(dad$prod)%*%dad$prod)-correcao)
```

```
(SQRes <- round(SQTotal - SQBloco - SQTrat,0))
```

```
#Número de comparações  
#m número de comparações duas a duas  
a <- 8  
(m <- a*(a-1)/2)
```

```
# Tukey base  
TukeyHSD(mod1)  
plot(TukeyHSD(mod1), las=1)
```

```
# Médias em ordem decrescente  
(medias <- tapply(dad$prod,dad$trat,mean))  
sort(medias,decreasing = T)  
# A média do tratamento 3 recebe a letra a  
# A média do tratamento 3 difere da média  
# do tratamento 2? Não! Recebe a letra a
```

```
out <- TukeyHSD(mod1)
names(out)
out$dad$trat
dim(out$dad$trat)
out$dad$trat[1,]
# Sim, a média do tratamento 3 difere da média do
# tratamento 8? Não! Recebe a letra a
sort(medias,decreasing = T)

# A média do trat 3 difere da média do trat 7? sim

# e por ai vai até finalizar
```

```
#Teste Tukey agricolae
```

```
require(agricolae)
```

```
(gl.Res <- df.residual(mod1))
```

```
(QMRes <- sigma(mod1)^2)
```

```
out <- HSD.test(dad$prod, dad$trat, DFerror=gl.Res,  
               MSerror=QMRes,alpha = 0.05,  
               group=TRUE, main = NULL,  
               unbalanced=FALSE,console=FALSE)
```

```
out$statistics
```

```
media.geral
```

```
# Coeficiente de variação  
(CV <- sqrt(QMRes)/media.geral*100)
```

```
alpha <- 0.05
```

```
# q tabelado
```

```
(q.crit <- qtukey(1-alpha,a,gl.Res))
```

```
#Amplitude studentizada
```

```
#abs(yb_i-yb_j)/sqrt(QMREs/b)
```

```
#abs(yb_i-yb_j)/sqrt(QMREs/b) >qt_crit
```

```
#abs(yb_i-yb_j) > qt_crit*sqrt(QMREs=MSError/b)
```

```
#Diferença mínima significativa
```

```
(MSD <- q.crit*sqrt(QMRes/b))
```

```
print(out$groups)
```

```
out.comp <- HSD.test(dad$prod,dad$trat,group=F,  
                    MSError = QMRes,DFerror = gl.Res)
```

```
names(out)
```

```
print(out$statistics)
```

```
print(out$comparison)
```

```
out.graf <- HSD.test(dad$prod,dad$trat,group=T,  
                    MSError = QMRes,DFerror = gl.Res)
```

```
# Veja as diferenças e similaridades  
plot(out, las=1)  
#plot(out.comp, las=1)  
plot(out.graf, las=1)
```