

Estatística e Análise de Dados em Zootécnia

FORMULÁRIO PARTE I

Descrição	Fórmula
Transformações Linearizantes	
Modelo Exponencial, equação $y = ce^{dx}$	$y^* = \ln(y), x^* = x$
Modelo Potência, equação $y = cx^d$	$y^* = \ln(y), x^* = \ln(x)$
Modelo Logístico, equação $y = \frac{1}{1+e^{-(c+dx)}}$	$y^* = \ln\left(\frac{y}{1-y}\right), x^* = x$
Propriedades de Matrizes	$(\mathbf{A} \pm \mathbf{B})^t = \mathbf{A}^t \pm \mathbf{B}^t \quad (\mathbf{AB})^t = \mathbf{B}^t \mathbf{A}^t \quad (\mathbf{A}^t)^{-1} = (\mathbf{A}^{-1})^t$
Propriedades da Multinormal	Se $\vec{\mathbf{W}} \cap \mathcal{N}_n(\vec{\boldsymbol{\mu}}, \boldsymbol{\Sigma})$, $\vec{\mathbf{a}}$ vector $k \times 1$ (não aleatório) e $\mathbf{C}$ matriz $k \times n$ (não aleatória, de característica k) então $\mathbf{C}\vec{\mathbf{W}} + \vec{\mathbf{a}} \cap \mathcal{N}_k(\mathbf{C}\vec{\boldsymbol{\mu}} + \vec{\mathbf{a}}, \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^t)$
MODELO LINEAR-Regressão, p preditores	
Equação do modelo	$\vec{\mathbf{Y}} = \mathbf{X}\vec{\boldsymbol{\beta}} + \vec{\boldsymbol{\epsilon}}$
Vector dos estimadores dos parâmetros	$\vec{\hat{\boldsymbol{\beta}}} = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \vec{\mathbf{Y}}$ Para $p = 1$: $\hat{\beta}_1 = \frac{Cov_{xy}}{s_x^2} = \sum_{i=1}^n c_i Y_i$, com $c_i = \frac{x_i - \bar{x}}{(n-1)s_x^2}$ $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x} = \sum_{i=1}^n d_i Y_i$, com $d_i = \frac{1}{n} - \frac{(x_i - \bar{x})\bar{x}}{(n-1)s_x^2}$
Matriz de projecção ortogonal	$\mathbf{H} = \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t$
Vector dos valores estimados de Y	$\vec{\hat{\mathbf{Y}}} = \mathbf{X}\vec{\hat{\boldsymbol{\beta}}} = \mathbf{H}\vec{\mathbf{Y}}$
Matriz de (co-)variâncias dos estimadores $\hat{\beta}_i$	$Var(\vec{\hat{\boldsymbol{\beta}}}) = \sigma^2 \cdot (\mathbf{X}^t \mathbf{X})^{-1}$ Para $p = 1$: $Var(\hat{\beta}_1) = \frac{\sigma^2}{(n-1)s_x^2}$, $Var(\hat{\beta}_0) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{(n-1)s_x^2} \right)$ $Cov(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\bar{x}\sigma^2}{(n-1)s_x^2}$
IC a $(1 - \alpha) \times 100\%$ para combinações lineares dos parâmetros: $\vec{\mathbf{a}}^t \vec{\boldsymbol{\beta}} = \sum_{i=0}^p a_i \beta_i$: Para $p = 1$, Intervalo de predição a $(1 - \alpha) \times 100\%$ para observação individual de Y , dado $\mathbf{X} = \mathbf{x}$:	$\left[\vec{\mathbf{a}}^t \vec{\mathbf{b}} - t_{\frac{\alpha}{2}; n-(p+1)} \cdot \hat{\sigma}_{\vec{\mathbf{a}}^t \vec{\boldsymbol{\beta}}}, \vec{\mathbf{a}}^t \vec{\mathbf{b}} + t_{\frac{\alpha}{2}; n-(p+1)} \cdot \hat{\sigma}_{\vec{\mathbf{a}}^t \vec{\boldsymbol{\beta}}} \right]$ com $\hat{\sigma}_{\vec{\mathbf{a}}^t \vec{\boldsymbol{\beta}}} = \sqrt{QMRE \cdot \vec{\mathbf{a}}^t (\mathbf{X}^t \mathbf{X})^{-1} \vec{\mathbf{a}}}$ $\left[(b_0 + b_1 x) - t_{\frac{\alpha}{2}; n-2} \cdot \sqrt{QMRE \cdot \left[1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{(n-1)s_x^2} \right]}, \right. \\ \left. (b_0 + b_1 x) + t_{\frac{\alpha}{2}; n-2} \cdot \sqrt{QMRE \cdot \left[1 + \frac{1}{n} + \frac{(x - \bar{x})^2}{(n-1)s_x^2} \right]} \right]$
Estatística do Teste de Ajustamento Global	$F = \frac{QMR}{QMRE} = \frac{n-(p+1)}{p} \cdot \frac{R^2}{1-R^2}$
Estatística Teste aos Modelos Encaixados (modelo completo p preditores, submodelo k preditores)	$F = \frac{(SQRE_s - SQRE_c)/(p-k)}{(SQRE_c)/(n-(p+1))} = \frac{n-(p+1)}{p-k} \cdot \frac{R_C^2 - R_s^2}{1-R_C^2}$
AIC (Critério de Informação de Akaike)	$AIC = n \ln\left(\frac{SQRE_p}{n}\right) + 2(p+1)$.
Distribuição dos resíduos Para $p = 1$, valor do efeito alavanca Resíduos (internamente) estandardizados Distância de Cook R^2 modificado	$E_i \cap \mathcal{N}(0, \sigma^2 \cdot (1 - h_{ii}))$ com $h_{ii} = \mathbf{H}_{(i,i)}$ (efeito alavanca) $h_{ii} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{(n-1)s_x^2}$ $R_i = \frac{E_i}{\sqrt{QMRE \cdot (1 - h_{ii})}}$ $D_i = R_i^2 \cdot \left(\frac{h_{ii}}{1 - h_{ii}} \right) \cdot \frac{1}{p+1}$ $R_{mod}^2 = 1 - \frac{QMRE}{QMT} = 1 - (1 - R^2) \cdot \frac{n-1}{n-(p+1)}$