

BIOESTATÍSTICA

M.I. Eng. Biomédica

2015-2016

Aula Teórica 11



Regressão logística

- A regressão logística é uma técnica de regressão indicada para a quando a variável dependente é dicotômica.
 - As variáveis independentes podem ser qualitativas e/ou quantitativas;
 - O modelo logístico permite avaliar a significância de cada uma das variáveis independentes;

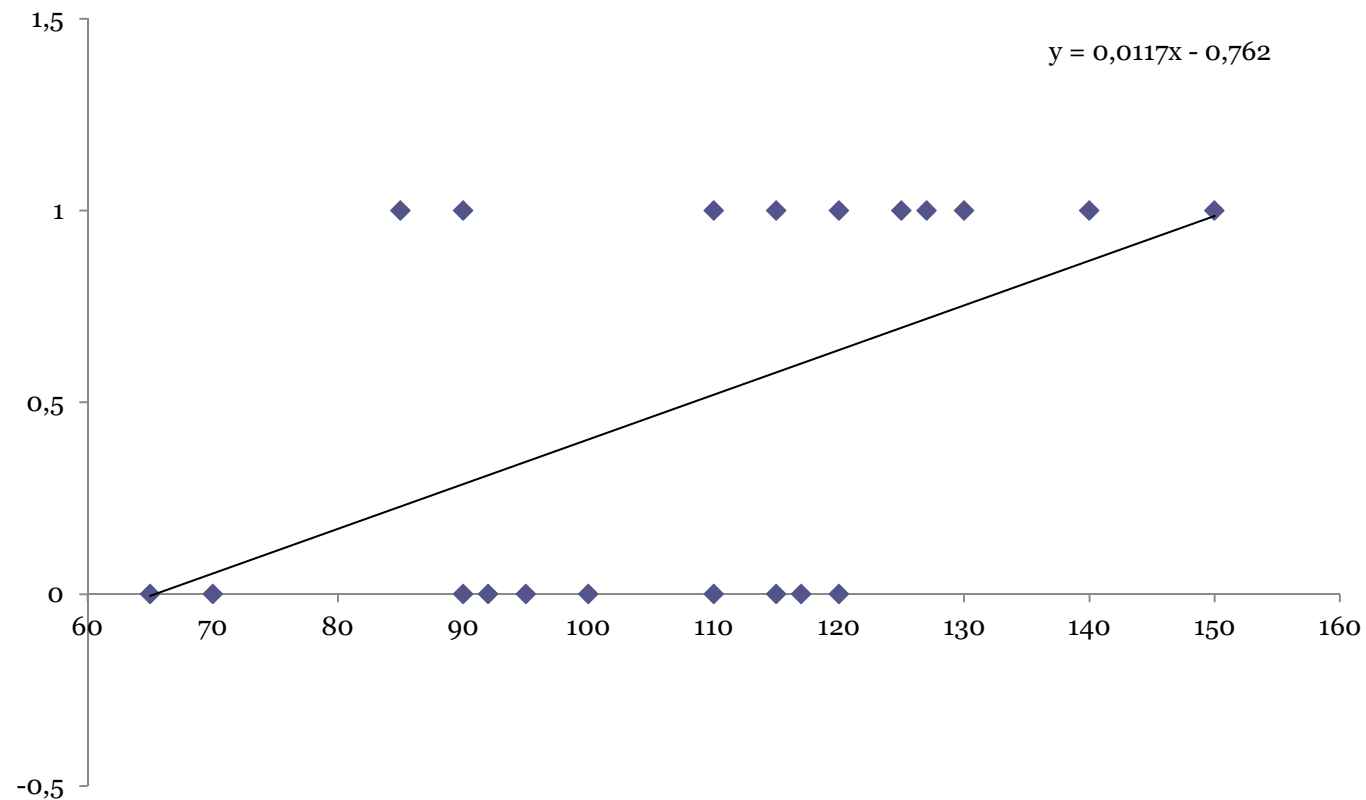
Regressão logística

Sexo	Idade	Controlo velocidade (km/h)	Taxa de alcoolémia (g/l)	Tempo de carta (anos)	Teve algum acidente sério
F	27	90	0,00	< 5	Não
M	69	125	0,10	> 20	Sim
M	27	80	0,20	< 5	Não
M	48	114	0,10	5-20	Sim
F	60	112	0,10	> 20	Não
F	32	100	0,00	< 5	Não
F	37	95	0,10	5-20	Não

adaptado de:
Análise Estatística com utilização do SPSS
João Maroco

Regressão logística

- Acidente *vs.* velocidade

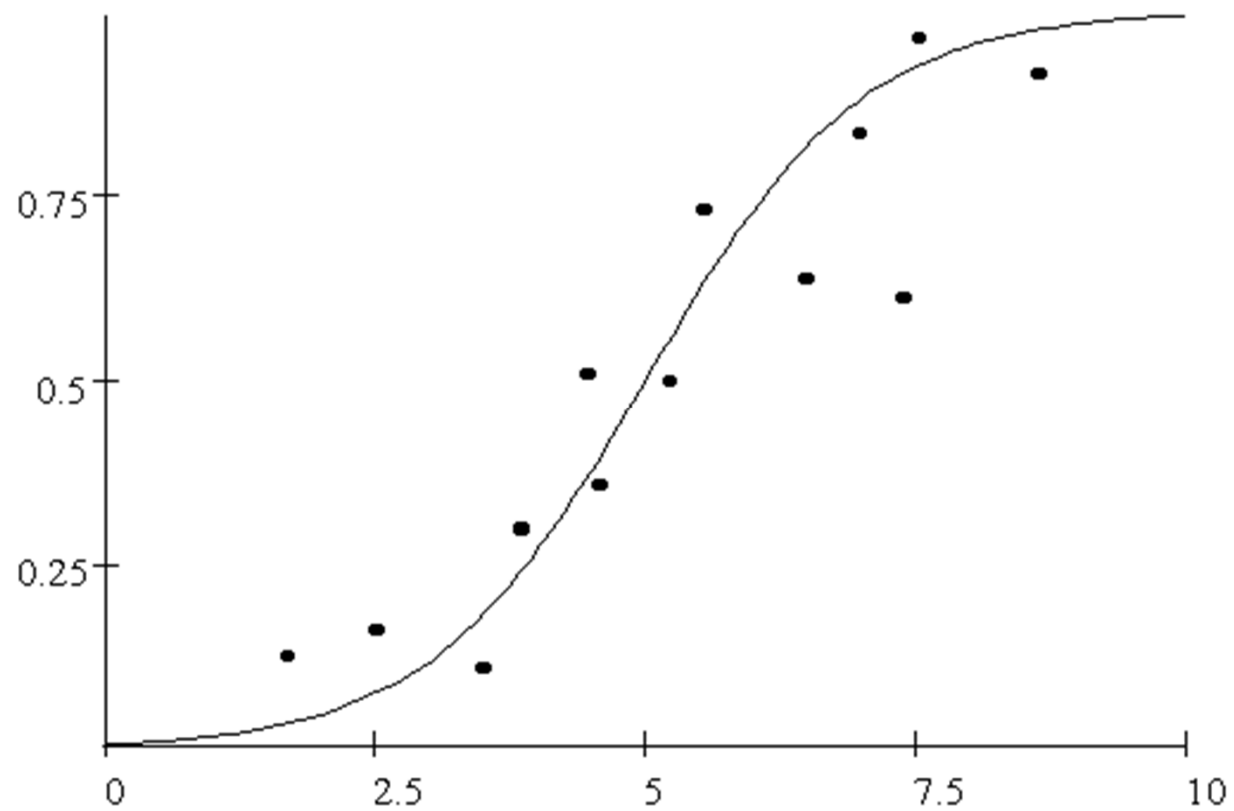




Regressão logística

- A variabilidade da variável dependente para os valores da variável independentes tornam difícil perceber a relação;
 - Dividir a velocidade em classes mutuamente exclusivas;
 - Determinar a probabilidade de ocorrência em cada uma das classes;
 - Representar graficamente a probabilidade em função das classes;

Regressão logística



Regressão logística

- A curva em 'S' pode ser modelada de diversas formas.
- A forma genérica usada é

$$\hat{\pi}_j = \frac{e^{\beta_0 + \beta_1 X_j}}{1 + e^{\beta_0 + \beta_1 X_j}}$$

Regressão logística

- Para mais do que uma variável independente

$$\hat{\pi}_j = \frac{e^{\beta_0 + \beta_1 X_{1j} + \dots + \beta_k X_{kj}}}{1 + e^{\beta_0 + \beta_1 X_{1j} + \dots + \beta_k X_{kj}}}$$

Regressão logística

- Geralmente lineariza-se para o ajuste

$$\textit{Logit}(\hat{\pi}_j) = \ln \left(\frac{\hat{\pi}_j}{1 - \hat{\pi}_j} \right)$$

Regressão logística

- Pressupostos

- Linearidade e aditividade: a escala de Logit é aditiva e linear
- Proporcionalidade: a contribuição de cada X_i é proporcional ao seu valor β ;
- Constância de efeito: a contribuição de uma variável independente é constante e independente das outras variáveis;
- Os erros são independentes e apresentam distribuição binomial;
- Os predictores não são multicolineares;