Soluções de exercícios da Lista 3

1) Regressão logística

a) Pela saída do R obtemos o seguinte modelo ajustado:

$$\text{logito}(p(x)) = -3.777 + 0.145 \text{ LI}$$

Portanto, o valor predito de p(26) é dado por

$$\exp\{-3.77 + 0.145 \text{ x } 26\} / (1 + \exp\{-3.77 + 0.145 \text{ x } 26\}) = 0.497$$

b) $\exp\{0.145\} = 1.16$

c) Primeiro vamos obter os quartis de LI usando a função do R

> quantile(LI, c(0.25,0.75))

Resulta em Q1= 13 e Q3= 25

Obtendo as probabilidades preditas associadas aos quartis:

$p(13) = \exp\{-3.77 + 0.145 \text{ x } 13\} / (1 + \exp\{-3.77 + 0.145 \text{ x } 13\}) = 0.131$

$p(25) = \exp\{-3.77 + 0.145 \text{ x } 25\} / (1 + \exp\{-3.77 + 0.145 \text{ x } 25\}) = 0.461$

Diferença das probabilidades: 0.461 – 0.131 = 0.33

d) Intervalo de Wald para intercepto

(-3.777 – 1.645 x 1.3786 ; 3.777 + 1.645 x 1.3786) = (- 6.04 ; -1.51)

e) Estatística da RV

$$Q_{RV} = 34.372 – 26.073 = 8.3$$

Graus de liberdade: 26 - 25 = 1

Valor-P < 0.004 . Rejeita-se H_0 .

O index LI é significante para o modelo.

3) Modelo Poisson/loglinear

Entrando com os dados:RV

x=c(rep(0,10),rep(1,10))
y=c(8,7,6,6,3,4,7,2,3,4,9,9,8,14,8,13,11,5,7,6)
ajuste=glm(y~x, family=poisson)

Saída:

```
summary(ajuste)

Call:
glm(formula = y ~ x, family = poisson)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5280  -0.7622  -0.1699   0.6938   1.5399

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   1.6094     0.1414  11.380  < 2e-16 ***
x             0.5878     0.1764   3.332 0.000861 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 27.857  on 19  degrees of freedom
Residual deviance: 16.268  on 18  degrees of freedom
AIC: 94.349

Number of Fisher Scoring iterations: 4
```

Resposta:

A equação de predição: $\log \mu = 1.061 + 0.588\ X$

Interpretação da estimativa de β :

A diferença dos logaritmos das taxas médias de imperfeições entre o tratamento B e tratamento A é aproximadamente 0.588.

Interpretação da estimativa de $\exp\{\beta\}$: $\exp\{0.588\} = 1.8$

A taxa média de imperfeições para o tratamento B é aproximadamente 1.8 vezes a taxa média do tratamento A. Indicando que, em média, o tratamento A é melhor que o tratamento B.

Vamos adicionar algumas questões ao exercício.

(b) Obtenha a estatística da razão de verossimilhança para testar $H_0 : \beta = 0$. Qual a sua conclusão ao nível de significância 0.01?

$$Q_{RV} = D_0 - D_1 = 27.857 - 16.268 = 11.589$$

Valor-P < 0.0007 . A evidência em favor $H_0 : \beta=0$ é muito pequena.

Considerando níveis de significância usuais, rejeita-se H_0 .

Concluímos que as taxas médias de erros nos processos são diferentes.

(c) Obtenha um IC(0.99) para a razão μ_B / μ_A .

Intervalo de Wald:

$$(\exp\{0.588 - 2.58 \times 0.1764\} ; \exp\{0.588 + 2.58 \times 0.1764\}) = (1.14 ; 2.84)$$

Indicando que $\mu_B > \mu_A$.

7) Comparação entre dados agrupados e não agrupados.

Dados não agrupados:

```
> x=c(rep(0,4),rep(1,4),rep(2,4))
> x
 [1] 0 0 0 0 1 1 1 1 2 2 2 2
> y=c(1,0,0,0,1,1,0,0,1,1,1,1)
> y
 [1] 1 0 0 0 1 1 0 0 1 1 1 1
```

Saída:

```
Call:
glm(formula = y ~ x, family = binomial)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4216  -0.6339   0.3752   0.5193   1.8459

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -1.503      1.181  -1.272   0.2033
x              2.060      1.130   1.823   0.0682 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 16.301  on 11  degrees of freedom
Residual deviance: 11.028  on 10  degrees of freedom
AIC: 15.028

Number of Fisher Scoring iterations: 4
```

Dados agrupados:

```
> rsim=c(1,2,4)
> rnao=c(3,2,0)
> x=c(0,1,2)
```

Saída:

```
Call:
glm(formula = cbind(rsim, rnao) ~ x, family = binomial)

Deviance Residuals:
      1        2        3
 0.3377  -0.5543   0.7504

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)   -1.503      1.181  -1.272   0.2034
x              2.060      1.130   1.823   0.0683 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 6.2568  on 2  degrees of freedom
Residual deviance: 0.9844  on 1  degrees of freedom
AIC: 8.6722

Number of Fisher Scoring iterations: 4
```

Resposta:

a) Os valores das funções desvios são diferentes nos dois casos.

Para dados não agrupados: D0 = 16.301 e D1 = 11.028

Para dados agrupados: D0= 6.2568 e D1= 0.9844

b) Para dados não agrupados: D0-D1= 5.273

Para dados agrupados: D0-D1=5.2724

Pequena diferença devido a arredondamento de valores.

As medidas devem ser iguais.