

Delineamento em blocos casualizados

Distribuição Bernoulli

2022

Os dados são do Exemplo 4.4, pag. 93 do livro Demétrio, C. G. B. (2002), *Modelos Lineares Generalizados em Experimentação Agronômica*, ESALQ (<https://docs.ufpr.br/~niveam/micro%20da%20sala/bom/Apostila%20de%20MLG.pdf>).

A coleta dos dados foi em um delineamento em blocos casualizados com 10 blocos. Os tratamentos formam um esquema fatorial $2 \times 3 \times 3$ com fatores

- A: 2 tipos de citocinina (BAP e TDZ),
- B: 3 níveis de citocinina (5,0, 1,0 e 0,1) e
- C: 3 tipos de auxina (NAA, IBA e 2-4D).

Recipientes foram preparados com as 18 combinações dos níveis dos três fatores. Em cada recipiente foi adicionado um fragmento de maçã (explante). A variável resposta Y é tal que $Y = 1$, se o explante regenerou após quatro semanas, e $Y = 0$, caso contrário.

O objetivo consiste em avaliar a influência dos três fatores e de suas interações sobre a regeneração dos fragmentos. A linguagem R é utilizada no exemplo.

```
# Separador decimal nos resultados: ","
options(OutDec = ",")

# Dados (Tabela 31, pag. 94)
y <- c(1, 1, 0, 0, 1, 0, 1, 0, 1, 1,
      0, 1, 1, 1, 1, 1, 0, 1, 1, 1,
      1, 1, 1, 1, 1, 1, 1, 0, 0, 1,
      0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
      1, 1, 1, 0, 0, 1, 1, 0, 1, 1,
      1, 0, 1, 1, 0, 1, 1, 1, 1, 1,
      0, 0, 1, 1, 1, 0, 1, 0, 0, 0,
      0, 0, 0, 1, 1, 1, 1, 0, 1, 0,
      0, 0, 1, 1, 1, 1, 1, 0, 1, 1,
      1, 1, 1, 1, 1, 0, 1, 1, 1, 1,
      1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
      1, 0, 1, 1, 1, 1, 1, 1, 1, 1,
      1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
      1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
      1, 1, 1, 1, 1, 1, 1, 1, 1, 0,
      1, 1, 1, 1, 1, 1, 0, 1, 1, 1,
      1, 1, 1, 1, 1, 1, 0, 1, 1, 1,
      0, 0, 1, 0, 1, 1, 1, 1, 1, 1)
```

```
n <- length(y)
# Fator A
tipoc <- factor(c(rep("BAP", n / 2), rep("TDZ", n / 2)))
# Fator B
```

```
nivelc <- factor(rep(c(5, 1, 0.1), each = n / 6, 2))
# Fator C
tipoa <- factor(rep(c("NAA", "IBA", "2-4D"), each = n / 18, 6))
bloco <- factor(rep(1:10, n / 10))
```

A função `factor` cria variáveis qualitativas (da classe *factor*). Para uma variável qualitativa com k níveis são criadas $k - 1$ variáveis auxiliares (*dummy*), sendo que o primeiro nível da variável é a referência e os coeficientes das variáveis auxiliares representam diferenças no preditor linear em relação ao nível de referência.

Nota 1. Explique o código acima, verificando que `y` corresponde às linhas da Tabela 31 em um vetor.

O tamanho da amostra é $n = 2 \times 3 \times 3 \times 10 = 180$.

Em seguida são ajustados os modelos da Tabela 32, notando que não é uma sequência de 18 modelos encaixados. A função `update` é utilizada para incluir termos em um modelo já ajustado (e também pode ser usada para excluir termos), ou seja, geramos modelos encaixados (por exemplo, `mbnivel` está encaixado em `mbloco`). A função de ligação logito (canônica, ligação *default*) é adotada.

```
# Modelos
mnulo <- glm(y ~ 1, family = binomial)
mbloco <- update(mnulo, . ~ . + bloco)
mbtipo <- update(mbloco, . ~ . + tipoc)
mbnivel <- update(mbloco, . ~ . + nivelc)
mbauxina <- update(mbloco, . ~ . + tipoa)
mbtnivel <- update(mbtipo, . ~ . + nivelc)
mbtauxina <- update(mbtipo, . ~ . + tipoa)
mbnauxina <- update(mbnivel, . ~ . + tipoa)
mbtnivel2 <- update(mbtnivel, . ~ . + tipoc:nivelc)
mbtauxina2 <- update(mbtauxina, . ~ . + tipoc:tipoa)
mbnauxina2 <- update(mbnauxina, . ~ . + nivelc:tipoa)
mbtnivel2auxina <- update(mbtnivel2, . ~ . + tipoa)
mbtauxina2nivel <- update(mbtauxina2, . ~ . + nivelc)
mbnauxina2tipo <- update(mbnauxina2, . ~ . + tipoc)
mbtnivel2auxina2 <- update(mbtnivel2auxina, . ~ . + tipoc:tipoa)
mbtnivel2auxina3 <- update(mbtnivel2auxina, . ~ . + nivelc:tipoa)
mbtauxina2nivel2 <- update(mbtauxina2nivel, . ~ . + nivelc:tipoa)
mbtnauxina3 <- update(mbtauxina2nivel2, . ~ . + tipoc:nivelc + tipoc:nivelc:tipoa)
```

Nota 2. Compare os modelos ajustados acima com os modelos da Tabela 32. Por exemplo, para o modelo `mbtipo`, a desviância e os graus de liberdade podem ser obtidos com `mbtipo$deviance` e `mbtipo$df.residual` ou `deviance(mbtipo)` e `df.residual(mbtipo)`.

Outros modelos são ajustados de modo a testar a sequência de modelos encaixados partindo do modelo nulo e incluindo os efeitos de bloco, tipo de citocinina (A), nível de citocinina (B | A), tipo de auxina (C | A, B), interação A:B (AB | A, B, C), interação A:C (AC | A*B, C), interação B:C (BC | A*B, A*C) e interação A:B:C (ABC | A*B, A*C, B*C). O operador “*” indica que são incluídos os efeitos individuais e a interação; por exemplo, `A*B` se desdobra em `A + B + A:B`. A tabela ANODEV é apresentada.

```
## Outros modelos
# tipoc + nivelc + tipoa
mbtnauxina <- update(mbtnivel, . ~ . + tipoa)
# sem interação de segunda ordem
mbtnauxina2o <- update(mbtnauxina3, . ~ . - tipoc:nivelc:tipoa)

# Tabela ANODEV
anova(mnulo, mbloco, mbtipo, mbtnivel, mbtnauxina, mbtnivel2auxina,
      mbtnivel2auxina2, mbtnauxina2o, mbtnauxina3, test = "Chi")
```

```
## Analysis of Deviance Table
##
## Model 1: y ~ 1
## Model 2: y ~ bloco
## Model 3: y ~ bloco + tipoc
## Model 4: y ~ bloco + tipoc + nivelc
## Model 5: y ~ bloco + tipoc + nivelc + tipoa
## Model 6: y ~ bloco + tipoc + nivelc + tipoa + tipoc:nivelc
## Model 7: y ~ bloco + tipoc + nivelc + tipoa + tipoc:nivelc + tipoc:tipoa
## Model 8: y ~ bloco + tipoc + tipoa + nivelc + tipoc:tipoa + tipoa:nivelc +
## tipoc:nivelc
## Model 9: y ~ bloco + tipoc + tipoa + nivelc + tipoc:tipoa + tipoa:nivelc +
## tipoc:nivelc + tipoc:tipoa:nivelc
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1 179 202,44
## 2 170 193,79 9 8,6495 0,470237
## 3 169 165,65 1 28,1394 1,129e-07 ***
## 4 167 160,83 2 4,8175 0,089929 .
## 5 165 152,46 2 8,3751 0,015184 *
## 6 163 149,03 2 3,4250 0,180414
## 7 161 138,66 2 10,3706 0,005598 **
## 8 157 132,65 4 6,0143 0,198079
## 9 153 127,00 4 5,6453 0,227252
## ---
## Signif. codes: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1
```

Nota 3. O modelo mbtntaxina3 (Model 9 na tabela ANODEV) pode ser ajustado com o comando
`mbtntaxina3 <- glm(y ~ bloco + (tipoc + nivelc + tipoa)^3, family = binomial)`

ou

```
mbtntaxina3 <- glm(y ~ bloco + tipoc * nivelc * tipoa, family = binomial).
```

Nota 4. Compare os resultados acima com a Tabela 33. Com base nos resultados acima, justifique que o efeito do nível de citocinina (`nivelc`) não é significativo, ao passo que os efeitos do tipo de citocinina (`tipoc`), tipo de auxina (`auxina`) e da interação entre estes dois (`tipoc:auxina`) são significativos a um nível de 5%. Assim, foi selecionado o modelo mbtntaxina2. Cada termo deste modelo é analisado em seguida testando modelos encaixados.

```
# Tabela ANODEV
anova(mnulo, mbloco, mbtipo, mbtntaxina, mbtntaxina2, test = "Chi")
```

```
## Analysis of Deviance Table
##
## Model 1: y ~ 1
## Model 2: y ~ bloco
## Model 3: y ~ bloco + tipoc
## Model 4: y ~ bloco + tipoc + tipoa
## Model 5: y ~ bloco + tipoc + tipoa + tipoc:tipoa
## Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1 179 202,44
## 2 170 193,79 9 8,6495 0,470237
## 3 169 165,65 1 28,1394 1,129e-07 ***
## 4 167 157,57 2 8,0795 0,017602 *
## 5 165 147,82 2 9,7530 0,007624 **
## ---
```

```
## Signif. codes:  0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1
```

As estimativas dos coeficientes do modelo selecionado (mbtauxina2) são apresentadas abaixo.

```
summary(mbtauxina2)
```

```
##
## Call:
## glm(formula = y ~ bloco + tipoc + tipoa + tipoc:tipoa, family = binomial)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2,91001  0,00647  0,37263  0,64961  1,95423
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      0,6748     0,7103   0,950 0,342110
## bloco2           -0,3401     0,8263  -0,412 0,680653
## bloco3            1,2469     0,9399   1,327 0,184633
## bloco4            0,7738     0,8894   0,870 0,384293
## bloco5            1,2469     0,9399   1,327 0,184633
## bloco6            0,7738     0,8894   0,870 0,384293
## bloco7            1,2469     0,9399   1,327 0,184633
## bloco8           -0,6643     0,8210  -0,809 0,418410
## bloco9            1,2469     0,9399   1,327 0,184633
## bloco10           0,7738     0,8894   0,870 0,384293
## tipocTDZ          0,4585     0,6824   0,672 0,501631
## tipoaIBA         -0,5508     0,6114  -0,901 0,367615
## tipoaNAA         -2,0838     0,6181  -3,372 0,000747 ***
## tipocTDZ:tipoaIBA  2,3902     1,3023   1,835 0,066446 .
## tipocTDZ:tipoaNAA  3,1742     1,0995   2,887 0,003891 **
## ---
## Signif. codes:  0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 202,44  on 179  degrees of freedom
## Residual deviance: 147,82  on 165  degrees of freedom
## AIC: 177,82
##
## Number of Fisher Scoring iterations: 6
```

A matriz do modelo (\mathbf{X}) pode ser obtida com a função `model.matrix`.

```
dim(model.matrix(mbtauxina2))
```

```
## [1] 180 15
```

O número de linhas é o tamanho da amostra n . As 15 colunas correspondem aos coeficientes (β) assim quantificados: intercepto (1), blocos (9), tipos de citocinina (1), tipos de auxina (2) e interações entre tipos de citocinina e tipos de auxina (2).

Em seguida apresentamos estimativas da razão de chances da citocinina TDZ em relação à citocinina BAP (lembrando que BAP é o nível de referência). O modelo selecionado inclui a interação entre tipo de citocinina e tipo de auxina (`tipoc:tipoa`), de modo que a razão de chances depende do tipo de auxina.

```
## Razão de chances de TDZ em relação a BAP
# auxina 2-4D (nível de referência)
```

```
exp(coef(mbtauxina2)["tipocTDZ"])
```

```
## tipocTDZ
```

```
## 1,581699
```

```
# auxina IBA
```

```
exp(coef(mbtauxina2)["tipocTDZ"] + coef(mbtauxina2)["tipocTDZ:tipoaIBA"])
```

```
## tipocTDZ
```

```
## 17,26558
```

```
# auxina NAA
```

```
exp(coef(mbtauxina2)["tipocTDZ"] + coef(mbtauxina2)["tipocTDZ:tipoaNAA"])
```

```
## tipocTDZ
```

```
## 37,81452
```

Nota 5. Não foi verificado se o modelo faz um bom ajuste aos dados.

Nota 6. Refaça o exemplo em linguagem Python.