

Universidade Federal do Paraná - Departamento de Estatística

CE225 - Modelos Lineares Generalizados

01/09/2017

Os dados analisados referem-se a 81 crianças submetidas a uma cirurgia de correção da coluna.

- Variável resposta: **kyphosis** - fator com dois níveis (ausente ou presente), indicando se uma cifose (tipo de deformação) foi verificada após o procedimento cirúrgico.
- Variáveis explicativas:
 - **Age** - idade (em meses);
 - **Number** - número de vértebras envolvidas na cirurgia;
 - **Start** - número da primeira vértebra operada (de cima para baixo).

Na sequência são apresentados os resumos de dois modelos lineares generalizados ajustados. Para ambos se considerou distribuição binomial e função de ligação logito.

Modelo 1 - Modelo aditivo considerando as três covariáveis.

```
ajuste1 <- glm(Kyphosis ~ Age + Number + Start, family = binomial(link = logit),
              data = kyphosis)
summary(ajuste1)
```

```
##
## Call:
## glm(formula = Kyphosis ~ Age + Number + Start, family = binomial(link = logit),
##      data = kyphosis)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.3124  -0.5484  -0.3632  -0.1659   2.1613
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.036934   1.449575  -1.405  0.15996
## Age          0.010930   0.006446   1.696  0.08996 .
## Number       0.410601   0.224861   1.826  0.06785 .
## Start       -0.206510   0.067699  -3.050  0.00229 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 83.234  on 80  degrees of freedom
## Residual deviance: 61.380  on 77  degrees of freedom
## AIC: 69.38
##
## Number of Fisher Scoring iterations: 5
```

Modelo 2 - Modelo considerando apenas a primeira vértebra operada como covariável.

```
ajuste2 <- glm(Kyphosis ~ Start, family = binomial(link = logit), data = kyphosis)
summary(ajuste2)
```

```
##
## Call:
## glm(formula = Kyphosis ~ Start, family = binomial(link = logit),
##      data = kyphosis)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4729  -0.5176  -0.4211  -0.3413   2.1305
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.89007    0.62996   1.413 0.157686
## Start       -0.21789    0.06044  -3.605 0.000312 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 83.234  on 80  degrees of freedom
## Residual deviance: 68.072  on 79  degrees of freedom
## AIC: 72.072
##
## Number of Fisher Scoring iterations: 5
```

Seguem resultados referentes às predições fornecidas pelo primeiro modelo para dois perfis de pacientes.

```
perfil1 <- data.frame(Age = 150, Number = 6, Start = 3)
predict(ajuste1, newdata = perfil1, type = 'link', se.fit = TRUE)
```

```
## $fit
##      1
## 1.446716
##
## $se.fit
## [1] 0.7830221
##
## $residual.scale
## [1] 1
```

```
perfil2 <- data.frame(Age = 50, Number = 2, Start = 12)
predict(ajuste1, newdata = perfil2, type = 'link', se.fit = TRUE)
```

```
## $fit
##      1
## -3.147328
##
## $se.fit
## [1] 0.7980145
##
## $residual.scale
## [1] 1
```

- a) Escreva a expressão do modelo ajustado (`ajuste1`) nas escalas do preditor e da resposta;
- b) Com base nos sinais das estimativas, quais fatores estão associados a uma maior incidência de cifose no pós-operatório?
- c) Qual a probabilidade estimada de cifose para uma criança com 6 anos recém completos, cuja cirurgia envolva 4 vértebras a partir da 3ª?
- d) Apresente um intervalo de confiança (95%) para o parâmetro de regressão correspondente à idade;
- e) Apresente intervalos de confiança (95%) para a probabilidade de cifose para crianças com os dois perfis definidos nos códigos;
- f) Teste a significância do modelo ($H_0 : \beta_1 = \beta_2 = \beta_3 = 0$) usando o teste da razão de verossimilhanças;
- g) Teste a hipótese $H_0 : \beta_2 = 0$ com base na estatística de Wald. Qual seria sua conclusão aos níveis de 10 e 5% de significância? O mesmo para β_3 .
- h) Agora considere o `ajuste2`. Esse modelo pode ser obtido a partir de alguma restrição aplicada aos parâmetros do primeiro modelo? Justifique.
- i) Teste a hipótese $H_0 : \beta_2 = \beta_3 = 0$. Qual seria sua decisão ao nível de significância de 5%?
- j) Compare os resultados dos testes apresentados nos itens (g) e (i). Explique possíveis diferenças.
- k) Qual a diferença dos testes sequencias apresentados nos seguintes quadros de análise de deviance? Interprete-os.

```
anova(ajuste1)
```

```
## Analysis of Deviance Table
##
## Model: binomial, link: logit
##
## Response: Kyphosis
##
## Terms added sequentially (first to last)
##
##
##          Df Deviance Resid. Df Resid. Dev
## NULL                80      83.234
## Age                1    1.302      79      81.932
## Number             1   10.306      78      71.627
## Start              1   10.247      77      61.380
```

```
Anova(ajuste1)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: Kyphosis
##          LR Chisq Df Pr(>Chisq)
## Age          3.1565 1 0.075623 .
## Number       3.9191 1 0.047739 *
## Start       10.2466 1 0.001369 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```