

“ligadas” por $g(\cdot)$ na forma:

$$g(E[Y]) = X\beta \Leftrightarrow g(\mu_i) = x_i^t \beta = \sum_{j=0}^p \beta_j x_{ij} \quad (i = 1, \dots, n) \quad (1)$$

sendo $\mu = E[Y]$ o vector dos valores esperados das n observações de Y . Caso g admita função inversa, resulta:

$$E[Y] = g^{-1}(X\beta) \Leftrightarrow \mu_i = g^{-1}(x_i^t \beta) = g^{-1}\left(\sum_{j=0}^p \beta_j x_{ij}\right) \quad (i = 1, \dots, n). \quad (2)$$

A Tabela 2 contém as funções de ligação utilizadas no presente trabalho. A função de ligação Aranda-Ordaz (1981), para determinados valores atribuídos ao parâmetro α , poderá constituir um caso particular das funções de ligação Logit e Log-Log do Complementar. Assim:

(i) quando $\alpha \rightarrow 0$ tem-se a função ligação Log-Log do Complementar. Recorrendo ao limite notável:

$$\lim_{x \rightarrow 0} \left(\frac{a^x - 1}{x} \right) = \log(a) \quad (\text{se } a > 0). \quad (3)$$

Considerando $x = \alpha$; $a = (1 - \mu)^{-1}$ na expressão (3) e aplicando ao cálculo de $\lim_{\alpha \rightarrow 0} g^{-1}(\mu)$ teremos:

$$\lim_{\alpha \rightarrow 0} \log \left\{ \frac{(1 - \mu)^{-\alpha} - 1}{\alpha} \right\} = \log(-\log(1 - \mu)), \quad 0 < \mu < 1.$$

(ii) quando $\alpha = 1$ tem-se a função de ligação Logit

$$\log \left\{ \frac{(1 - \mu)^{-1} - 1}{1} \right\} = \log \left\{ \frac{1}{1 - \mu} - 1 \right\} = \log \left\{ \frac{1 - (1 - \mu)}{1 - \mu} \right\} = \log \left\{ \frac{\mu}{1 - \mu} \right\}. \quad (4)$$

A Figura 2 representa graficamente as curvas estimadas pela função de ligação Aranda-Ordaz para diferentes valores de α .

O estudo tem como base o modelo de Regressão Logística (RL). A RL irá procurar determinar um modelo com interpretação biológica, bem ajustado e parcimonioso que permita descrever com sensatez a relação entre a variável resposta ICC e um conjunto de covariáveis. A secção que se segue destina-se a explicar a metodologia subjacente à escolha das variáveis predictoras a figurarem no modelo de RL.

2.3 Selecção das covariáveis

Muitos são os casos onde métodos sequenciais na determinação de modelos se revelam ferramentas úteis pois permitem o “rastreo” das variáveis candidatas a serem incluídas/excluídas no modelo de uma forma célere. Apesar deste atributo,