

BIOESTATÍSTICA

M.I. Eng. Biomédica

2015-2016

Aula Teórica 12

Regressão logística

- Avaliação do ajuste
 - Teste de Hosmer-Lemeshow

$$H = \sum_{g=1}^G \frac{(O_g - E_g)^2}{N_g \pi_g (1 - \pi_g)}.$$

em que O indica observações, E valores esperados e o índice g refere-se ao grupo no qual é determinado o risco. A estatística teste segue uma distribuição qui-quadrado com $G-2$ graus de liberdade.

Regressão logística

- Pseudo R^2
 - Cox – Snell

$$R^2 = 1 - \left(\frac{-2LL_{null}}{-2LL_k} \right)^{\frac{2}{n}}$$

- Nagelkerke

$$R^2 = \frac{1 - \left(\frac{-2LL_{null}}{-2LL_k} \right)^{\frac{2}{n}}}{1 - (-2LL_{null})^{2/n}}$$

Regressão logística

- Significância das variáveis independentes
 - Teste de Wald

$$\frac{(\hat{\theta} - \theta_0)^2}{\text{var}(\hat{\theta})}$$

$$\frac{\hat{\theta} - \theta_0}{\text{se}(\hat{\theta})}$$

- o parâmetro estimado é comparado com o parâmetro de modelo nulo, pelo que a hipótese nula é que os dois são iguais entre si o que significa que o parâmetro estimado não tem significado estatístico.
- a estatística de teste pode ser calculada tendo em conta uma distribuição qui-quadrado(à esquerda) ou normal (à direita).



Aprendizagem supervisionada

- Grupo de treino
 - ajustamento do modelo estatístico
- Grupo de validação cruzada
 - optimização do modelo (sobre/sub ajustamento)
 - generalização
- Grupo de teste
 - teste e validação do modelo
 - generalização

Aprendizagem supervisionada

- Curva de aprendizagem

