Lista 5

Guilherme Navarro NºUSP: 8943160

Exercício 1

Regressão logistica ajustada:

$$logito[\pi_i(x_i, w_i)] = \alpha + \beta x_i + \gamma(w_i - 5)$$

Tabela de estimativas para os parâmetros:

Parâmetro	Estimativa	Erro-Padão	Valor-p
α	0.69	0.12	<0.01
β	0.33	0.10	<0.01
γ	-0.03	0.005	<0.01

a)

O Parâmetro α em termos de razão de chances é: $e^{\widehat{\alpha}} = e^{0.69} = 1.99$ que siginifica que uma criança do gênero feminino com 5 anos de idade preferir Kcola tem 2 vezes a chance quando se tem a mesma condição para o gênero masculino.

O Parâmetro β em termos de razão de chances é: $e^{\widehat{\beta}} = e^{0.33} = 1.39$ que siginifica que a chance de uma criança de 5 anos, do genero masculino tem 1.4 vezes a chance quando se tem a mesma condição para o gênero feminino.

O Parâmetro γ em termos de razão de chances é: $e^{\widehat{\gamma}} = e^{-0.03} = 0.97$ que siginifica a chance de uma criança do gênero feminino com idade maior ou igual a 5 preferir Kcola é de 97% da chance quando se tem a mesma condição para o gênero masculino.

b) Para crianças do gênero feminino, temos:

I-) Com 10 anos:

$logito[\pi_i(0, 10)] = 0.69 + 0.33 * 0 - 0.03 * (10 - 5) = 0.54 \Rightarrow e^{0.54}$

II-) Com 15 anos:

$logito[\pi_i(0, 15)] = 0.69 + 0.33 * 0 - 0.03 * (15 - 5) = 0.39 \Rightarrow e^{0.39}$

Assim temos :

$logito[\pi_i(0, 10)] - logito[\pi_i(0, 15)] = 0.54 - 0.39 = 0.15 \Rightarrow e^{0.15} = 1.16$

Para crianças do gênero masculino, temos:

I-) Com 10 anos:

$logito[\pi_i(0, 10)] = 0.69 + 0.33 * 1 - 0.03 * (10 - 5) = 0.87$

II-) Com 15 anos:

$logito[\pi_i(0, 15)] = 0.69 + 0.33 * 1 - 0.03 * (15 - 5) = 0.72$

Assim temos :

$logito[\pi_i(0, 10)] - logito[\pi_i(0, 15)] = 0.87 - 0.72 = 0.15 \Rightarrow e^{0.15} = 1.16$

Logo posso concluir que com as informações acima, a razão de chances estimada de preferência por Kcola para crianças do mesmo gênero com 10 e 15 anos é a mesma.

c)

Como visto em aula pelo método 2 (método delta), temos que:

$$\frac{e^{\widehat{\beta}} - e^{\beta}}{e^{\widehat{\beta}}\widehat{ep}(\widehat{\beta})} \sim N(0,1)$$

Portanto:

$$IC(e^{\beta}; 0.95) = [e^{\widehat{\beta}} - 1.96 * e^{\widehat{\beta}}\widehat{ep}(\widehat{\beta}), e^{\widehat{\beta}} + 1.96 * e^{\widehat{\beta}}\widehat{ep}(\widehat{\beta})]$$

Substituindo os valores:

$$IC(e^{\beta}; 0.95) = [1.391 - 1.96 * 0.01 * 1.391, 1.391 + 1.96 * 0.01 * 1.391] \Rightarrow IC(e^{\beta}; 0.95) = [1.364; 1.418]$$

Análogamnte para e^{γ}, temos que:

$$IC(e^{\gamma}; 0.95) = [0.97 - 1.96 * 0.97 * 0.005; 0.97 - 1.96 * 0.97 * 0.005] \Rightarrow IC(e^{\gamma}; 0.95) = [0.96; 0.98]$$

d)

Probabilidade de meninos com 15 anos preferirem Kcola:

$$Y = logito[\pi_i(1, 15)] = 0.69 + 0.33 * 1 - 0.03(15 - 5) = 0.72$$

Mas como:

$$P(Y = 1) = \frac{e^{\alpha + \beta x_i + \gamma(w_i - 5)}}{e^{\alpha + \beta x_i + \gamma(w_i - 5)} + 1} \Rightarrow P(Y = 1) = \frac{e^{0.72}}{e^{0.72} + 1} = 0.67$$

Exercício 2

a)

Nesse modelo temos: $log(\frac{\pi(x)}{1-\pi(x)}) = \alpha + \beta x \Rightarrow \frac{\pi(x)}{1-\pi(x)} = e^{\alpha+\beta x} \Rightarrow \pi(x) = e^{\alpha+\beta x}(1 - \pi(x))$

$\Rightarrow \pi(x) = e^{\alpha+\beta x} - e^{\alpha+\beta x}\pi(x) \Rightarrow \pi(x) + e^{\alpha+\beta x}\pi(x) = e^{\alpha+\beta x} \Rightarrow \pi(x)(1 + e^{\alpha+\beta x}) = e^{\alpha+\beta x} \Rightarrow \pi(x) = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$

Como a função $e^{\alpha+\beta x}$ é sempre estritamente positiva logo podemos concluir que $\pi(x) > 0$ Agora basta mostrar que $\pi(x) < 1$ Para isso vamos chamar $e^{\alpha+\beta x} + 1 = k$ Podemos notar que:

$\pi(x) = \frac{k-1}{k} \Rightarrow \pi(x) = 1 - \frac{1}{k}$ Como $0 < \frac{1}{k} \leq 1$ $\forall k \in \mathbb{R}^*$ $\therefore$ a diferença $0 < 1 - \frac{1}{k} < 1 \therefore 0 < \pi(x) < 1\blacksquare$

b)

Como sabemos $Y = log(\frac{\pi(x)}{1-\pi(x)}) = \alpha + \beta x$ e que $P(Y = 1) = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}} \Rightarrow P(Y = 1) = \frac{1}{2} = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$

Seja $K = \frac{1}{1+e^{\alpha+\beta x}} \Rightarrow \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}} = 1 - \frac{1}{1+e^{\alpha+\beta x}} = 1 - K$

E como $K = \frac{1}{1+e^{\alpha+\beta x}} \Rightarrow 1 - K = \frac{1}{2} \Rightarrow K = \frac{1}{2}$

Assim, temos que:

$$\frac{1}{1 + e^{\alpha+\beta x}} = \frac{1}{2} \Rightarrow e^{\alpha+\beta x} + 1 = 2 \Rightarrow e^{\alpha+\beta x} = 1 \Rightarrow ln(e^{\alpha+\beta x}) = ln(1) = 0 \Rightarrow \alpha + \beta x = 0 \Rightarrow x = \frac{-\alpha}{\beta}$$

c)

Supondo que num determinado exemplo as estimativas são:

$\widehat{\alpha} = -4.57 \pm 0.87$

$\widehat{\beta} = 1.16 \pm 0.22$

$\widehat{corr}(\alpha, \beta) = 0.17$

E sabendo que $cov(\alpha, \beta) = corr(\alpha, \beta) * \sigma_\alpha * \sigma_\beta$ e que $x_0 = \widehat{\alpha}$, $y_0 = \widehat{\beta}$, $x = \alpha$, $y = \beta$

Utilizando o método Delta Bivariado, temos:

Para uma função $g(x, y) = \frac{x}{y}$,usando Taylor (1a.ordem) podemos escrever:

$g(x, y) = \frac{x}{y} = g(x_0, y_0) + \frac{\partial g(x_0,y_0)}{\partial x}(x - x_0) + \frac{\partial g(x_0,y_0)}{\partial y}(y - y_0)$

Assim:

$g(\widehat{\alpha}, \widehat{\beta}) = \frac{\widehat{\alpha}}{\widehat{\beta}} + \frac{1}{\widehat{\beta}}(\widehat{\alpha} - \alpha) - \frac{\widehat{\alpha}}{\widehat{\beta}^2}(\widehat{\beta} - \beta) = \frac{\widehat{\alpha}}{\widehat{\beta}} + \frac{\widehat{\alpha}}{\widehat{\beta}} - \frac{\alpha}{\widehat{\beta}} - \frac{\widehat{\alpha}}{\widehat{\beta}} + \frac{\widehat{\alpha}\beta}{\widehat{\beta}^2} = \frac{\widehat{\alpha}}{\widehat{\beta}} - \frac{\alpha}{\widehat{\beta}} + \frac{\widehat{\alpha}\beta}{\widehat{\beta}^2}$

Logo

$Var(g(\widehat{\alpha}, \widehat{\beta})) = Var(\frac{\widehat{\alpha}}{\widehat{\beta}} - \frac{\alpha}{\widehat{\beta}} + \frac{\widehat{\alpha}\beta}{\widehat{\beta}^2}) = Var(\frac{\alpha}{\widehat{\beta}} + \frac{\widehat{\alpha}\beta}{\widehat{\beta}^2}) + 2Corr(\widehat{\alpha}, \widehat{\beta}) * \widehat{\sigma}_\alpha * \widehat{\sigma}_\beta \Rightarrow Var(g(\widehat{\alpha}, \widehat{\beta})) =$ $Var(\alpha)(\frac{1}{\widehat{\beta}})^2 + Var(\beta)(\frac{\widehat{\alpha}}{\widehat{\beta}^2})^2 \Rightarrow Var(g(\widehat{\alpha}, \widehat{\beta})) = (0.87)^2(\frac{1}{1.16})^2 + (0.22)^2(\frac{-4.57}{1.16^2})^2 + 2*0.17*0.22*0.87 = 1.185848$

Assim o $IC(DL50^*; 0,95) = \frac{\widehat{\alpha}}{\widehat{\beta}} \pm 1.96\sqrt{1.186} \Rightarrow IC(DL50^*; 0,95) = -3.9396 \pm 2.1334 \Rightarrow IC(DL50^*; 0,95) =$ $[-6.0739; -1.8052]$

d)

Com as mesmas informações do item (c), a razão de chances estimada em comparação de indivíduos que fumam 3 maços de cigarros por dia com indivíduos que fumam apenas 1 maço de cigarros diariamente é:

$log(\frac{\pi(x)}{1-\pi(x)}) = \alpha + \beta x \Rightarrow RC = \frac{\pi(x)}{1-\pi(x)} = e^{\alpha + \beta x}$

Substituindo os valores a razão de chances estimada é:

$\widehat{RC} = \frac{\pi(x)}{1-\pi(x)} = \frac{e^{-4.57+1.16*3}}{e^{-4.57+1.16*1}} = \frac{0.3362}{0.03304} = 10.1755$

$DL50^*$ = Dose letal 50