

**Ajustes para o teste  
da razão de verossimilhanças  
em modelos de regressão beta**

Eliane Cantinho Pinheiro

DISSERTAÇÃO APRESENTADA  
AO  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DA  
UNIVERSIDADE DE SÃO PAULO  
PARA  
OBTENÇÃO DO TÍTULO  
DE  
MESTRE EM CIÊNCIAS

Programa: Estatística

Orientador: Prof<sup>a</sup>. Dr<sup>a</sup>. Silvia Lopes de Paula Ferrari

Durante o desenvolvimento deste trabalho a autora recebeu auxílio financeiro da CAPES e do CNPQ.

São Paulo, 7 de julho de 2009

**Ajustes para o teste  
da razão de verossimilhanças  
em modelos de regressão beta**

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Eliane Cantinho Pinheiro e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Profa. Dra. Silvia Lopes de Paula Ferrari (Orientadora) - IME/USP.
- Profa. Dra. Denise Aparecida Botter - IME/USP.
- Prof. Dr. Raydonal Ospina Martínez - CCEN/UFPE.

*"O que é impossível aos homens, é possível a Deus."*

Jesus  
Lc 18,27

*"Tudo o que fizeres, faze-o na medida do teu talento, e terás louvado a Deus."*

Sto Agostinho

*"Eu não vejo graça em outras coisas como eu vejo em cantar."*

Elis Regina

*Dedico este trabalho a todos os que me instruíram na Verdade, a todos os que me indicaram boas leituras, a todos os que me incentivaram, a todos os que me fizeram crescer e a todos os que tornaram meu caminho mais suave.*



## **Agradecimentos**

A Deus pelo dom da vida e pelo dom de amá-la.

Ao meu marido pelo incentivo e parceria.

Aos meus filhos por alegrarem minha vida.

Aos meus pais por sua dedicação.

À minha sogra por seu exemplo de fé.

Ao meu sogro por seu exemplo de estudante perene.

À professora Silvia pelo entusiasmo contagiante e disponibilidade.

À CAPES e ao CNPQ pelo apoio financeiro.

Aos participantes da banca examinadora.



## Resumo

O presente trabalho considera o problema de fazer inferência com acurácia para pequenas amostras, tomando por base a estatística da razão de verossimilhanças em modelos de regressão beta. Estes, por sua vez, são úteis para modelar proporções contínuas que são afetadas por variáveis independentes. Deduzem-se as estatísticas da razão de verossimilhanças ajustadas de Skovgaard (*Scandinavian Journal of Statistics* 28 (2001) 3-32) nesta classe de modelos. Os termos do ajuste, que têm uma forma simples e compacta, podem ser implementados em um software estatístico. São feitas simulações de Monte Carlo para mostrar que a inferência baseada nas estatísticas ajustadas propostas é mais confiável do que a inferência usual baseada na estatística da razão de verossimilhanças. Aplicam-se os resultados a um conjunto real de dados.

**Palavras-chave:** Regressão beta; Proporções contínuas; Modelos não lineares; Teste da razão de verossimilhanças; Ajustes para pequenas amostras.





## Abstract

We consider the issue of performing accurate small-sample likelihood-based inference in beta regression models, which are useful for modeling continuous proportions that are affected by independent variables. We derive Skovgaard's (*Scandinavian Journal of Statistics* 28 (2001) 3-32) adjusted likelihood ratio statistics in this class of models. We show that the adjustment terms have simple compact form that can be easily implemented from standard statistical software. We present Monte Carlo simulations showing that inference based on the adjusted statistics we propose is more reliable than that based on the usual likelihood ratio statistic. A real data example is presented.

**Keywords:** Beta regression; Continuous proportions; Likelihood ratio test; Nonlinear models; Small-sample adjustments.



# Sumário

|  |             |
|--|-------------|
| <b>Lista de Figuras</b>  | <b>xi</b>   |
| <b>Lista de Tabelas</b>  | <b>xiii</b> |
| <b>1 Introdução</b>  | <b>1</b>    |
| 1.1 Considerações preliminares . . . . .                       | 1           |
| 1.2 Objetivos . . . . .  | 3           |
| 1.3 Organização do trabalho . . . . .                          | 3           |
| <b>2 Modelo de regressão beta</b>                              | <b>5</b>    |
| 2.1 Modelo de regressão beta linear . . . . .                  | 5           |
| 2.2 Modelo de regressão beta não linear . . . . .              | 8           |
| <b>3 Ajuste de Skovgaard</b>                                   | <b>11</b>   |
| <b>4 Simulações</b>  | <b>15</b>   |
| 4.1 Modelo de regressão beta com precisão constante . . . . .  | 15          |
| 4.2 Modelo de regressão beta com precisão variável . . . . .   | 19          |
| <b>5 Aplicação</b>   | <b>23</b>   |
| <b>6 Considerações finais</b>                                  | <b>27</b>   |
| <b>A Cálculo da função escore e das matrizes de informação</b> | <b>29</b>   |
| <b>B Cálculo de <math>q</math> e <math>\Upsilon</math></b>     | <b>37</b>   |
| <b>C Programas</b>   | <b>41</b>   |



## Lista de Figuras

|     |  |    |
|-----|--|----|
| 1.1 | Densidades beta para diferentes combinações de $(\mu, \phi)$ . . . . . | 2  |
| 4.1 | Gráficos quantil-quantil, modelo 1, $\phi = 30$ . . . . .              | 18 |
| 4.2 | Gráficos quantil-quantil, modelo 2, $\gamma_1 = \log 30$ . . . . .     | 20 |
| 4.3 | Gráficos quantil-quantil, modelo 3. . . . .                            | 22 |



## Lista de Tabelas

|     |  |    |
|-----|--|----|
| 4.1 | Taxas de rejeição nula (%); modelo 1 . . . . .                                 | 16 |
| 4.2 | Taxas de rejeição não nula (%); modelo 1, $n = 20$ , $\alpha = 10\%$ . . . . . | 17 |
| 4.3 | Taxas de rejeição nula (%); modelo 2 . . . . .                                 | 19 |
| 4.4 | Taxas de rejeição não nula (%); modelo 2, $n = 40$ , $\alpha = 10\%$ . . . . . | 20 |
| 4.5 | Taxas de rejeição nula; modelo 3 . . . . .                                     | 21 |
| 4.6 | Taxas de rejeição não nula (%); modelo 3, $n = 40$ , $\alpha = 10\%$ . . . . . | 22 |
| 5.1 | Dados do estudo de Oxidação de Amônia. . . . .                                 | 24 |
| 5.2 | Estimativas dos parâmetros e erros padrão : modelo I . . . . .                 | 24 |
| 5.3 | Estimativas dos parâmetros e erros padrão : modelo II . . . . .                | 25 |
| 5.4 | Estimativas dos parâmetros e erros padrão : modelo III . . . . .               | 26 |





# Capítulo 1

## Introdução

### 1.1 Considerações preliminares

Modelos normais de regressão podem não ser apropriados quando a variável resposta assume valores pertencentes a um intervalo limitado da reta, tais como proporções contínuas. A família beta é uma alternativa à distribuição normal, sendo flexível para modelar dados cujos valores estejam restritos ao intervalo unitário padrão. Ferrari e Cribari-Neto (2004) propuseram um modelo de regressão baseado na suposição de que a resposta tem distribuição beta. Eles usaram uma parametrização adequada, baseada na média e em um parâmetro de precisão. A Figura 1.1 mostra a função densidade de probabilidade da distribuição beta para vários valores dos parâmetros  $\mu$  e  $\phi$ . A estrutura do modelo proposto é similar à estrutura de modelos lineares generalizados (McCullagh e Nelder, 1989).

Em Ferrari e Cribari-Neto (2004) e Espinheira, Ferrari e Cribari-Neto (2008a, 2008b) encontra-se estudo acerca de inferência e diagnóstico no modelo de regressão beta com precisão constante. Nesta dissertação consideramos generalizações deste modelo, permitindo que a precisão varie nas observações, de acordo com Smithson e Verkuilen (2006), e permitindo que as funções de regressão sejam não lineares.

A inferência em modelo de regressão beta para grandes amostras está baseada em teoria assintótica. Contudo, a inferência pode ser incorreta se a amostra é pequena. Por exemplo, o teste da razão de verossimilhanças assintótico pode ser bastante distorcido se a amostra não for suficientemente grande. Com o objetivo de melhorar a aproximação da distribuição da estatística do teste pela distribuição de referência, uma distribuição  $\chi^2$ , a estratégia usual é substituir a estatística da razão de verossimilhanças por sua versão corrigida de Bartlett (Lawley, 1956). Se a estatística de escore é usada, uma estratégia similar pode ser utilizada, isto é, a estatística de escore pode ser substituída por sua versão corrigida de tipo-Bartlett (Cordeiro e Ferrari, 1991). As correções de Bartlett e as de tipo-Bartlett dependem de alguns cumulantes de derivadas da log-verossimilhança e sua obtenção, para a maioria dos modelos, é trabalhosa. Um caso mais simples é quando o parâmetro de precisão é constante e globalmente ortogonal aos parâmetros da regressão como em modelos lineares generalizados; ver Cordeiro (1987) e Cribari-Neto e Ferrari (1995). Se o parâmetro de precisão é modelado, mas os parâmetros de regressão são globalmente ortogonais aos parâmetros restantes, a dedução das correções de Bartlett e tipo-Bartlett é factível. No modelo de regressão beta, no entanto, os parâmetros de regressão não são globalmente ortogonais aos parâmetros de precisão, isto é, a matriz de informação de Fisher não é diagonal em blocos. Por esta razão busca-se, neste trabalho, um ajuste para a estatística da

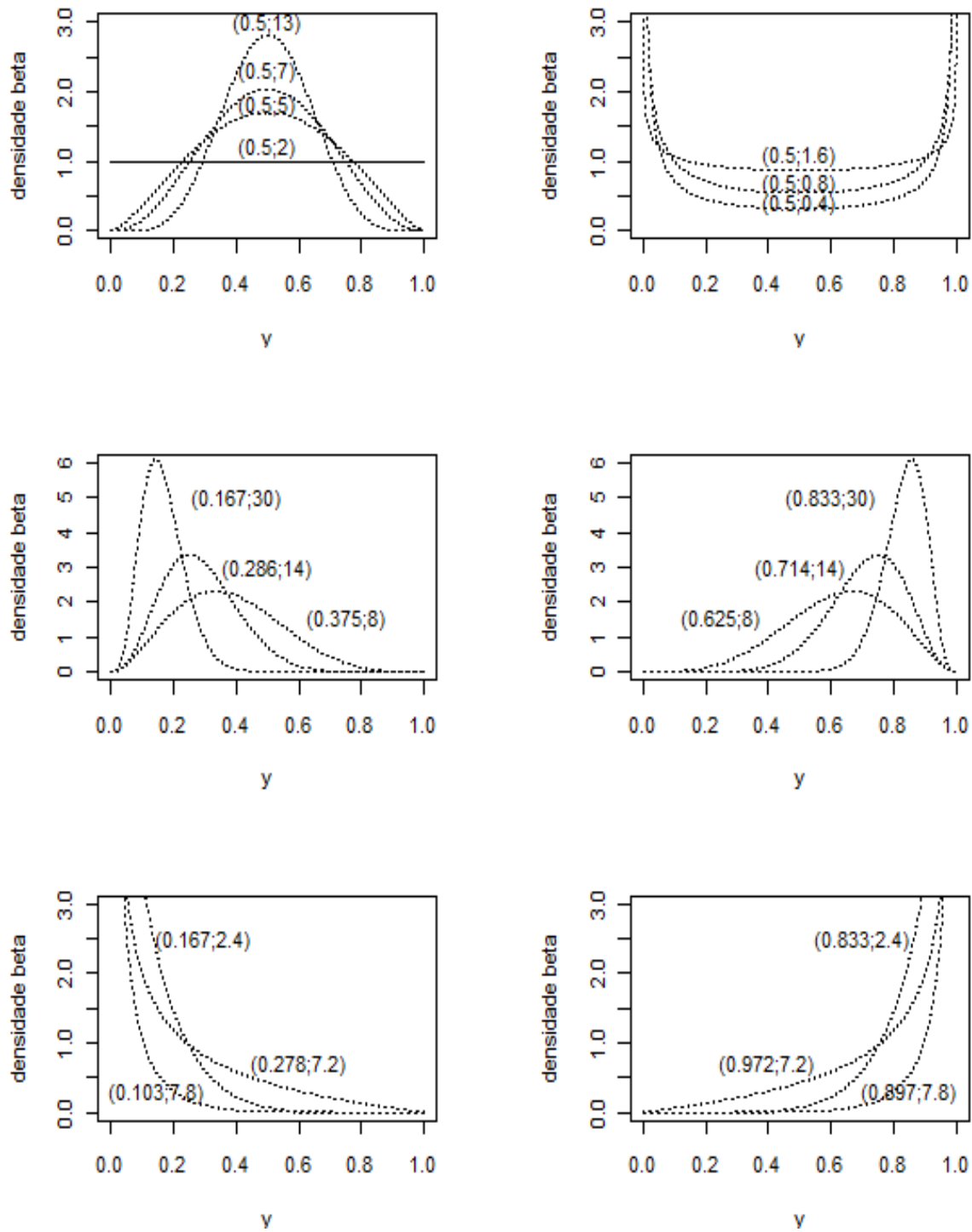


Figura 1.1: Densidades beta para diferentes combinações de  $(\mu, \phi)$ .

razão de verossimilhanças, que seja mais facilmente calculado no modelo de regressão beta.

## 1.2 Objetivos

Neste trabalho, segue-se Skovgaard (2001) para obter estatísticas da razão de verossimilhanças ajustadas. A dedução é bem mais simples que a da correção de Bartlett e as estatísticas ajustadas são aproximadamente distribuídas como  $\chi^2$  com alto grau de acurácia. Resultados empíricos mostram que os testes baseados nas estatísticas ajustadas de Skovgaard têm desempenho muito melhor que a estatística da razão de verossimilhanças em amostras pequenas; ver Ferrari e Cysneiros (2009).

## 1.3 Organização do trabalho

A dissertação se desenvolve como segue. No Capítulo 2 apresentam-se o modelo de regressão beta, a função escore e a matriz de informação. No Capítulo 3, obtêm-se um teste da razão de verossimilhanças ajustada baseado no trabalho de Skovgaard (2001) para o modelo de regressão beta com precisão variável e para um modelo mais geral em que se admite não linearidade na especificação tanto de média quanto de precisão. No Capítulo 4 apresentam-se alguns resultados, obtidos com simulação de Monte Carlo, comparando a performance, para pequenas amostras, de diferentes testes baseados em verossimilhança. No Capítulo 5, apresenta-se uma aplicação para um conjunto real de dados. Algumas conclusões são apresentadas no Capítulo 6. Detalhes técnicos e programas computacionais são apresentados nos apêndices.



## Capítulo 2

# Modelo de regressão beta

### 2.1 Modelo de regressão beta linear

Sejam  $y_1, \dots, y_n$  variáveis aleatórias independentes, em que cada  $y_t$ ,  $t = 1, \dots, n$ , tem a função densidade

$$f(y_t; \mu_t, \phi_t) = \frac{\Gamma(\phi_t)}{\Gamma(\mu_t\phi_t)\Gamma((1-\mu_t)\phi_t)} y_t^{\mu_t\phi_t-1} (1-y_t)^{(1-\mu_t)\phi_t-1}, \quad 0 < y_t < 1, \quad (2.1)$$

onde  $\Gamma(\cdot)$  é a função gama,  $0 < \mu_t < 1$  e  $\phi_t > 0$ . Aqui,  $E(y_t) = \mu_t$  e  $\text{var}(y_t) = \mu_t(1-\mu_t)/(1+\phi_t)$ . Portanto,  $\mu_t$  é a média de  $y_t$  e  $\phi_t$  pode ser visto como um parâmetro de precisão. Analogamente,  $1/(1+\phi_t)$  pode ser visto como um parâmetro de dispersão. O modelo proposto por Ferrari e Cribari Neto (2004) assume  $\phi_t = \phi$ , para  $t = 1, \dots, n$  (precisão constante) e que  $\mu_t$  pode ser escrito como

$$g(\mu_t) = \eta_t = \sum_{i=1}^k x_{ti}\beta_i, \quad (2.2)$$

onde  $\beta = (\beta_1, \dots, \beta_k)^\top$  é um vetor de parâmetros de regressão desconhecidos ( $\beta \in \mathbb{R}^k$ ) e  $x_{t1}, \dots, x_{tk}$  são observações em  $k$  covariadas ( $k < n$ ), as quais são assumidas como fixas e conhecidas. Finalmente,  $g(\cdot)$  é uma função de ligação estritamente monótona e duas vezes diferenciável que transforma  $(0, 1)$  em  $\mathbb{R}$ . Entre as possíveis funções de ligação da média estão a especificação logito  $g(\mu) = \log\{\mu/(1-\mu)\}$ , a função probito  $g(\mu) = \Phi^{-1}(\mu)$ , onde  $\Phi(\cdot)$  é a função de distribuição acumulada de uma variável aleatória normal padrão, a função de ligação log-log complementar  $g(\mu) = \log\{-\log(1-\mu)\}$ , a função de ligação log-log  $g(\mu) = -\log\{-\log(\mu)\}$ , entre outras.

Nesta dissertação trabalha-se com generalizações deste modelo. Admite-se inicialmente que a precisão varia nas observações de acordo com a seguinte estrutura:

$$h(\phi_t) = \delta_t = \sum_{i=1}^m z_{ti}\gamma_i, \quad (2.3)$$

onde  $\gamma = (\gamma_1, \dots, \gamma_m)^\top$  é um vetor de parâmetros de regressão desconhecidos ( $\gamma \in \mathbb{R}^m$ ) e  $z_{t1}, \dots, z_{tm}$  são observações em  $m$  covariadas fixas ( $m < n$ ); ver Smithson e Verkuilen (2006). Aqui,  $h(\cdot)$  é uma função de ligação estritamente monótona e duas vezes diferenciável com domínio  $(0, +\infty)$  e imagem  $\mathbb{R}$ . Uma possível escolha para a função de ligação para a precisão é  $h(\phi) = \log(\phi)$ . Uma generalização mais ampla do modelo (2.1)-(2.3) é obtida permitindo que  $\eta_t$  e  $\delta_t$  sejam funções não lineares de  $\beta$  e  $\gamma$  respectivamente.

Esse modelo mais geral também será tratado nessa dissertação.

Seja  $\omega = (\beta^\top, \gamma^\top)^\top$  o vetor de parâmetros desconhecidos do modelo de regressão (2.1)-(2.3). No que segue,  $\nu = (\nu_1, \dots, \nu_r)^\top$  representa o parâmetro de interesse e  $\psi = (\psi_1, \dots, \psi_s)^\top$  é o parâmetro de perturbação (note que  $r + s = k + m$ ). Consideraremos no Capítulo 3 testes baseados na função de verossimilhança, da hipótese nula  $\mathcal{H}_0 : \nu = \nu_0$  onde  $\nu_0$  é um vetor fixo de dimensão  $r$ . Tais testes podem ser invertidos para obter intervalos de confiança para  $\nu$ .

Seja  $\ell(\omega)$  a função de log-verossimilhança do modelo definido por (2.1)-(2.3) dado o vetor de observações  $y = (y_1, \dots, y_n)$ . Podemos escrever o logaritmo da função de verossimilhança como:

$$\ell(\omega) = \sum_{t=1}^n \ell_t(\mu_t, \phi_t), \quad (2.4)$$

onde

$$\ell_t(\mu_t, \phi_t) = \log \Gamma(\phi_t) - \log \Gamma(\mu_t \phi_t) - \log \Gamma((1 - \mu_t) \phi_t) + (\mu_t \phi_t - 1) y_t^* + (\phi_t - 2) y_t^\dagger, \quad (2.5)$$

com  $\mu_t$  e  $\phi_t$  definidos de tal forma que (2.2) e (2.3) sejam válidos e

$$y_t^* = \log \frac{y_t}{1 - y_t}$$

e

$$y_t^\dagger = \log(1 - y_t).$$

É fácil obter os momentos de  $y_t^*$  e  $y_t^\dagger$ . Temos que

$$\begin{aligned} \mu_t^* &= E_\omega(y_t^*) = \psi(\mu_t \phi_t) - \psi((1 - \mu_t) \phi_t), \\ \mu_t^\dagger &= E_\omega(y_t^\dagger) = \psi((1 - \mu_t) \phi_t) - \psi(\phi_t), \\ v_t^* &= \text{var}_\omega(y_t^*) = \psi'(\mu_t \phi_t) + \psi'((1 - \mu_t) \phi_t), \\ v_t^\dagger &= \text{var}_\omega(y_t^\dagger) = \psi'((1 - \mu_t) \phi_t) - \psi'(\phi_t), \\ c_t^{*\dagger} &= \text{cov}_\omega(y_t^*, y_t^\dagger) = -\psi'((1 - \mu_t) \phi_t), \end{aligned}$$

onde  $\psi(\cdot) = \Gamma'(\cdot)/\Gamma(\cdot)$  é a função digamma. Portanto,  $\ell(\omega)$  pode ser escrito como

$$\ell(\omega) = [(y^* - \mu^*)^\top (\Phi \mathcal{M} - \mathcal{I}) + (y^\dagger - \mu^\dagger)^\top (\Phi - 2\mathcal{I}) + a^\top] \iota \quad (2.6)$$

onde  $a = (a_1, \dots, a_n)^\top$  com

$$a_t = \log \Gamma(\phi_t) - \log \Gamma(\mu_t \phi_t) - \log \Gamma((1 - \mu_t) \phi_t) + \mu_t^* (\mu_t \phi_t - 1) + \mu_t^\dagger (\phi_t - 2),$$

$y^* = (y_1^*, \dots, y_n^*)^\top$ ,  $\mu^* = (\mu_1^*, \dots, \mu_n^*)^\top$ ,  $y^\dagger = (y_1^\dagger, \dots, y_n^\dagger)^\top$ ,  $\mu^\dagger = (\mu_1^\dagger, \dots, \mu_n^\dagger)^\top$ ,  $\Phi = \text{diag}(\phi_1, \dots, \phi_n)$ ,  $\mathcal{I}$  é a matriz identidade  $n \times n$ ,  $\mathcal{M} = \text{diag}\{\mu_1, \dots, \mu_n\}$  e  $\iota$  é um vetor coluna de uns de dimensão  $n$  (ver Apêndice A).

A função escore, obtida diferenciando-se o logaritmo da função de verossimilhança com respeito aos parâmetros desconhecidos, é dada por  $U = (U_\beta(\beta, \gamma)^\top, U_\gamma(\beta, \gamma)^\top)^\top$ , em que

$$U_\beta(\beta, \gamma) = X^\top \Phi T (y^* - \mu^*) \quad (2.7)$$

e

$$U_\gamma(\beta, \gamma) = Z^\top H [\mathcal{M}(y^* - \mu^*) + (y^\dagger - \mu^\dagger)] \quad (2.8)$$

onde  $X$  é uma matriz  $n \times k$  cuja  $t$ -ésima linha é  $x_t^\top$ ,  $Z$  é uma matriz  $n \times m$  cuja  $t$ -ésima linha é  $z_t^\top$ ,  $T = \text{diag}\{1/g'(\mu_1) \dots, 1/g'(\mu_n)\}$ ,  $H = \text{diag}\{1/h'(\phi_1) \dots, 1/h'(\phi_n)\}$  (ver Apêndice A).

A matriz de informação observada, segunda derivada do logaritmo da função de verossimilhança em relação aos parâmetros desconhecidos, é dada por

$$J = \begin{bmatrix} J_{\beta\beta} & J_{\beta\gamma} \\ J_{\gamma\beta} & J_{\gamma\gamma} \end{bmatrix}$$

com

$$J_{\beta\beta} = X^\top [\Phi T V^* + S T^2 (Y^* - \mathcal{M}^*)] T \Phi X,$$

$$J_{\beta\gamma} = J_{\gamma\beta}^\top = -X^\top [(Y^* - \mathcal{M}^*) - \Phi (\mathcal{M} V^* + C)] T H Z$$

e

$$J_{\gamma\gamma} = Z^\top [H (\mathcal{M}^2 V^* + 2\mathcal{M}C + V^\dagger) + [\mathcal{M}(Y^* - \mathcal{M}^*) + (Y^\dagger - \mathcal{M}^\dagger)] Q H^2] H Z,$$

onde  $Y^* = \text{diag}\{y_1^*, \dots, y_n^*\}$ ,  $Y^\dagger = \text{diag}\{y_1^\dagger, \dots, y_n^\dagger\}$ ,  $\mathcal{M}^* = \text{diag}\{\mu_1^*, \dots, \mu_n^*\}$ ,  $\mathcal{M}^\dagger = \text{diag}\{\mu_1^\dagger, \dots, \mu_n^\dagger\}$ ,  $V^* = \text{diag}\{v_1^*, \dots, v_n^*\}$ ,  $V^\dagger = \text{diag}\{v_1^\dagger, \dots, v_n^\dagger\}$ ,  $C = \text{diag}\{c_1^{\dagger}, \dots, c_n^{\dagger}\}$ ,  $S = \text{diag}\{g''(\mu_1), \dots, g''(\mu_n)\}$  e  $Q = \text{diag}\{h''(\phi_1), \dots, h''(\phi_n)\}$  (ver Apêndice A).

A matriz de informação de Fisher, que é a esperança da matriz de informação observada, é

$$I = \begin{bmatrix} I_{\beta\beta} & I_{\beta\gamma} \\ I_{\gamma\beta} & I_{\gamma\gamma} \end{bmatrix}$$

com

$$I_{\beta\beta} = X^\top \Phi T V^* T \Phi X,$$

$$I_{\beta\gamma} = I_{\gamma\beta}^\top = X^\top \Phi (\mathcal{M} V^* + C) T H Z$$

e

$$I_{\gamma\gamma} = Z^\top H (\mathcal{M}^2 V^* + 2\mathcal{M}C + V^\dagger) H Z,$$

(ver Apêndice A).

## 2.2 Modelo de regressão beta não linear

Cosideramos, por fim, modelos de regressão beta não lineares, sendo as especificações para a média e o parâmetro de precisão dadas em (2.2) e (2.3) substituídas por

$$g(\mu_t) = \eta_t = \eta(x_t, \beta), \quad (2.9)$$

e

$$h(\phi_t) = \delta_t = \delta(z_t, \gamma). \quad (2.10)$$

Aqui,  $\eta(\cdot, \cdot)$  e  $\delta(\cdot, \cdot)$  podem ser funções não lineares no segundo argumento. Seja  $\mathcal{X}$  a matriz das derivadas de  $\eta = (\eta_1, \dots, \eta_n)^\top$  com respeito a  $\beta^\top$ . Analogamente, seja  $\mathcal{Z}$  a matriz das derivadas de  $\delta = (\delta_1, \dots, \delta_m)^\top$  em relação a  $\gamma^\top$ . No caso linear,  $\mathcal{X} = X$  e  $\mathcal{Z} = Z$ . Assume-se que  $\text{posto}(\mathcal{X}) = k$  e  $\text{posto}(\mathcal{Z}) = m$  para todo  $\beta$  e  $\gamma$ .

Para este modelo os componentes da função escore tornam-se

$$U_\beta(\beta, \gamma) = \mathcal{X}^\top \Phi T(y^* - \mu^*)$$

e

$$U_\gamma(\beta, \gamma) = \mathcal{Z}^\top H[\mathcal{M}(y^* - \mu^*) + (y^\dagger - \mu^\dagger)].$$

É fácil mostrar que as fórmulas para os componentes da matriz de informação observada tornam-se

$$J_{\beta\beta} = \mathcal{X}^\top \{\Phi T V^* + S T^2 (Y^* - \mathcal{M}^*)\} T \Phi \mathcal{X} + [(y^* - \mu^*)^\top \Phi T][\dot{\mathcal{X}}], \quad (2.11)$$

$$J_{\beta\gamma} = J_{\gamma\beta}^\top = -\mathcal{X}^\top \{(Y^* - \mathcal{M}^*) - \Phi(\mathcal{M} V^* + C)\} T H \mathcal{Z} \quad (2.12)$$

e

$$\begin{aligned} J_{\gamma\gamma} &= \mathcal{Z}^\top \{H(\mathcal{M}^2 V^* + 2\mathcal{M}C + V^\dagger) + \{\mathcal{M}(Y^* - \mathcal{M}^*) + (Y^\dagger - \mathcal{M}^\dagger)\} Q H^2\} H \mathcal{Z} \\ &+ [(y^* - \mu^*)^\top \mathcal{M} + (y^\dagger - \mu^\dagger)^\top] H [\dot{\mathcal{Z}}] \end{aligned} \quad (2.13)$$

com  $\dot{\mathcal{X}} = \partial^2 \eta / \partial \beta \partial \beta^\top$  e  $\dot{\mathcal{Z}} = \partial^2 \delta / \partial \gamma \partial \gamma^\top$  representando  $n \times k \times k$  e  $n \times m \times m$  arrays respectivamente. Além disso,  $[\cdot][\cdot]$  representa o produto colchete de uma matriz por um array, como definido por Wei (1998, p. 188). Mais especificamente, se  $A$  é uma matriz  $m \times n$  e  $B$  é um array  $n \times p \times q$ , então  $C = [A][B]$  é chamado o produto colchete de  $A$  e  $B$ , que é um array  $m \times p \times q$  com elementos  $C_{tij} = \sum_{k=1}^n A_{tk} B_{kij}$  (ver Apêndice A).

Como  $E(Y^* - \mathcal{M}^*) = 0$ ,  $E(Y^\dagger - \mathcal{M}^\dagger) = 0$ ,  $E(y^* - \mu^*) = 0$  e  $E(y^\dagger - \mu^\dagger) = 0$  concluímos a partir de (2.11), (2.12) e (2.13) que os componentes da matriz de informação de Fisher são dados por

$$I_{\beta\beta} = \mathcal{X}^\top \Phi T V^* T \Phi \mathcal{X},$$



$$I_{\beta\gamma} = \bar{I}_{\gamma\beta} = \mathcal{X}^\top \Phi(\mathcal{M}V^* + C)THZ$$

e

$$I_{\gamma\gamma} = Z^\top H(\mathcal{M}^2V^* + 2\mathcal{M}C + V^\dagger)HZ.$$



## Capítulo 3

### Ajuste de Skovgaard

A estatística da razão de verossimilhanças para testar  $\mathcal{H}_0 : \nu = \nu_0$  pode ser escrita como

$$w = 2\{\ell(\hat{\nu}, \hat{\psi}) - \ell(\nu_0, \tilde{\psi})\}.$$

Aqui,  $(\hat{\nu}^\top, \hat{\psi}^\top)$  é o estimador de máxima verossimilhança sem restrição de  $(\nu^\top, \psi^\top)$  e  $(\nu_0^\top, \tilde{\psi}^\top)$  é o estimador de máxima verossimilhança restrito de  $(\nu^\top, \psi^\top)$  sob  $\mathcal{H}_0$ . Em grandes amostras,  $w$  é aproximadamente distribuído como  $\chi_r^2$  sob  $\mathcal{H}_0$  com erro de ordem  $n^{-1}$ . Contudo, sabe-se que o teste da razão de verossimilhanças pode ter tamanho substancialmente distorcido se o tamanho da amostra não for suficientemente grande para garantir uma boa aproximação da distribuição nula de  $w$  pela distribuição  $\chi^2$  de referência. O objetivo deste trabalho é obter estatísticas da razão de verossimilhanças ajustadas cujas distribuições sob  $\mathcal{H}_0$  sejam bem aproximadas por uma distribuição  $\chi^2$  mesmo em amostras pequenas ou de tamanho moderado.

Skovgaard (2001) deduziu uma estatística da razão de verossimilhança ajustada dada por

$$w^* = w - 2 \log \xi,$$

onde

$$\xi = \frac{\{|\tilde{I}| |\hat{I}| |\tilde{J}_{\psi\psi}|\}^{1/2}}{|\tilde{\Upsilon}| |\tilde{I}\tilde{\Upsilon}^{-1}\hat{I}\tilde{\Upsilon}^{-1}\tilde{J}_{\psi\psi}|^{1/2}} \frac{\{\tilde{U}^\top \tilde{\Upsilon}^{-1} \hat{I} \tilde{J}^{-1} \tilde{\Upsilon} \tilde{I}^{-1} \tilde{U}\}^{r/2}}{w^{r/2-1} \tilde{U}^\top \tilde{\Upsilon}^{-1} \bar{q}},$$

e  $\bar{q}$  e  $\tilde{\Upsilon}$  são obtidos a partir de

$$q = E_{\omega_1}[U(\omega_1) (\ell(\omega_1) - \ell(\omega))] \quad (3.1)$$

e

$$\Upsilon = E_{\omega_1}[U(\omega_1) U^\top(\omega)] \quad (3.2)$$

substituindo  $\omega_1$  pelo estimador de máxima verossimilhança de  $\omega$  sem restrição ( $\hat{\omega}$ ) e  $\omega$  pelo estimador de máxima verossimilhança de  $\omega$  com restrição sob  $\mathcal{H}_0$  ( $\tilde{\omega}$ ), respectivamente, depois que os valores esperados são calculados. Uma matriz  $A_{\psi\psi}$  é uma matriz formada a partir da matriz  $A$ ,  $(r+s) \times (r+s)$ , excluindo as linhas e colunas correspondentes ao parâmetro de interesse. Assim  $J_{\psi\psi}$  é a matriz de informação observada  $s \times s$  correspondente a  $\psi$ . Além disso,  $\hat{\cdot}$  e  $\tilde{\cdot}$  indicam avaliação no estimador de máxima verossimilhança de  $\omega$  sem restrição ( $\hat{\omega}$ ) e com restrição sob  $\mathcal{H}_0$  ( $\tilde{\omega}$ ) respectivamente. Por exemplo,  $\hat{I} = I(\hat{\omega})$ ,  $\tilde{I} = I(\tilde{\omega})$  e  $\hat{J} = J(\hat{\omega})$ .

Note que  $\bar{q}$  é um vetor de dimensão  $(r + s)$  e  $\bar{\Upsilon}$  é uma matriz de dimensão  $(r + s) \times (r + s)$ . Sob  $\mathcal{H}_0$ ,  $w^*$  é aproximadamente distribuída como uma  $\chi_r^2$  com alto grau de acurácia (Skovgaard, 2001, p. 7). A estatística

$$w^{**} = w \left( 1 - \frac{1}{w} \log \xi \right)^2$$

é assintoticamente equivalente a  $w^*$ . Como assinalado por Skovgaard, a versão  $w^*$  é a que surge a partir do desenvolvimento teórico. Através de simulações, Skovgaard mostrou que  $w^*$  tem uma performance um pouco melhor que  $w^{**}$  em alguns casos. Contudo, diferentemente de  $w^*$ ,  $w^{**}$  é sempre não negativa e se reduz a  $R^{*2}$  quando  $r = 1$ . A estatística  $R^*$  (Barndorff-Nielsen, 1986, 1991) é uma versão modificada da estatística da razão de verossimilhanças que tem distribuição normal sob a hipótese nula com erro de ordem  $n^{-3/2}$ , enquanto o erro desta aproximação é  $n^{-1/2}$  quando a estatística não ajustada é usada. Estas informações podem ser conferidas em Skovgaard (2001, Seção 5.7), onde são apresentadas com mais detalhes. Assim como  $w$ , as estatísticas ajustadas são invariantes sob reparametrização da forma  $(\nu, \psi) \longrightarrow (\nu, \zeta(\nu, \psi))$ .

Com o objetivo de obter as estatísticas da razão de verossimilhanças ajustadas  $w^*$  e  $w^{**}$  no modelo de regressão beta (2.1)-(2.3), é necessário obter o vetor  $q$  e a matriz  $\Upsilon$ . Os valores esperados que aparecem em  $q$  e  $\Upsilon$ , podem ser obtidos a partir dos dois primeiros momentos de  $y_t^*$ 's e  $y_t^\dagger$ 's e das covariâncias entre  $y_t^*$  e  $y_t^\dagger$ , para  $t = 1, \dots, n$ . A obtenção detalhada dos resultados é dada no Apêndice B. Temos que

$$\bar{q} = \begin{bmatrix} X^\top \hat{\Phi} \hat{T} [\hat{V}^* (\hat{\Phi} \hat{\mathcal{M}} - \tilde{\Phi} \tilde{\mathcal{M}}) + \hat{C} (\hat{\Phi} - \tilde{\Phi})]_\iota \\ Z^\top \hat{H} [(\hat{\mathcal{M}} \hat{V}^* + \hat{C}) (\hat{\Phi} \hat{\mathcal{M}} - \tilde{\Phi} \tilde{\mathcal{M}}) + (\hat{\mathcal{M}} \hat{C} + \hat{V}^\dagger) (\hat{\Phi} - \tilde{\Phi})]_\iota \end{bmatrix}$$

e

$$\bar{\Upsilon} = \begin{bmatrix} X^\top \hat{\Phi} \hat{T} \hat{V}^* \tilde{T} \tilde{\Phi} X & X^\top \hat{\Phi} \hat{T} (\hat{\mathcal{M}} \hat{V}^* + \hat{C}) \tilde{H} Z \\ Z^\top \hat{H} (\hat{\mathcal{M}} \hat{V}^* + \hat{C}) \tilde{T} \tilde{\Phi} X & Z^\top \hat{H} [\hat{\mathcal{M}} \hat{V}^* \tilde{\mathcal{M}} + (\hat{\mathcal{M}} + \tilde{\mathcal{M}}) \hat{C} + \hat{V}^\dagger] \tilde{H} Z \end{bmatrix},$$

onde as matrizes foram definidas anteriormente.

Sobre o que foi visto até aqui, vale ressaltar que as estatísticas da razão de verossimilhança ajustadas podem ser facilmente calculadas com os recursos de software existentes. Estas requerem apenas a estimação dos parâmetros desconhecidos sob a hipótese nula e álgebra usual de matrizes.

Quando as estimativas restrita e irrestrita são iguais, temos que  $w = 0$  e  $\bar{q}$  é um vetor nulo, e as fórmulas para as estatísticas ajustadas não fazem sentido. Usando recurso similar ao utilizado por Severini (2000, p. 244) faz-se  $w^*$  e  $w^{**}$  iguais a  $w$  quando  $w < 0.1$ . Em situações práticas, valores de  $w$  próximos de zero indicam que a hipótese nula não deve ser rejeitada, isto é, os  $p$ -valores são claramente grandes. Em tais situações, não há necessidade de ajustes.

Como observação final, mostra-se que as fórmulas das estatísticas da razão de verossimilhanças ajustadas dadas acima podem ser estendidas para modelos de regressão beta não lineares, sendo as especificações para a média e o parâmetro de precisão dadas em (2.9) e (2.10). De modo similar obtêm-se  $\bar{q}$  e  $\bar{\Upsilon}$ :

$$\bar{q} = \begin{bmatrix} \hat{\mathcal{X}}^\top \hat{T} \hat{\Phi} [\hat{V}^* (\hat{\Phi} \hat{\mathcal{M}} - \tilde{\Phi} \tilde{\mathcal{M}}) + \hat{C} (\hat{\Phi} - \tilde{\Phi})]_\iota \\ \hat{Z}^\top \hat{H} [(\hat{\mathcal{M}} \hat{V}^* + \hat{C}) (\hat{\Phi} \hat{\mathcal{M}} - \tilde{\Phi} \tilde{\mathcal{M}}) + (\hat{\mathcal{M}} \hat{C} + \hat{V}^\dagger) (\hat{\Phi} - \tilde{\Phi})]_\iota \end{bmatrix}$$

e

$$\bar{\Upsilon} = \begin{bmatrix} \hat{\chi}^\top \hat{T} \hat{\Phi} \hat{V}^* \tilde{\Phi} \tilde{T} \tilde{\chi} & \hat{\chi}^\top \hat{T} \hat{\Phi} (\tilde{\mathcal{M}} \hat{V}^* + \hat{C}) \tilde{H} \tilde{Z} \\ \hat{Z}^\top \hat{H} (\tilde{\mathcal{M}} \hat{V}^* + \hat{C}) \tilde{\Phi} \tilde{T} \tilde{\chi} & \hat{Z}^\top \hat{H} [\tilde{\mathcal{M}} \hat{V}^* \tilde{\mathcal{M}} + (\tilde{\mathcal{M}} + \tilde{\mathcal{M}}) \hat{C} + \hat{V}^\dagger] \tilde{H} \tilde{Z} \end{bmatrix}$$

(ver Apêndice B).



## Capítulo 4

### Simulações

Este capítulo apresenta o comportamento dos testes baseados em verossimilhança para amostras pequenas e de tamanho moderado, através dos resultados das simulações de Monte Carlo, que são: o teste da razão de verossimilhanças e os testes da razão de verossimilhanças ajustadas ( $w^*$  e  $w^{**}$ ). Os parâmetros são estimados através da maximização do logaritmo da função de verossimilhança utilizando o método quasi-Newton BFGS com primeiras derivadas analíticas. Na Seção 4.1 consideramos o modelo de regressão beta com precisão constante. Na Seção 4.2 consideramos o modelo de regressão beta com precisão variável. Programas na linguagem Ox (Doornik, 2006) utilizados para as simulações podem ser vistos no Apêndice C.

#### 4.1 Modelo de regressão beta com precisão constante

Considera-se o modelo (2.1) com parâmetro de precisão constante e o componente sistemático da média dado por

$$\log \frac{\mu_t}{1 - \mu_t} = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5},$$

o qual denominamos de ‘modelo 1’. Consideram-se três hipóteses nulas distintas, a saber  $\mathcal{H}_0 : \beta_2 = 0$  ( $r = 1$ ),  $\mathcal{H}_0 : \beta_2 = \beta_3 = 0$  ( $r = 2$ ) e  $\mathcal{H}_0 : \beta_2 = \beta_3 = \beta_4 = 0$  ( $r = 3$ ), a serem testadas contra uma alternativa bilateral. Para o primeiro caso, assume-se  $\beta_1 = 1$ ,  $\beta_3 = 1$ ,  $\beta_4 = 5$ ,  $\beta_5 = -4$  e  $\beta_2 = 0$ . Para o segundo caso,  $\beta_1 = 1$ ,  $\beta_4 = 5$ ,  $\beta_5 = -4$  e  $\beta_2 = \beta_3 = 0$ . Finalmente, para o terceiro caso  $\beta_1 = 1$ ,  $\beta_5 = -4$  e  $\beta_2 = \beta_3 = \beta_4 = 0$ . Os valores das covariadas foram obtidos aleatoriamente de uma distribuição  $\mathcal{U}(-0.5, 0.5)$ ,  $\phi$  foi definido como 5, 10 e 30 e os tamanhos das amostras como 15, 20, 30 e 40. O número de réplicas de Monte Carlo usado foi 50000 e os níveis nominais dos testes foram  $\alpha = 10\%$ ,  $5\%$  e  $1\%$ . Os resultados são resumidos nas Tabelas 4.1 e 4.2 e na Figura 4.1.

Algumas conclusões interessantes podem ser tiradas da Tabela 4.1. A primeira é que o teste da razão de verossimilhanças é notadamente liberal para amostras pequenas. Por exemplo, para  $\phi = 30$ ,  $n = 15$ ,  $r = 1$  e  $\alpha = 5\%$ , a taxa de rejeição nula é 12.2%. A segunda é que os testes ajustados se comportam muito melhor do que o teste da razão de verossimilhanças usual em relação ao erro do tipo I. Para o caso mencionado acima, a taxa de rejeição nula dos testes ajustados são 5.0% ( $w^*$ ) e 5.9% ( $w^{**}$ ). A terceira é que o teste baseado em  $w^{**}$  é um pouco mais liberal que o teste que usa  $w^*$ .

Tabela 4.1: Taxas de rejeição nula (%); modelo 1

| $\phi$ | $r$ | $n$ | $\alpha = 10\%$ |       |          | $\alpha = 5\%$ |       |          | $\alpha = 1\%$ |       |          |
|--------|-----|-----|-----------------|-------|----------|----------------|-------|----------|----------------|-------|----------|
|        |     |     | $w$             | $w^*$ | $w^{**}$ | $w$            | $w^*$ | $w^{**}$ | $w$            | $w^*$ | $w^{**}$ |
| 30     | 1   | 15  | 19.2            | 10.2  | 11.5     | 12.2           | 5.0   | 5.9      | 4.1            | 1.0   | 1.3      |
|        |     | 20  | 16.1            | 9.7   | 10.4     | 9.3            | 5.0   | 5.5      | 2.8            | 1.0   | 1.1      |
|        |     | 30  | 13.8            | 10.0  | 10.4     | 7.6            | 4.9   | 5.1      | 2.1            | 1.0   | 1.1      |
|        |     | 40  | 12.7            | 10.0  | 10.2     | 6.8            | 5.0   | 5.1      | 1.8            | 1.0   | 1.1      |
|        | 2   | 15  | 22.0            | 9.9   | 11.3     | 13.8           | 5.0   | 5.8      | 4.6            | 1.0   | 1.3      |
|        |     | 20  | 18.0            | 10.0  | 10.6     | 10.7           | 5.0   | 5.5      | 3.3            | 1.1   | 1.3      |
|        |     | 30  | 14.9            | 10.0  | 10.4     | 8.3            | 5.0   | 5.3      | 2.2            | 1.0   | 1.1      |
|        |     | 40  | 13.3            | 10.1  | 10.2     | 7.3            | 5.0   | 5.2      | 1.9            | 1.1   | 1.2      |
|        | 3   | 15  | 22.5            | 9.8   | 11.0     | 14.2           | 5.0   | 5.7      | 4.8            | 1.0   | 1.4      |
|        |     | 20  | 18.5            | 10.1  | 10.7     | 11.2           | 5.1   | 5.5      | 3.4            | 1.0   | 1.2      |
|        |     | 30  | 15.0            | 9.7   | 10.0     | 8.3            | 5.1   | 5.3      | 2.2            | 0.9   | 1.0      |
|        |     | 40  | 13.2            | 9.8   | 10.0     | 7.2            | 5.0   | 5.0      | 1.9            | 1.0   | 1.1      |
| 10     | 1   | 15  | 19.8            | 10.0  | 11.5     | 12.5           | 5.1   | 6.1      | 4.4            | 1.0   | 1.4      |
|        |     | 20  | 16.4            | 10.0  | 10.8     | 9.8            | 5.0   | 5.5      | 2.9            | 1.0   | 1.2      |
|        |     | 30  | 13.7            | 10.0  | 10.4     | 7.7            | 5.0   | 5.2      | 2.1            | 1.0   | 1.2      |
|        |     | 40  | 12.5            | 9.9   | 10.1     | 6.8            | 5.0   | 5.2      | 1.6            | 1.0   | 1.1      |
|        | 2   | 15  | 21.9            | 9.7   | 11.1     | 13.9           | 5.0   | 6.0      | 4.9            | 0.9   | 1.4      |
|        |     | 20  | 17.8            | 9.8   | 10.6     | 10.7           | 5.0   | 5.5      | 3.2            | 1.0   | 1.2      |
|        |     | 30  | 14.8            | 10.0  | 10.4     | 8.2            | 5.0   | 5.3      | 2.1            | 1.0   | 1.2      |
|        |     | 40  | 13.3            | 9.9   | 10.1     | 7.3            | 5.0   | 5.2      | 1.8            | 1.0   | 1.1      |
|        | 3   | 15  | 22.9            | 10.0  | 11.2     | 14.5           | 5.0   | 5.8      | 5.0            | 1.0   | 1.3      |
|        |     | 20  | 18.9            | 10.1  | 10.7     | 11.3           | 5.1   | 5.6      | 3.5            | 1.1   | 1.3      |
|        |     | 30  | 15.1            | 9.9   | 10.2     | 8.4            | 5.0   | 5.2      | 2.2            | 1.0   | 1.1      |
|        |     | 40  | 13.5            | 10.0  | 10.1     | 7.5            | 5.0   | 5.1      | 1.9            | 1.0   | 1.1      |
| 5      | 1   | 15  | 19.8            | 10.4  | 12.1     | 12.7           | 5.2   | 6.4      | 4.4            | 1.0   | 1.5      |
|        |     | 20  | 16.2            | 10.3  | 11.1     | 9.8            | 5.1   | 5.7      | 2.8            | 0.9   | 1.1      |
|        |     | 30  | 13.3            | 9.8   | 10.3     | 7.3            | 5.0   | 5.3      | 1.8            | 0.9   | 1.0      |
|        |     | 40  | 11.8            | 9.4   | 9.7      | 6.2            | 4.7   | 4.8      | 1.5            | 0.9   | 1.0      |
|        | 2   | 15  | 22.2            | 9.9   | 11.4     | 14.0           | 5.0   | 6.0      | 5.0            | 1.1   | 1.5      |
|        |     | 20  | 17.6            | 9.8   | 10.7     | 10.5           | 4.9   | 5.6      | 3.2            | 1.0   | 1.3      |
|        |     | 30  | 14.2            | 9.8   | 10.4     | 7.9            | 4.9   | 5.3      | 2.0            | 1.0   | 1.2      |
|        |     | 40  | 12.7            | 9.7   | 10.0     | 6.9            | 4.9   | 5.1      | 1.6            | 0.9   | 1.1      |
|        | 3   | 15  | 22.8            | 9.8   | 11.0     | 14.4           | 4.9   | 5.8      | 5.1            | 1.0   | 1.3      |
|        |     | 20  | 18.4            | 10.0  | 10.6     | 11.1           | 5.0   | 5.5      | 3.4            | 1.0   | 1.2      |
|        |     | 30  | 15.0            | 10.1  | 10.3     | 8.4            | 5.1   | 5.3      | 2.2            | 1.0   | 1.2      |
|        |     | 40  | 13.6            | 10.2  | 10.4     | 7.5            | 5.1   | 5.2      | 1.8            | 1.0   | 1.1      |



Tabela 4.2: Taxas de rejeição não nula (%); modelo 1,  $n = 20$ ,  $\alpha = 10\%$ 

| $\phi$ | $r$ | $\epsilon$ | -3.0  | -2.5  | -2.0  | -1.5 | -1.0 | -0.5 | 0.5  | 1.0  | 1.5   | 2.0   | 2.5   | 3.0   |
|--------|-----|------------|-------|-------|-------|------|------|------|------|------|-------|-------|-------|-------|
| 30     | 1   | $w^*$      | 100.0 | 100.0 | 100.0 | 98.9 | 73.2 | 26.1 | 29.4 | 87.3 | 100.0 | 100.0 | 100.0 | 100.0 |
|        |     | $w^{**}$   | 100.0 | 100.0 | 100.0 | 99.3 | 78.8 | 35.5 | 38.0 | 89.8 | 100.0 | 100.0 | 100.0 | 100.0 |
|        | 2   | $w^*$      | 100.0 | 100.0 | 100.0 | 99.1 | 83.6 | 34.9 | 35.1 | 86.5 | 99.7  | 100.0 | 100.0 | 100.0 |
|        |     | $w^{**}$   | 100.0 | 100.0 | 100.0 | 99.2 | 84.3 | 35.7 | 36.1 | 86.9 | 99.7  | 100.0 | 100.0 | 100.0 |
|        | 3   | $w^*$      | 100.0 | 100.0 | 100.0 | 99.9 | 94.8 | 53.1 | 55.4 | 96.8 | 100.0 | 100.0 | 100.0 | 100.0 |
|        |     | $w^{**}$   | 100.0 | 100.0 | 100.0 | 99.9 | 95.0 | 53.8 | 55.8 | 96.9 | 100.0 | 100.0 | 100.0 | 100.0 |
| 10     | 1   | $w^*$      | 99.4  | 97.1  | 85.4  | 56.8 | 31.6 | 14.3 | 15.4 | 32.9 | 65.4  | 94.5  | 99.5  | 99.9  |
|        |     | $w^{**}$   | 99.8  | 98.3  | 89.5  | 64.4 | 40.6 | 24.7 | 27.8 | 41.6 | 72.4  | 96.4  | 99.8  | 100.0 |
|        | 2   | $w^*$      | 99.6  | 98.0  | 90.6  | 70.8 | 42.5 | 18.5 | 17.7 | 41.7 | 70.9  | 91.4  | 98.3  | 99.6  |
|        |     | $w^{**}$   | 99.7  | 98.3  | 91.4  | 72.0 | 43.8 | 19.1 | 18.6 | 43.2 | 72.8  | 92.5  | 98.7  | 99.8  |
|        | 3   | $w^*$      | 99.3  | 98.8  | 96.3  | 87.1 | 63.3 | 27.9 | 29.4 | 70.6 | 93.5  | 99.3  | 99.9  | 100.0 |
|        |     | $w^{**}$   | 99.8  | 99.5  | 97.3  | 88.2 | 64.3 | 28.9 | 30.2 | 70.9 | 93.7  | 99.3  | 100.0 | 100.0 |
| 5      | 1   | $w^*$      | 87.3  | 72.8  | 52.0  | 34.9 | 22.1 | 12.4 | 12.7 | 25.3 | 39.3  | 67.3  | 90.4  | 97.9  |
|        |     | $w^{**}$   | 93.0  | 81.3  | 62.2  | 44.1 | 32.9 | 22.1 | 24.1 | 33.8 | 48.5  | 74.7  | 94.0  | 99.2  |
|        | 2   | $w^*$      | 92.7  | 85.0  | 70.3  | 50.8 | 30.9 | 15.8 | 14.4 | 28.1 | 47.3  | 67.7  | 83.5  | 92.4  |
|        |     | $w^{**}$   | 94.3  | 87.2  | 72.8  | 52.6 | 32.0 | 16.5 | 15.4 | 29.7 | 50.2  | 71.5  | 87.4  | 95.2  |
|        | 3   | $w^*$      | 89.3  | 85.6  | 77.6  | 64.2 | 44.5 | 20.4 | 22.0 | 53.2 | 79.3  | 93.1  | 98.0  | 99.3  |
|        |     | $w^{**}$   | 95.3  | 91.4  | 82.4  | 66.9 | 45.7 | 21.4 | 22.7 | 53.7 | 79.8  | 93.7  | 98.3  | 99.6  |

A Tabela 4.2 apresenta taxas de rejeição dos testes ajustados das três hipóteses nulas para o modelo 1 com  $n = 20$ ,  $\alpha = 10\%$ ,  $\phi = 5$ ,  $\phi = 10$  e  $\phi = 30$ . Aqui, as taxas de rejeição foram obtidas sob a hipótese alternativa  $\beta_2 = \epsilon$  ( $r = 1$ ),  $\beta_2 = \beta_3 = \epsilon$  ( $r = 2$ ) e  $\beta_2 = \beta_3 = \beta_4 = \epsilon$  ( $r = 3$ ), para diferentes valores de  $\epsilon$ . O teste da razão de verossimilhanças não está incluído na comparação de poder já que ele é muito liberal. Os poderes dos dois testes ajustados são similares, com pequena vantagem para o teste que usa  $w^{**}$ , no entanto esse é um pouco mais liberal. Quando a precisão é maior o poder cresce mais rapidamente conforme  $\epsilon$  se afasta da hipótese nula.

A Figura 4.1 apresenta gráficos quantil-quantil (quantil empírico versus quantil assintótico) das diferentes estatísticas para teste das hipóteses  $\mathcal{H}_0 : \beta_2 = 0$  ( $r = 1$ ),  $\mathcal{H}_0 : \beta_2 = \beta_3 = 0$  ( $r = 2$ ) e  $\mathcal{H}_0 : \beta_2 = \beta_3 = \beta_4 = 0$  ( $r = 3$ ) para  $\phi = 30$ . Os quantis assintóticos são obtidos de uma distribuição  $\chi_r^2$ . As figuras confirmam a tendência da estatística da razão de verossimilhanças usual de rejeitar a hipótese nula com mais frequência do que o esperado, baseado no nível nominal. Nota-se que as distribuições das estatísticas ajustadas são muito mais próximas da distribuição de referência do que a distribuição de  $w$ . A melhor concordância entre os quantis verdadeiros e assintóticos é alcançada por  $w^*$ . Nas três situações, número de parâmetros de interesse igual a  $r = 1$ ,  $r = 2$  e  $r = 3$ , as distribuições empíricas das estatísticas ajustadas estão muito próximas das respectivas distribuições limite mesmo quando a amostra é pequena, isto é,  $n = 20$ .

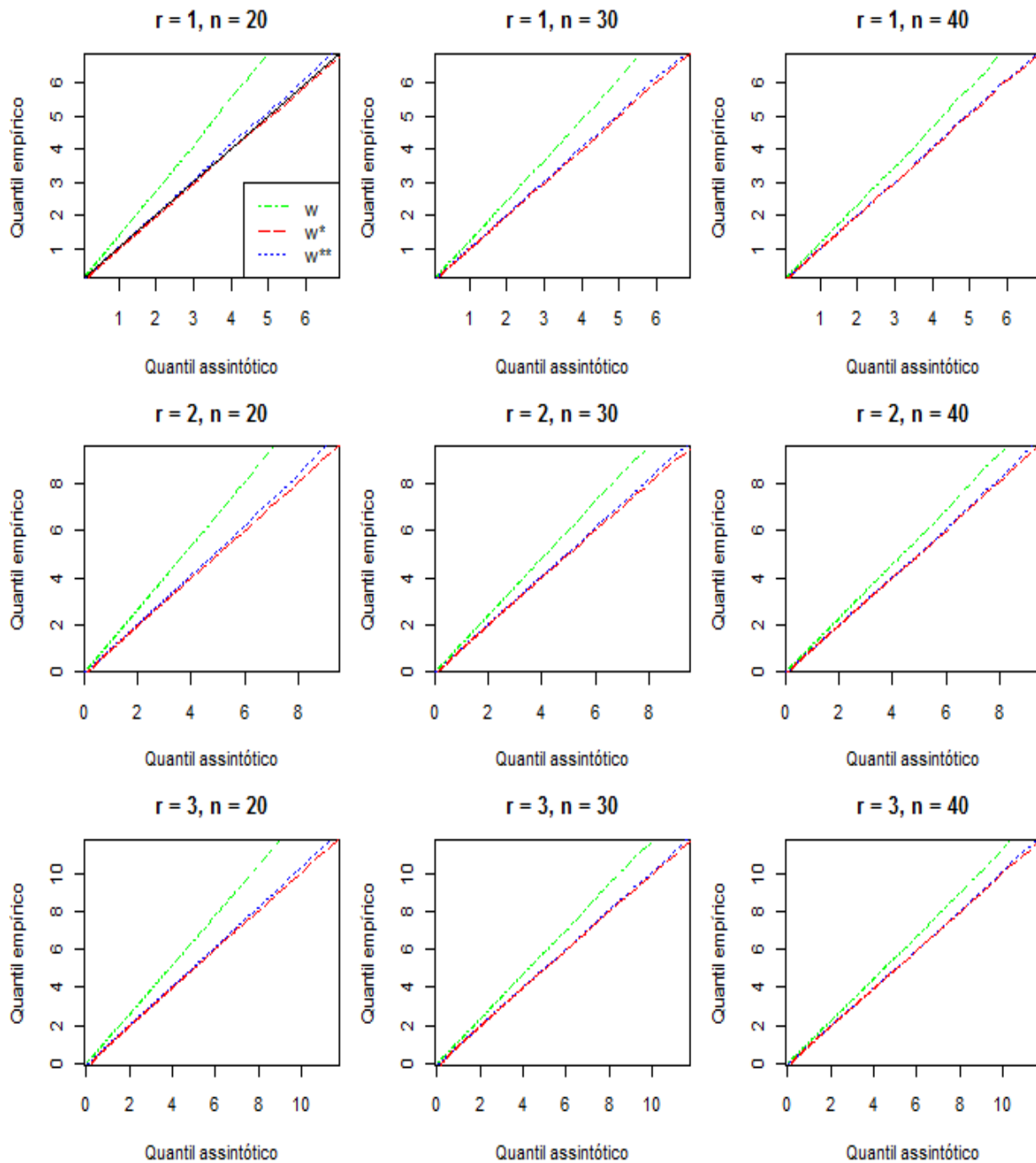


Figura 4.1: Gráficos quantil-quantil, modelo 1,  $\phi = 30$ .

## 4.2 Modelo de regressão beta com precisão variável

Considera-se agora o modelo (2.1) com o componente sistemático da média dado por

$$\log \frac{\mu_t}{1 - \mu_t} = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4},$$

e especificação da precisão dada por

$$\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4},$$

o qual denominamos de ‘modelo 2’. Considera-se a hipótese nula  $\mathcal{H}_0 : \gamma_2 = \gamma_3 = \gamma_4 = 0$  ( $r = 3$ )(precisão constante), que é testada contra uma alternativa bilateral. Novamente os valores das covariadas  $x$ 's foram escolhidos aleatoriamente de uma distribuição  $\mathcal{U}(-0.5, 0.5)$  e faz-se  $z_{ti} = x_{ti}$ . Os valores da variável resposta foram gerados usando  $\beta_1 = 1, \beta_2 = 1, \beta_3 = 5, \beta_4 = -4, \gamma_1 = \log(30), \gamma_2 = \gamma_3 = \gamma_4 = 0$ ,  $\gamma_1 = \log(10)$  e  $\gamma_1 = \log(5)$  (simulações de tamanho) e  $\gamma_2 = \gamma_3 = \gamma_4 = \epsilon$  para diferentes valores de  $\epsilon$  (simulações de poder). Os tamanhos das amostras foram 20, 30, 40, 50 e 60. Os resultados numéricos estão resumidos nas Tabelas 4.3 e 4.4 e Figura 4.2.

Tabela 4.3: Taxas de rejeição nula (%); modelo 2

| $\gamma_1$ | $n$ | $\alpha = 10\%$ |       |          | $\alpha = 5\%$ |       |          | $\alpha = 1\%$ |       |          |
|------------|-----|-----------------|-------|----------|----------------|-------|----------|----------------|-------|----------|
|            |     | $w$             | $w^*$ | $w^{**}$ | $w$            | $w^*$ | $w^{**}$ | $w$            | $w^*$ | $w^{**}$ |
| log(30)    | 20  | 31.8            | 11.0  | 13.5     | 22.1           | 6.1   | 7.7      | 9.4            | 1.7   | 2.3      |
|            | 30  | 20.9            | 10.8  | 11.6     | 12.9           | 5.7   | 6.2      | 4.2            | 1.2   | 1.3      |
|            | 40  | 16.8            | 10.0  | 10.4     | 9.8            | 5.0   | 5.3      | 2.8            | 1.1   | 1.2      |
|            | 50  | 15.0            | 9.8   | 10.1     | 8.3            | 4.9   | 5.1      | 2.3            | 1.0   | 1.1      |
|            | 60  | 14.0            | 10.0  | 10.2     | 7.7            | 5.0   | 5.1      | 1.9            | 1.1   | 1.1      |
| log(10)    | 20  | 27.7            | 10.1  | 11.9     | 18.6           | 5.4   | 6.6      | 7.5            | 1.4   | 1.8      |
|            | 30  | 19.3            | 11.0  | 11.6     | 11.5           | 5.6   | 6.0      | 3.5            | 1.3   | 1.4      |
|            | 40  | 15.3            | 10.1  | 10.3     | 8.7            | 5.1   | 5.2      | 2.3            | 1.0   | 1.1      |
|            | 50  | 14.4            | 10.1  | 10.2     | 7.8            | 4.9   | 5.0      | 1.9            | 1.0   | 1.0      |
|            | 60  | 13.4            | 10.1  | 10.2     | 7.2            | 5.0   | 5.0      | 1.7            | 1.0   | 1.1      |
| log(5)     | 20  | 23.7            | 11.4  | 12.8     | 15.2           | 6.2   | 7.1      | 5.6            | 1.6   | 1.9      |
|            | 30  | 18.7            | 10.1  | 10.7     | 11.1           | 5.3   | 5.6      | 3.5            | 1.1   | 1.2      |
|            | 40  | 14.9            | 10.4  | 10.6     | 8.3            | 5.4   | 5.5      | 2.2            | 1.1   | 1.1      |
|            | 50  | 13.5            | 10.5  | 10.6     | 7.3            | 5.3   | 5.3      | 1.8            | 1.1   | 1.1      |
|            | 60  | 13.3            | 11.1  | 11.2     | 7.3            | 5.7   | 5.7      | 1.8            | 1.2   | 1.2      |

Observa-se na Tabela 4.3 que há um ganho significativo em relação à probabilidade do erro de tipo I quando se utilizam os testes ajustados em lugar do teste da razão de verossimilhanças usual, principalmente quando a amostra é pequena. Por exemplo, para  $\gamma_1 = \log(30)$ ,  $n = 20$  e  $\alpha = 10\%$ , a probabilidade do erro de tipo I do teste usual é 31.8%, enquanto o do teste ajustado que usa  $w^*$  é 11.0%. Como no ‘modelo 1’, o teste baseado em  $w^{**}$  é um pouco mais liberal do que o teste que utiliza  $w^*$  sendo a probabilidade do erro

de tipo I destes mais próximas dos níveis nominais. Por exemplo, para  $\gamma_1 = \log 30$ ,  $n = 20$  e  $\alpha = 1\%$  as taxas de rejeição dos testes são 9.4% ( $w$ ), 1.7% ( $w^*$ ) e 2.3% ( $w^{**}$ ). Como esperado os ajustes são menos necessários para amostras grandes. Todavia, note que não é desprezível o ganho obtido para  $n = 60$  nas três configurações testadas.

Tabela 4.4: Taxas de rejeição não nula (%); modelo 2,  $n = 40$ ,  $\alpha = 10\%$ .

| $\gamma_1$ | $\epsilon$ | -3.0 | -2.5 | -2.0 | -1.5 | -1.0 | -0.5 | 0.5  | 1.0  | 1.5  | 2.0  | 2.5  | 3.0  |
|------------|------------|------|------|------|------|------|------|------|------|------|------|------|------|
| log(30)    | $w^*$      | 99.9 | 99.1 | 95.6 | 82.3 | 55.5 | 24.0 | 21.2 | 51.8 | 76.8 | 91.0 | 96.7 | 98.6 |
|            | $w^{**}$   | 99.9 | 99.2 | 95.6 | 82.5 | 55.9 | 24.4 | 21.9 | 52.4 | 77.3 | 91.8 | 97.4 | 99.2 |
| log(10)    | $w^*$      | 99.9 | 99.2 | 96.0 | 82.3 | 53.8 | 23.6 | 21.2 | 50.4 | 76.0 | 90.5 | 96.4 | 98.5 |
|            | $w^{**}$   | 99.9 | 99.3 | 96.0 | 82.5 | 54.1 | 24.0 | 21.9 | 51.1 | 76.7 | 91.7 | 97.5 | 99.3 |
| log(5)     | $w^*$      | 99.5 | 98.1 | 92.9 | 76.7 | 48.2 | 21.3 | 21.1 | 48.0 | 73.5 | 88.9 | 95.9 | 97.9 |
|            | $w^{**}$   | 99.5 | 98.2 | 92.9 | 76.9 | 48.5 | 21.7 | 21.9 | 48.7 | 74.6 | 90.5 | 97.3 | 99.1 |

As conclusões obtidas para o ‘modelo 2’, a partir da Tabela 4.4, não se alteram em relação às conclusões para o ‘modelo 1’. Os testes ajustados são quase igualmente poderosos, sendo que  $w^{**}$  é ligeiramente mais poderoso devido ao fato de ser um pouco mais liberal.

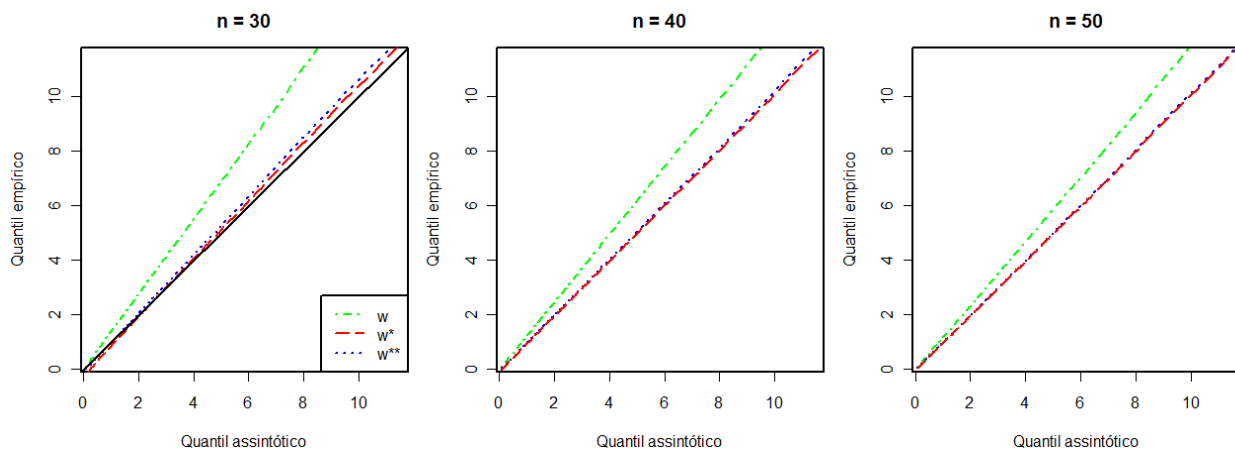


Figura 4.2: Gráficos quantil-quantil, modelo 2,  $\gamma_1 = \log 30$ .

A Figura 4.2 apresenta gráficos quantil-quantil (quantil empírico versus quantil assintótico) das diferentes estatísticas para teste da hipótese  $\mathcal{H}_0 : \gamma_2 = \gamma_3 = \gamma_4 = 0$  ( $r = 3$ ) para  $\gamma_1 = \log(30)$ . Da mesma forma que no ‘modelo 1’, os quantis assintóticos são obtidos de uma distribuição  $\chi_r^2$ .

Nota-se que no ‘modelo 2’ a distribuição  $\chi_r^2$  não é uma aproximação satisfatória para a distribuição nula de  $w$  para diferentes tamanhos de amostra ( $n$ ). De fato, o teste que usa  $w$  é consideravelmente liberal. A distribuição de referência está razoavelmente próxima das distribuições das estatísticas ajustadas. O melhor desempenho é alcançado pelo teste que se baseia na estatística  $w^*$ . Nota-se que à medida em que o número de observações cresce há uma melhora significativa em relação à proximidade entre a distribuição

de referência e as distribuições das estatísticas, isto é, as distorções de tamanho dos testes são reduzidas, em especial no que se refere às estatísticas ajustadas  $w^*$  e  $w^{**}$ . Para  $n = 50$  as distribuições das estatísticas ajustadas são praticamente iguais às da distribuição de referência  $\chi_3^2$ .

Considera-se agora o modelo (2.1) com o componente sistemático da média dado por

$$\log \frac{\mu_t}{1 - \mu_t} = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5},$$

e especificação da precisão dada por

$$\log(\phi_t) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5},$$

o qual denominamos de ‘modelo 3’. A hipótese nula é  $\mathcal{H}_0 : \gamma_2 = \dots = \gamma_5 = 0$  (precisão constante), que é testada contra a alternativa bilateral. Novamente os valores das covariadas  $x$ 's foram escolhidos aleatoriamente de uma distribuição  $\mathcal{U}(-0.5, 0.5)$  e faz-se  $z_{ti} = x_{ti}$ . Os valores da resposta foram gerados usando  $\beta_1 = 1$ ,  $\beta_2 = 0$ ,  $\beta_3 = 5$ ,  $\beta_4 = -4$ ,  $\beta_5 = 1$  e  $\gamma_1 = \log(30)$ ,  $\gamma_2 = \dots = \gamma_5 = 0$  (simulações de tamanho) e  $\gamma_1 = \log(30)$ ,  $\gamma_2 = \dots = \gamma_5 = \epsilon$  para diferentes valores de  $\epsilon$  (simulações de poder). Os tamanhos das amostras foram 20, 30, 40, 50 e 60. Os resultados são apresentados nas Tabelas 4.5 e 4.6 e Figura 4.3.

Tabela 4.5: Taxas de rejeição nula; modelo 3

| $n$ | $\alpha = 10\%$ |       |          | $\alpha = 5\%$ |       |          | $\alpha = 1\%$ |       |          |
|-----|-----------------|-------|----------|----------------|-------|----------|----------------|-------|----------|
|     | $w$             | $w^*$ | $w^{**}$ | $w$            | $w^*$ | $w^{**}$ | $w$            | $w^*$ | $w^{**}$ |
| 20  | 50.2            | 11.4  | 15.6     | 39.4           | 7.1   | 9.9      | 22.0           | 2.7   | 3.8      |
| 30  | 29.1            | 9.7   | 11.5     | 19.7           | 5.3   | 6.3      | 7.7            | 1.3   | 1.7      |
| 40  | 21.1            | 9.2   | 10.0     | 12.9           | 4.7   | 5.2      | 4.0            | 1.0   | 1.2      |
| 50  | 17.8            | 9.6   | 10.0     | 9.2            | 4.7   | 5.0      | 2.5            | 0.9   | 1.0      |
| 60  | 16.2            | 9.6   | 10.0     | 9.2            | 4.7   | 5.0      | 2.5            | 0.9   | 1.0      |

A Tabela 4.5 mostra que no ‘modelo 3’ o comportamento das estatísticas em estudo confirma as observações feitas até o momento. O teste que usa  $w$  é tipicamente liberal enquanto as taxas de rejeição dos testes que usam as estatísticas ajustadas estão bem mais próximas dos níveis nominais. Por exemplo, para  $n = 30$  e  $\alpha = 10\%$ , a taxa de rejeição nula do teste que utiliza  $w$  é 29.1%, enquanto a do teste baseado na estatística  $w^*$  é 9.7% e a do teste que usa a estatística  $w^{**}$  é 11.5%. Os testes ajustados rejeitam a hipótese nula com frequência bastante próxima do nível nominal, sendo que o teste que se baseia em  $w^*$  apresenta desempenho melhor em pequenas amostras.

Os resultados obtidos para o ‘modelo 3’, mostrados na Tabela 4.6, confirmam que os testes ajustados são quase igualmente poderosos, sendo que aquele que utiliza  $w^{**}$  é um pouco mais poderoso devido ao fato de ser um pouco mais liberal.

Tabela 4.6: Taxas de rejeição não nula (%); modelo 3,  $n = 40$ ,  $\alpha = 10\%$ .

| $\epsilon$ | -3.0 | -2.5 | -2.0 | -1.5 | -1.0 | -0.5 | 0.5  | 1.0  | 1.5  | 2.0  | 2.5  | 3.0  |
|------------|------|------|------|------|------|------|------|------|------|------|------|------|
| $w^*$      | 99.2 | 98.6 | 95.9 | 86.1 | 61.8 | 26.2 | 24.6 | 61.2 | 84.1 | 93.7 | 96.2 | 96.3 |
| $w^{**}$   | 99.6 | 99.0 | 96.3 | 86.5 | 62.5 | 27.3 | 25.9 | 62.2 | 85.3 | 95.2 | 98.1 | 98.8 |

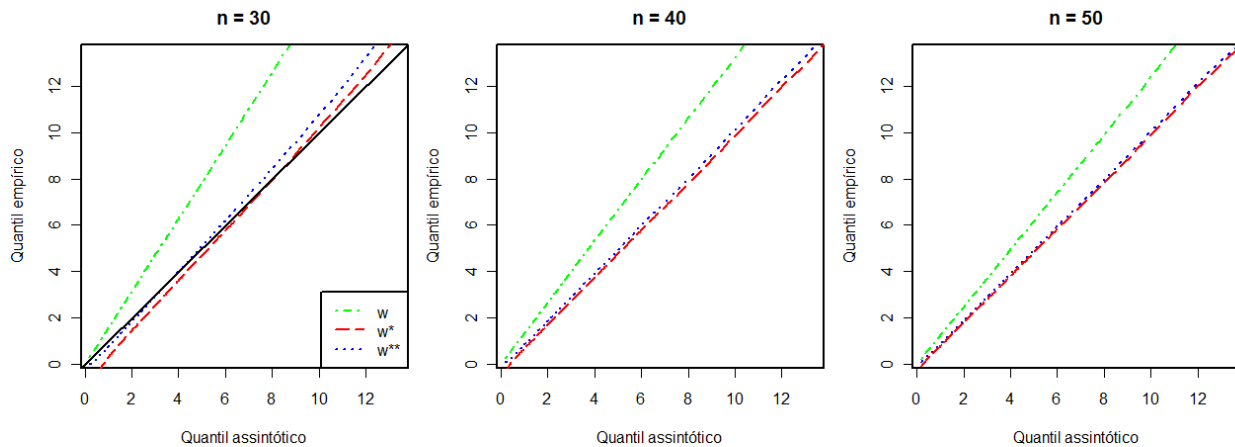


Figura 4.3: Gráficos quantil-quantil, modelo 3.

A Figura 4.3 mostra que a distribuição  $\chi_4^2$  não é uma aproximação satisfatória para a distribuição nula de  $w$ . Por outro lado, a distribuição de referência está próxima das distribuições das estatísticas ajustadas  $w^*$  e  $w^{**}$ .

Em resumo, dentre os três testes analisados ( $w$ ,  $w^*$  e  $w^{**}$ ), o teste da razão de verossimilhanças ajustada que utiliza  $w^*$  é o que produz, em geral, taxas de rejeição nula mais próximas dos níveis nominais especialmente quando  $n$  é pequeno. Em outras palavras,  $w^*$  é o teste mais confiável no que diz respeito à probabilidade de erro de tipo I. Adicionalmente, ele é quase igualmente tão poderoso quanto o teste que se baseia em  $w^{**}$ . Assim, recomendamos o uso da estatística  $w^*$  nas aplicações práticas.

## Capítulo 5

### Aplicação

Neste capítulo aplicaremos os resultados desenvolvidos até aqui a um conjunto real de dados obtido em Brownlee (1965, p. 454) e apresentado na Tabela 5.1. Trata-se de um conjunto de 21 observações provenientes de 21 dias de operação de uma fábrica no processo de oxidação de amônia como etapa na produção de ácido nítrico. As variáveis explicativas são *corrente de ar* ( $x_2$ ), *temperatura da água* ( $x_3$ ) e *concentração de ácido* ( $x_4$ ) e a variável resposta é a *proporção de perda na conversão* ( $y$ ). Segundo Atkinson (1985, p. 129), a *corrente de ar* mede a taxa de operação da fábrica, a *temperatura* da água é a da corrente de água de resfriamento de uma torre que absorve o óxido nítrico produzido no processo, *concentração de ácido* ( $10 \times (\text{concentração de ácido} - 50)$ ) se refere ao óxido nítrico e a *proporção de perda na conversão* é a proporção de amônia que não foi convertida em ácido nítrico. Para uma descrição detalhada dos dados ver a dissertação de mestrado de Oliveira (2004).

Este conjunto de dados tem sido extensivamente analisado na literatura estatística. Para uma descrição detalhada da evolução das análises ver Atkinson (1985, p. 266-268). Resumimos aqui o que orientou a escolha do nosso modelo inicial.

O estudo da oxidação da amônia foi analisado pela primeira vez por Brownlee (1965) quando propôs um modelo de regressão linear múltipla com três variáveis explicativas, sem utilizar transformação na resposta. O autor concluiu que a variável *concentração de ácido* ( $x_4$ ) poderia ser retirada do modelo. Esses dados foram analisados por Atkinson (1985, p. 129), que propôs um modelo de regressão linear (homoscedástico) para  $\log(y)$  com as covariadas  $x_2$ ,  $x_3$ ,  $x_2x_3$  e  $x_2^2$ . Ele mostrou que esse modelo se ajusta bem aos dados, mas a observação 21 mostra-se muito influente. De fato, a eliminação desta observação torna a covariada  $x_2^2$  irrelevante. Aqui, consideramos o modelo de regressão beta que inclui as covariadas *corrente de ar* ( $x_2$ ), *temperatura da água* ( $x_3$ ) e *produto* ( $x_2x_3$ ). O modelo inicial tem o componente sistemático da média dado por

$$\log \frac{\mu_t}{1 - \mu_t} = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t2} x_{t3}$$

e especificação da precisão dada por

$$\log(\phi_t) = \gamma_1 + \gamma_2 x_{t2} + \gamma_3 x_{t3},$$

o qual chamamos de ‘modelo I’. As estimativas de máxima verossimilhança dos parâmetros para o ‘modelo I’ estão na Tabela 5.2. Programas na linguagem Ox (Doornik, 2006) utilizados para cálculo das estimativas

Tabela 5.1: Dados do estudo de Oxidação de Amônia.

| Observação | Perda na conversão<br>$y$ | Corrente de ar<br>$x_2$ | Temperatura da água<br>$x_3$ | Concentração de ácido<br>$x_4$ |
|------------|---------------------------|-------------------------|------------------------------|--------------------------------|
| 1          | 0.042                     | 80                      | 27                           | 89                             |
| 2          | 0.037                     | 80                      | 27                           | 88                             |
| 3          | 0.037                     | 75                      | 25                           | 90                             |
| 4          | 0.028                     | 62                      | 24                           | 87                             |
| 5          | 0.018                     | 62                      | 22                           | 87                             |
| 6          | 0.018                     | 62                      | 23                           | 87                             |
| 7          | 0.019                     | 62                      | 24                           | 93                             |
| 8          | 0.020                     | 62                      | 24                           | 93                             |
| 9          | 0.015                     | 58                      | 23                           | 87                             |
| 10         | 0.014                     | 58                      | 18                           | 80                             |
| 11         | 0.014                     | 58                      | 18                           | 89                             |
| 12         | 0.013                     | 58                      | 17                           | 88                             |
| 13         | 0.011                     | 58                      | 18                           | 82                             |
| 14         | 0.012                     | 58                      | 19                           | 93                             |
| 15         | 0.008                     | 50                      | 18                           | 89                             |
| 16         | 0.007                     | 50                      | 18                           | 86                             |
| 17         | 0.008                     | 50                      | 19                           | 72                             |
| 18         | 0.008                     | 50                      | 19                           | 79                             |
| 19         | 0.009                     | 50                      | 20                           | 80                             |
| 20         | 0.015                     | 56                      | 20                           | 82                             |
| 21         | 0.015                     | 70                      | 20                           | 91                             |

e estatísticas podem ser vistos no Apêndice C.

Tabela 5.2: Estimativas dos parâmetros e erros padrão : modelo I

| Parâmetros | $\beta_1$  | $\beta_2$      | $\beta_3$   | $\beta_4$ | $\gamma_1$ | $\gamma_2$     | $\gamma_3$  |
|------------|------------|----------------|-------------|-----------|------------|----------------|-------------|
| Descrição  | Intercepto | Corrente de ar | Temperatura | Produto   | Intercepto | Corrente de ar | Temperatura |
| EMV        | -12.239    | 0.119          | 0.232       | -0.003    | 19.220     | -0.211         | 0.079       |
| EP         | 2.333      | 0.040          | 0.115       | 0.002     | 2.248      | 0.055          | 0.160       |

Inicialmente, testamos a hipótese nula  $\mathcal{H}_0 : \gamma_2 = \gamma_3 = 0$  (precisão constante). A estatística da razão de verossimilhanças ( $w$ ) é igual a 6.44 (valor- $p = 0.04$ ). Ao nível nominal de 5%, a hipótese nula é rejeitada. No entanto, o fato do valor- $p$  ser próximo de 5% e o tamanho da amostra ser pequeno ( $n = 21$ ) leva a



dúvida quanto à decisão a tomar. Em situações como essa é bastante útil calcular as estatísticas ajustadas. Obtivemos  $w^* = 10.37$  (valor- $p = 0.01$ ) e  $w^{**} = 10.97$  (valor- $p = 0.00$ ). Note que os testes ajustados conduzem a forte evidência contra a hipótese nula confirmando que essa deve ser rejeitada ao nível nominal escolhido.

Testamos, a seguir, a hipótese  $\mathcal{H}_0 : \gamma_3 = 0$ . A estatística do teste da razão de verossimilhanças ( $w$ ) é igual a 0.34 (valor- $p = 0.56$ ), e as duas estatísticas ajustadas são  $w^* = 0.07$  (valor- $p = 0.80$ ) e  $w^{**} = 0.12$  (valor- $p = 0.73$ ). Claramente, a inferência baseada nos três testes não rejeita a hipótese nula e por isso removemos o efeito de temperatura da especificação da precisão. Assumimos, então, o modelo com mesmo componente sistemático para a média e especificação da precisão dada por

$$\log(\phi_t) = \gamma_1 + \gamma_2 x_{t2},$$

o qual chamamos de ‘modelo II’. As estimativas de máxima verossimilhança dos parâmetros para o ‘modelo II’ estão na Tabela 5.3.

Tabela 5.3: Estimativas dos parâmetros e erros padrão : modelo II

| Parâmetros | $\beta_1$  | $\beta_2$      | $\beta_3$   | $\beta_4$ | $\gamma_1$ | $\gamma_2$     |
|------------|------------|----------------|-------------|-----------|------------|----------------|
| Descrição  | Intercepto | Corrente de ar | Temperatura | Produto   | Intercepto | Corrente de ar |
| EMV        | -12.223    | 0.119          | 0.232       | -0.003    | 20.071     | -0.198         |
| EP         | 2.432      | 0.042          | 0.122       | 0.002     | 2.108      | 0.035          |

Testamos a hipótese  $\mathcal{H}_0 : \beta_4 = 0$ . A estatística do teste da razão de verossimilhanças ( $w$ ) é igual a 2.22 (valor- $p = 0.14$ ), e as duas estatísticas ajustadas são  $w^* = 1.18$  (valor- $p = 0.28$ ) e  $w^{**} = 1.30$  (valor- $p = 0.25$ ). Os três testes levam a não rejeição de  $\mathcal{H}_0$  aos níveis de significância usuais.

Removemos o efeito do produto ( $x_2 x_3$ ) do componente sistemático da média (como sugerido pelos três testes) e assumimos o modelo com componente sistemático da média dado por

$$\log \frac{\mu_t}{1 - \mu_t} = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} \quad (5.1)$$

e especificação da precisão dada por

$$\log(\phi_t) = \gamma_1 + \gamma_2 x_{t2}, \quad (5.2)$$

o qual chamamos de ‘modelo III’. As estimativas de máxima verossimilhança dos parâmetros para o ‘modelo III’ estão na Tabela 5.4.

Testamos a hipótese  $\mathcal{H}_0 : \beta_3 = 0$ . A estatística do teste da razão de verossimilhanças ( $w$ ) é igual a 8.50 (valor- $p = 0.00$ ), e as duas estatísticas ajustadas são  $w^* = 6.45$  (valor- $p = 0.01$ ) e  $w^{**} = 6.57$  (valor- $p$

Tabela 5.4: Estimativas dos parâmetros e erros padrão : modelo III

| Parâmetros        |     | $\beta_1$  | $\beta_2$      | $\beta_3$   | $\gamma_1$ | $\gamma_2$     |
|-------------------|-----|------------|----------------|-------------|------------|----------------|
| Descrição         |     | Intercepto | Corrente de ar | Temperatura | Intercepto | Corrente de ar |
| Dados completos   | EMV | -8.671     | 0.058          | 0.051       | 21.070     | -0.217         |
|                   | EP  | 0.275      | 0.006          | 0.015       | 2.066      | 0.034          |
| Sem observação 21 | EMV | -8.683     | 0.061          | 0.045       | 20.405     | -0.204         |
|                   | EP  | 0.264      | 0.006          | 0.015       | 2.122      | 0.035          |

= 0.01). A inferência baseada no teste da razão de verossimilhanças rejeita a hipótese nula, assim como as estatísticas ajustadas, a nível nominal de 5%. Nesse mesmo modelo testamos a hipótese  $\mathcal{H}_0 : \beta_2 = 0$ . Obtivemos  $w = 20.70$  (valor- $p = 0.00$ ),  $w^* = 24.27$  (valor- $p = 0.00$ ) e  $w^{**} = 24.43$  (valor- $p = 0.00$ ). A inferência baseada nos três testes leva à rejeição da hipótese nula.

Uma análise de diagnóstico do ajuste do ‘modelo III’ aos dados foi elaborada (não mostrada aqui) e não foram observados indícios de má especificação do modelo. É de se destacar que a observação 21, que se mostrou altamente influente no modelo considerado por Atkinson (1985), não exerce alta influência no ajuste do modelo de regressão beta (5.1)-(5.2). De fato, a eliminação desta observação não altera significativamente as estimativas dos parâmetros como mostra a Tabela 5.4.

## Capítulo 6

### Considerações finais

A regressão beta pode ser uma opção apropriada para modelar proporções contínuas, além de ser bastante flexível para acomodar vários tipos de aplicações. A distribuição beta pode ser parametrizada em termos de sua média e um parâmetro de precisão. Estes, por sua vez, podem ser modelados usando diferentes tipos de função de ligação, preditores lineares ou não lineares e covariadas disponíveis. A inferência nestes modelos se baseia usualmente em teoria assintótica, que é frequentemente enganosa quando a amostra é pequena. É útil, portanto, deduzir ferramentas de inferência em regressão beta que sejam confiáveis mesmo em amostras pequenas ou de tamanho moderado. Uma estratégia é buscar ajustes para a estatística da razão de verossimilhanças que possibilitem uma boa aproximação entre a distribuição da estatística de teste e a distribuição assintótica. Neste trabalho foi seguida a proposta de Skovgaard (2001). Foram obtidos dois ajustes para a estatística da razão de verossimilhanças em modelos de regressão beta lineares e não lineares com precisão variável.

Resumimos nos itens a seguir as principais contribuições desse trabalho:

1. Apresentamos o modelo de regressão beta linear proposto por Ferrari e Cribari-Neto (2004) e generalizações deste modelo para acomodar precisão variável e especificações não lineares para a estrutura de regressão da média e da precisão. Obtivemos a função *score* e as matrizes de informação observada e de Fisher para esses modelos.
2. Obtivemos as quantidades necessárias para o cálculo das estatísticas de razão de verossimilhanças ajustadas propostas por Skovgaard ( $w^*$  e  $w^{**}$ ) para os modelos citados acima. A obtenção de tais quantidades é bem menos trabalhosa que a das correções de Bartlett. É de se destacar que nossos resultados são bem gerais. Por exemplo, no caso de modelos de regressão beta com precisão variável, os resultados não se limitam a testes de heteroscedasticidade (precisão variável), mas abrangem testes de quaisquer hipóteses nulas que estabelecem um vetor de valores fixados para uma parte do vetor de parâmetros do modelo.
3. Apresentamos resultados de simulações de Monte Carlo implementadas na linguagem Ox. O programa de simulação para o caso em que a precisão é constante foi obtido pela inclusão de algumas linhas de comando no programa apresentado em Oliveira (2004). Quanto ao programa para simulação para o caso em que a precisão é variável, há alterações estruturais significativas em relação ao anterior.

4. Os resultados das simulações mostraram que o teste da razão de verossimilhanças pode ser consideravelmente liberal se a amostra não for grande o suficiente para garantir a proximidade entre a distribuição da estatística do teste e a distribuição  $\chi^2$  limite. As simulações indicaram que os testes ajustados propostos são mais confiáveis, em termos de controle da probabilidade de erro de tipo I, do que o teste da razão de verossimilhanças original. As estatísticas ajustadas propostas, sob a hipótese nula, têm distribuição muito próxima da distribuição  $\chi^2$  de referência. Consequentemente, as taxas de rejeição da hipótese nula dos testes que usam as estatísticas ajustadas são próximos dos respectivos níveis nominais, com uma pequena vantagem para o teste baseado em  $w^*$ .
5. Simulações de poder dos testes mostram que o teste baseado na estatística  $w^{**}$  é ligeiramente mais poderoso do que aquele que usa  $w^*$ , no entanto, o primeiro é um pouco mais liberal.
6. Uma aplicação do modelo de regressão beta linear com precisão variável a um conjunto de dados reais é apresentada. A aplicação ilustra o fato de que o modelo de regressão beta é uma opção apropriada para ajustar dados que têm seu limite de variação no intervalo unitário  $(0, 1)$ . Mostramos, ainda, que alteração inferencial pode ocorrer dependendo da estatística em que esta se baseia. É recomendável que, quando o teste da razão de verossimilhanças produz valor- $p$  assintótico não muito distante dos níveis nominais usuais, as estatísticas ajustadas sejam utilizadas, especialmente se o tamanho da amostra não for muito grande. As dificuldades computacionais para a produção das estatísticas ajustadas para um conjunto de dados reais limitam-se a implementar a leitura destes no programa utilizado para as simulações.

Em resumo, recomendamos o uso do teste baseado na estatística de razão de verossimilhanças ajustada  $w^*$ .

Em trabalho futuro pretendemos elaborar estudos de simulação semelhantes aos que foram apresentados nessa dissertação, mas para modelos de regressão beta não lineares. Ainda, pretendemos implementar os resultados no software R (Ihaka e Gentleman, 1996).

## Apêndice A

### Cálculo da função escore e das matrizes de informação

Temos, por (2.1), que

$$\begin{aligned}\ell_t(\mu_t, \phi_t) &= \log f(y_t; \mu_t, \phi_t) \\ &= \log \left( \frac{\Gamma(\phi_t)}{\Gamma(\mu_t \phi_t) \Gamma((1 - \mu_t) \phi_t)} \right) + (\mu_t \phi_t - 1) \log y_t + ((1 - \mu_t) \phi_t - 1) \log(1 - y_t) \\ &= \log \left( \frac{\Gamma(\phi_t)}{\Gamma(\mu_t \phi_t) \Gamma((1 - \mu_t) \phi_t)} \right) + (\mu_t \phi_t - 1) y_t^* + (\phi_t - 2) y_t^\dagger \\ &= a_t + (\mu_t \phi_t - 1)(y_t^* - \mu_t^*) + (\phi_t - 2)(y_t^\dagger - \mu_t^\dagger)\end{aligned}$$

com  $a_t$  definido como para (2.6).

Podemos escrever (2.1) como

$$f(y; \mu, \phi) = \exp \left\{ \mu \phi y^* + \phi y^\dagger + \log \left( \frac{\Gamma(\phi)}{\Gamma(\mu \phi) \Gamma((1 - \mu) \phi)} \right) \right\} \frac{1}{y(1 - y)},$$

que tem a forma

$$p(x|\vartheta) = \exp \left[ \sum_{i=1}^s \vartheta_i T_i(x) - A(\vartheta) \right] h(x),$$

com  $(\vartheta_1, \vartheta_2) = (\mu \phi, \phi)$ ,  $(T_1(y), T_2(y)) = (y^*, y^\dagger)$  e  $A(\vartheta_1, \vartheta_2) = \log[\Gamma(\vartheta_1) \Gamma(\vartheta_2 - \vartheta_1) / \Gamma(\vartheta_2)]$ . Então a densidade pertence à família exponencial bidimensional. Lehmann e Casella (1998, p.27) mostram que se a densidade tem a forma acima, então

$$E_{\vartheta}(T_j) = \frac{\partial}{\partial \vartheta_j} A(\vartheta)$$

e

$$\text{cov}_{\vartheta}(T_j, T_k) = \frac{\partial^2}{\partial \vartheta_j \partial \vartheta_k} A(\vartheta).$$

Segue que

$$\mu^* = E_{\vartheta}(y^*) = \psi(\mu \phi) - \psi((1 - \mu) \phi)$$

e

$$\mu^\dagger = E_{\vartheta}(y^\dagger) = \psi((1 - \mu) \phi) - \psi(\phi)$$

com  $\psi(\cdot)$  definido como no Capítulo 2.

Ainda

$$\begin{aligned} v^* &= \text{cov}(y^*, y^*) = \psi'(\mu\phi) + \psi'((1-\mu)\phi), \\ v^\dagger &= \text{cov}(y^\dagger, y^\dagger) = \psi'((1-\mu)\phi) - \psi'(\phi), \\ c^{*\dagger} &= \text{cov}(y^*, y^\dagger) = -\psi'((1-\mu)\phi). \end{aligned}$$

onde  $\psi'(\cdot)$  é a função trigama.

É fácil obter que

$$\begin{aligned} \frac{\partial \mu^*}{\partial \mu} &= \phi v^*, \\ \frac{\partial \mu^\dagger}{\partial \mu} &= \phi c^{*\dagger}, \\ \frac{\partial \mu^*}{\partial \phi} &= \mu v^* + c^{*\dagger} \end{aligned}$$

e

$$\frac{\partial \mu^\dagger}{\partial \phi} = v^\dagger + \mu c^{*\dagger}.$$

As derivadas parciais do logaritmo da função densidade (2.5) são:

$$\frac{\partial l_t(\mu_t, \phi_t)}{\partial \mu_t} = \phi_t(y_t^* - \mu_t^*), \quad (\text{A.1})$$

$$\begin{aligned} \frac{\partial^2 l_t}{\partial \mu_t^2} &= -\phi_t \frac{\partial \mu_t^*}{\partial \mu_t} \\ &= -\phi_t^2 v_t^*, \end{aligned} \quad (\text{A.2})$$

$$\frac{\partial l_t(\mu_t, \phi_t)}{\partial \phi_t} = \mu_t(y_t^* - \mu_t^*) + (y_t^\dagger - \mu_t^\dagger), \quad (\text{A.3})$$

$$\begin{aligned} \frac{\partial^2 l_t}{\partial \phi_t^2} &= -\mu_t \frac{\partial \mu_t^*}{\partial \phi_t} - \frac{\partial \mu_t^\dagger}{\partial \phi_t} \\ &= -\mu_t(\mu_t v_t^* + c_t^{*\dagger}) - (v_t^\dagger + \mu_t c_t^{*\dagger}) \\ &= -\mu_t^2 v_t^* - 2\mu_t c_t^{*\dagger} - v_t^\dagger \end{aligned} \quad (\text{A.4})$$

e

$$\begin{aligned}
\frac{\partial^2 l_t}{\partial \mu_t \partial \phi_t} &= \frac{\partial}{\partial \phi_t} (\phi_t (y_t^* - \mu_t^*)) \\
&= (y_t^* - \mu_t^*) + \phi_t \frac{\partial}{\partial \phi_t} (y_t^* - \mu_t^*) \\
&= (y_t^* - \mu_t^*) - \phi_t (\mu_t v_t^* + c_t^{*\dagger}).
\end{aligned} \tag{A.5}$$

E ainda, por (2.2) e (2.3),

$$\frac{\partial^2 \mu_t}{\partial \eta_t \partial \mu_t} = \frac{\partial}{\partial \mu_t} \left( \frac{\partial \mu_t}{\partial \eta_t} \right) = \frac{\partial}{\partial \mu_t} \left( \frac{1}{\frac{\partial \eta_t}{\partial \mu_t}} \right) = \frac{\partial}{\partial \mu_t} \left( \frac{1}{g'(\mu_t)} \right) = -\frac{g''(\mu_t)}{g'(\mu_t)^2} \tag{A.6}$$

e, analogamente,

$$\frac{\partial^2 \phi_t}{\partial \delta_t \partial \phi_t} = \frac{\partial}{\partial \phi_t} \left( \frac{\partial \phi_t}{\partial \eta_t} \right) = \frac{\partial}{\partial \phi_t} \left( \frac{1}{\frac{\partial \eta_t}{\partial \phi_t}} \right) = \frac{\partial}{\partial \phi_t} \left( \frac{1}{h'(\phi_t)} \right) = -\frac{h''(\phi_t)}{h'(\phi_t)^2}. \tag{A.7}$$

De (2.2) e (2.4) obtemos, para  $i = 1, \dots, k$ , que

$$\frac{\partial l(\beta, \phi)}{\partial \beta_i} = \sum_{t=1}^n \frac{\partial l_t(\mu_t, \phi_t)}{\partial \beta_i} = \sum_{t=1}^n \left\{ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_i} \right\},$$

ou seja, substituindo (A.1) na equação acima e obtendo as outras derivadas parciais a partir de (2.2), temos que

$$\frac{\partial l(\beta, \gamma)}{\partial \beta_i} = \phi_t \sum_{t=1}^n \left\{ (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} x_{t_i} \right\},$$

cujas forma matricial é

$$U_\beta(\beta, \phi) = X^\top T \Phi (y^* - \mu^*).$$

De (2.3) e (2.4) obtemos, para  $i = 1, \dots, k$ , que

$$\frac{\partial l(\beta, \gamma)}{\partial \gamma_i} = \sum_{t=1}^n \frac{\partial l_t(\mu_t, \phi_t)}{\partial \gamma_i} = \sum_{t=1}^n \left\{ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial \delta_t}{\partial \gamma_i} \right\},$$

ou seja, substituindo (A.3) na equação acima e obtendo as outras derivadas parciais a partir de (2.3) temos que

$$\frac{\partial l(\beta, \gamma)}{\partial \gamma_i} = \sum_{t=1}^n \left\{ [\mu_t (y_t^* - \mu_t^*) + (y_t^\dagger - \mu_t^\dagger)] \frac{1}{h'(\phi_t)} z_{t_i} \right\},$$

cujas forma matricial é

$$U_\gamma(\beta, \gamma) = Z^\top H [\mathcal{M}(y^* - \mu^*) + (y_t^\dagger - \mu_t^\dagger)].$$

De (2.2) e (2.4) obtemos, para  $i, j = 1, \dots, k$ , que

$$\begin{aligned}
\frac{\partial^2 l(\beta, \phi)}{\partial \beta_i \partial \beta_j} &= \sum_{t=1}^n \left\{ \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \beta_i \partial \beta_j} \right\} \\
&= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \beta_j} \left[ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_i} \right] \right\} \\
&= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \mu_t} \left[ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \right] \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_j} x_{ti} \right\} \\
&= \sum_{t=1}^n \left\{ \left[ \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \mu_t^2} \frac{\partial \mu_t}{\partial \eta_t} + \frac{\partial l_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{\partial^2 \mu_t}{\partial \eta_t \partial \mu_t} \right] \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_j} x_{ti} \right\}.
\end{aligned}$$

Substituindo (A.1), (A.2) e (A.6) na equação acima e obtendo as outras derivadas parciais a partir de (2.2) temos que

$$\frac{\partial^2 l(\beta, \gamma)}{\partial \beta_i \partial \beta_j} = \sum_{t=1}^n \left\{ \left[ -\phi_t^2 v_t^* \frac{1}{g'(\mu_t)} + \phi_t (y_t * -\mu_t^*) \left( \frac{-g''(\mu_t)}{g'(\mu_t)^2} \right) \right] \frac{1}{g'(\mu_t)} x_{tj} x_{ti} \right\}.$$

Logo,

$$J_{\beta\beta} = -\frac{\partial^2 l(\beta, \gamma)}{\partial \beta \partial \beta^T}$$

tem a seguinte forma matricial

$$J_{\beta\beta} = X^T [\Phi T V^* + S T^2 (Y^* - \mathcal{M}^*)] T \Phi X. \quad (\text{A.8})$$

A partir de (2.2), (2.3) e (2.4) obtemos também que

$$\begin{aligned}
\frac{\partial^2 l(\beta, \gamma)}{\partial \beta_i \partial \gamma_j} &= \sum_{t=1}^n \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \beta_i \partial \gamma_j} \\
&= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \gamma_j} \left[ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_i} \right] \right\} \\
&= \sum_{t=1}^n \left\{ \left[ \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \mu_t \partial \gamma_j} \right] \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_i} \right\} \\
&= \sum_{t=1}^n \left\{ \left[ \frac{\partial}{\partial \phi_t} \left( \frac{\partial l_t}{\partial \mu_t} \right) \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial \delta_t}{\partial \gamma_j} \right] \frac{1}{g'(\mu_t)} \frac{\partial \eta_t}{\partial \beta_i} \right\} \\
&= \sum_{t=1}^n \left\{ \left[ \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \mu_t \partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial \delta_t}{\partial \gamma_j} \right] \frac{1}{g'(\mu_t)} x_{ti} \right\}.
\end{aligned}$$

Substituindo (A.5) na equação acima e obtendo as outras derivadas parciais a partir de (2.2) e (2.3) concluímos que

$$\frac{\partial^2 l(\beta, \gamma)}{\partial \beta_i \partial \gamma_j} = \sum_{t=1}^n \left\{ \left[ (y^* - \mu_t^*) - \phi_t (\mu_t v_t^* + c^{*\dagger}) \right] \frac{1}{h'(\phi_t)} z_{tj} \frac{1}{g'(\mu_t)} x_{ti} \right\}.$$



Logo,

$$J_{\beta\gamma} = J_{\beta\gamma}^\top = -\frac{\partial^2 l(\beta, \gamma)}{\partial \beta \partial \gamma^\top}$$

tem a seguinte forma matricial

$$J_{\beta\gamma} = -X^\top [(Y^* - \mathcal{M}^*) - \Phi(MV^* + C)] THZ. \quad (\text{A.9})$$

De (2.3) e (2.4) obtemos

$$\begin{aligned} \frac{\partial^2 l(\beta, \phi)}{\partial \gamma_i \partial \gamma_j} &= \sum_{t=1}^n \left\{ \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \gamma_i \partial \gamma_j} \right\} \\ &= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \gamma_j} \left[ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial \delta_t}{\partial \gamma_i} \right] \right\} \\ &= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \phi_t} \left[ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \right] \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial \delta_t}{\partial \gamma_j} z_{ti} \right\} \\ &= \sum_{t=1}^n \left\{ \left[ \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \phi_t^2} \frac{\partial \phi_t}{\partial \delta_t} + \frac{\partial l_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial^2 \phi_t}{\partial \delta_t \partial \phi_t} \right] \frac{1}{h'(\phi_t)} z_{tj} z_{ti} \right\}. \end{aligned}$$

Substituindo (A.3), (A.4) e (A.7) na equação acima e obtendo as outras derivadas parciais a partir da equação (2.3) obtemos

$$\begin{aligned} \frac{\partial^2 l(\beta, \gamma)}{\partial \gamma_i \partial \gamma_j} &= \\ \sum_{t=1}^n \left\{ \left[ (-\mu_t^2 v_t^* - 2\mu_t c_t^{*\dagger} - v_t^\dagger) \frac{1}{h'(\phi_t)} + (\mu_t(y_t^* - \mu_t^*) + (y_t^\dagger - \mu_t^\dagger)) \left( -\frac{h''(\phi_t)}{h'(\phi_t)^2} \right) \right] \frac{1}{h'(\phi_t)} z_{tj} z_{ti} \right\}. \end{aligned}$$

Logo,

$$J_{\gamma\gamma} = -\frac{\partial^2 l(\beta, \gamma)}{\partial \gamma \partial \gamma^\top}$$

tem a seguinte forma matricial

$$J_{\gamma\gamma} = Z^\top \left[ H(\mathcal{M}^2 V^* + 2\mathcal{M}C + V^\dagger) + (\mathcal{M}(Y^* - \mathcal{M}^*) + (Y^\dagger - \mathcal{M}^\dagger)) Q H^2 \right] HZ. \quad (\text{A.10})$$

Sendo  $E_\omega(Y^* - \mathcal{M}^*) = 0$  e  $E_\omega(Y^\dagger - \mathcal{M}^\dagger) = 0$ , concluímos a partir de (A.8), (A.9) e (A.10) que a informação de Fisher,

$$I = \begin{pmatrix} I_{\beta\beta} & I_{\beta\gamma} \\ I_{\gamma\beta} & I_{\gamma\gamma} \end{pmatrix} = E_\omega(J) = E_\omega \begin{pmatrix} J_{\beta\beta} & J_{\beta\gamma} \\ J_{\gamma\beta} & J_{\gamma\gamma} \end{pmatrix}$$

é

$$I = \begin{pmatrix} X^\top \Phi T V^* T \Phi X & X^\top \Phi (M V^* + C) T H Z \\ (X^\top \Phi (M V^* + C) T H Z)^\top & Z^\top H (\mathcal{M}^2 V^* + 2\mathcal{M}C + V^\dagger) H Z \end{pmatrix}.$$

Consideramos, por fim, o modelo de regressão não linear. De (2.9) e (2.4) obtemos, para  $i = 1, \dots, k$ ,

que

$$\frac{\partial l(\beta, \phi)}{\partial \beta_i} = \sum_{t=1}^n \frac{\partial l_t(\mu_t, \phi_t)}{\partial \beta_i} = \sum_{t=1}^n \left\{ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_i} \right\},$$

ou seja, substituindo (A.1) na equação acima e obtendo as outras derivadas parciais a partir de (2.9), temos que

$$\frac{\partial l(\beta, \gamma)}{\partial \beta_i} = \phi_t \sum_{t=1}^n \left\{ (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} \frac{\partial \eta_t}{\partial \beta_i} \right\},$$

cuja forma matricial é

$$U_\beta(\beta, \phi) = \mathcal{X}^\top T \Phi (y^* - \mu^*).$$

De (2.10) e (2.4) obtemos, para  $i = 1, \dots, k$ , que

$$\frac{\partial l(\beta, \gamma)}{\partial \gamma_i} = \sum_{t=1}^n \frac{\partial l_t(\mu_t, \phi_t)}{\partial \gamma_i} = \sum_{t=1}^n \left\{ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial \delta_t}{\partial \gamma_i} \right\},$$

ou seja, substituindo (A.3) na equação acima e obtendo as outras derivadas parciais a partir de (2.10) temos que

$$\frac{\partial l(\beta, \gamma)}{\partial \gamma_i} = \sum_{t=1}^n \left\{ [\mu_t (y_t^* - \mu_t^*) + (y_t^\dagger - \mu_t^\dagger)] \frac{1}{h'(\phi_t)} \frac{\partial \delta_t}{\partial \gamma_i} \right\},$$

cuja forma matricial é

$$U_\gamma(\beta, \gamma) = \mathcal{Z}^\top H [\mathcal{M}(y^* - \mu^*) + (y^\dagger - \mu^\dagger)].$$

De (2.9) e (2.4) obtemos, para  $i, j = 1, \dots, k$ , que

$$\begin{aligned} \frac{\partial^2 l(\beta, \phi)}{\partial \beta_i \partial \beta_j} &= \sum_{t=1}^n \left\{ \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \beta_i \partial \beta_j} \right\} \\ &= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \beta_j} \left[ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_i} \right] \right\} \\ &= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \beta_j} \left[ \frac{\partial l_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \right] \frac{\partial \eta_t}{\partial \beta_i} + \frac{\partial l_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial}{\partial \beta_j} \left[ \frac{\partial \eta_t}{\partial \beta_i} \right] \right\} \\ &= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \mu_t} \left[ \frac{\partial l_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \right] \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_j} \frac{\partial \eta_t}{\partial \beta_i} + \frac{\partial l_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial^2 \eta_t}{\partial \beta_i \partial \beta_j} \right\} \\ &= \sum_{t=1}^n \left\{ \left( \frac{\partial^2 l_t}{\partial \mu_t^2} \frac{\partial \mu_t}{\partial \eta_t} + \frac{\partial l_t}{\partial \mu_t} \frac{\partial \mu_t^2}{\partial \eta \partial \mu_t} \right) \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_j} \frac{\partial \eta_t}{\partial \beta_i} + \frac{\partial l_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial^2 \eta_t}{\partial \beta_i \partial \beta_j} \right\}. \end{aligned}$$

Substituindo (A.1), (A.2) e (A.6) na equação acima e obtendo as outras derivadas parciais a partir de (2.9) temos que

$$\frac{\partial^2 l(\beta, \phi)}{\partial \beta_i \partial \beta_j} =$$

$$\sum_{t=1}^n \left\{ \left[ -\phi_t^2 v_t^* \frac{1}{g'(\mu_t)} + \phi_t (y_t^* - \mu_t^*) \left( -\frac{g''(\mu_t)}{(g'(\mu_t))^2} \right) \right] \frac{1}{g'(\mu_t)} \frac{\partial \eta_t}{\partial \beta_j} \frac{\partial \eta_t}{\partial \beta_i} + \phi_t (y_t^* - \mu_t^*) \frac{1}{g'(\mu_t)} \frac{\partial^2 \eta_t}{\partial \beta_i \partial \beta_j} \right\}.$$

Logo,

$$J_{\beta\beta} = -\frac{\partial^2 l(\beta, \gamma)}{\partial \beta \partial \beta^\top}$$

tem a seguinte forma matricial

$$J_{\beta\beta} = \mathcal{X}^\top \{ \Phi T V^* + S T^2 (Y^* - \mathcal{M}^*) \} T \Phi \mathcal{X} + [(y^* - \mu^*)^\top \Phi T] [\dot{\mathcal{X}}];$$

a definição do produto colchete  $[\cdot][\cdot]$  é dada na Seção 2.2. A partir de (2.9), (2.10) e (2.4) obtemos também que

$$\begin{aligned} \frac{\partial^2 l(\beta, \gamma)}{\partial \beta_i \partial \gamma_j} &= \sum_{t=1}^n \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \beta_i \partial \gamma_j} \\ &= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \gamma_j} \left[ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_i} \right] \right\} \\ &= \sum_{t=1}^n \left\{ \left[ \frac{\partial l_t^2(\mu_t, \phi_t)}{\partial \mu_t \partial \gamma_j} \right] \frac{\partial \mu_t}{\partial \eta_t} \frac{\partial \eta_t}{\partial \beta_i} \right\} \\ &= \sum_{t=1}^n \left\{ \left[ \frac{\partial}{\partial \phi_t} \left( \frac{\partial l_t}{\partial \mu_t} \right) \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial \delta_t}{\partial \gamma_j} \right] \frac{1}{g'(\mu_t)} \frac{\partial \eta_t}{\partial \beta_i} \right\} \\ &= \sum_{t=1}^n \left\{ \left[ \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \mu_t \partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial \delta_t}{\partial \gamma_j} \right] \frac{1}{g'(\mu_t)} \frac{\partial \eta_t}{\partial \beta_i} \right\}. \end{aligned}$$

Substituindo (A.5) na equação acima e obtendo as outras derivadas parciais a partir de (2.9) e (2.10) concluímos que

$$\frac{\partial^2 l(\beta, \gamma)}{\partial \beta_i \partial \gamma_j} = \sum_{t=1}^n \left\{ \left[ y_t^* - \mu_t^* - \phi_t (\mu_t v_t^* + c_t^{*\dagger}) \right] \frac{1}{h'(\phi_t)} \frac{\partial \delta_t}{\partial \gamma_j} \frac{1}{g'(\mu_t)} \frac{\partial \eta_t}{\partial \beta_i} \right\}.$$

Logo,

$$J_{\beta\gamma} = J_{\beta\gamma}^\top = -\frac{\partial^2 l(\beta, \gamma)}{\partial \beta \partial \gamma^\top}$$

tem a seguinte forma matricial

$$J_{\beta\gamma} = J_{\gamma\beta}^\top = -\mathcal{X}^\top \{ (Y^* - \mathcal{M}^*) - \Phi (M V^* + C) \} T H Z.$$

De (2.10) e (2.4) obtemos

$$\begin{aligned}
\frac{\partial^2 l(\beta, \phi)}{\partial \gamma_i \partial \gamma_j} &= \sum_{t=1}^n \left\{ \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \gamma_i \partial \gamma_j} \right\} \\
&= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \gamma_j} \left[ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial \delta_t}{\partial \gamma_i} \right] \right\} \\
&= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \gamma_j} \left( \frac{\partial l_t}{\partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \right) \frac{\partial \delta_t}{\partial \gamma_i} + \frac{\partial l_t}{\partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial}{\partial \gamma_j} \left( \frac{\partial \delta_t}{\partial \gamma_i} \right) \right\} \\
&= \sum_{t=1}^n \left\{ \frac{\partial}{\partial \phi_t} \left[ \frac{\partial l_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \right] \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial \delta_t}{\partial \gamma_j} \frac{\partial \delta_t}{\partial \gamma_i} + \frac{\partial l_t}{\partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial^2 \delta_t}{\partial \gamma_i \partial \gamma_j} \right\} \\
&= \sum_{t=1}^n \left\{ \left[ \frac{\partial^2 l_t(\mu_t, \phi_t)}{\partial \phi_t^2} \frac{\partial \phi_t}{\partial \delta_t} + \frac{\partial l_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{\partial^2 \phi_t}{\partial \delta_t \partial \phi_t} \right] \frac{1}{h'(\phi_t)} \frac{\partial \delta_t}{\partial \gamma_j} \frac{\partial \delta_t}{\partial \gamma_i} + \frac{\partial l_t}{\partial \phi_t} \frac{\partial \phi_t}{\partial \delta_t} \frac{\partial^2 \delta_t}{\partial \gamma_i \partial \gamma_j} \right\}.
\end{aligned}$$

Substituindo (A.3), (A.4) e (A.7) na equação acima e obtendo as outras derivadas parciais a partir da equação (2.10) obtemos

$$\begin{aligned}
\frac{\partial^2 l(\beta, \gamma)}{\partial \gamma_i \partial \gamma_j} &= \\
\sum_{t=1}^n \left\{ \left[ \left( -\mu_t^2 v_t^* - 2\mu_t c_t^{*\dagger} - v_t^\dagger \right) \frac{1}{h'(\phi_t)} + (\mu_t (y_t^* - \mu_t^*) + (y_t^\dagger - \mu_t^\dagger)) \left( -\frac{h''(\phi_t)}{h'(\phi_t)^2} \right) \right] \frac{1}{h'(\phi_t)} \frac{\partial \delta_t}{\partial \gamma_j} \frac{\partial \delta_t}{\partial \gamma_i} \right\} \\
&\quad + \sum_{t=1}^n \left\{ (\mu_t (y_t^* - \mu_t^*) + (y_t^\dagger - \mu_t^\dagger)) \frac{1}{h'(\phi_t)} \frac{\partial^2 \delta_t}{\partial \gamma_i \partial \gamma_j} \right\}.
\end{aligned}$$

Logo,

$$J_{\gamma\gamma} = -\frac{\partial^2 l(\beta, \gamma)}{\partial \gamma \partial \gamma^\top}$$

tem a seguinte forma matricial

$$\begin{aligned}
J_{\gamma\gamma} &= \mathcal{Z}^\top \{ H(\mathcal{M}^2 V^* + 2\mathcal{M}C + V^\dagger) + \{ \mathcal{M}(Y^* - \mathcal{M}^*) + (Y^\dagger - \mathcal{M}^\dagger) \} QH^2 \} H \mathcal{Z} \\
&\quad + [(y^* - \mu^*)^\top \mathcal{M} + (y^\dagger - \mu^\dagger)^\top H][\dot{\mathcal{Z}}].
\end{aligned}$$

## Apêndice B

### Cálculo de $q$ e $\Upsilon$

A quantidade (3.1) pode ser escrita como

$$q = E_{\omega_1}[U(\omega_1)(l(\omega_1) - l(\omega))] = \begin{bmatrix} E_{\omega_1}[U_\beta(\omega_1)l(\omega_1)] - E_{\omega_1}[U_\beta(\omega_1)l(\omega)] \\ E_{\omega_1}[U_\gamma(\omega_1)l(\omega_1)] - E_{\omega_1}[U_\gamma(\omega_1)l(\omega)] \end{bmatrix}.$$

Por (2.6) e (2.7) e lembrando que  $\mu^* = E_\omega(y^*)$ , onde  $\omega = (\beta^\top, \gamma^\top)$ , temos que

$$\begin{aligned} E_\omega[U_\beta(\omega)l(\omega)] &= E_\omega\{X^\top T\Phi(y^* - \mu^*)[(y^* - \mu^*)^\top(\Phi\mathcal{M} - I) + (y^\dagger - \mu^\dagger)(\Phi - 2I) + a^\top]\}_\iota \\ &= X^\top T\Phi\{E_\omega[(y^* - \mu^*)(y^* - \mu^*)^\top](\Phi\mathcal{M} - I) + E_\omega[(y^* - \mu^*)(y^\dagger - \mu^\dagger)^\top](\Phi - 2I) \\ &\quad + E_\omega[(y^* - \mu^*)a^\top]\}_\iota \\ &= X^\top T\Phi\{V^*(\Phi\mathcal{M} - I) + C(\Phi - 2I)\}_\iota. \end{aligned}$$

Então,

$$\begin{aligned} E_{\omega_1}[U_\beta(\omega_1)l(\omega_1)] - E_{\omega_1}[U_\beta(\omega_1)l(\omega)] &= X^\top T^{(1)}\Phi^{(1)}\{V^{*(1)}(\Phi^{(1)}\mathcal{M}^{(1)} - I) + C^{(1)}(\Phi^{(1)} - 2I)\}_\iota \\ &\quad - X^\top T^{(1)}\Phi^{(1)}\{V^{*(1)}(\Phi\mathcal{M} - I) + C^{(1)}(\Phi - 2I)\}_\iota \\ &= X^\top T^{(1)}\Phi^{(1)}\{V^{*(1)}(\Phi^{(1)}\mathcal{M}^{(1)} - \Phi\mathcal{M}) + C^{(1)}(\Phi^{(1)} - \Phi)\}_\iota, \end{aligned}$$

pois se  $t \neq u$ ,  $y_t$  é independente de  $y_u$ ,  $E_{\omega_1}\{(y_t^* - \mu_t^{*(1)})(y_u - \mu_u^*)\} = E_{\omega_1}\{y_t^* - \mu_t^{*(1)}\}E_{\omega_1}\{y_u - \mu_u^*\}$  e  $E_{\omega_1}\{y_t^* - \mu_t^{*(1)}\} = 0$ . Ainda

$$\begin{aligned} E_{\omega_1}\{(y_t^* - \mu_t^{*(1)})(y_t - \mu_t^*)\} &= E_{\omega_1}\{(y_t^* - \mu_t^{*(1)})(y_t - \mu_t^{*(1)} + \mu_t^{*(1)} - \mu_t^*)\} \\ &= E_{\omega_1}\{(y_t^* - \mu_t^{*(1)})(y_t - \mu_t^{*(1)})\} + E_{\omega_1}\{(y_t^* - \mu_t^{*(1)})\}(\mu_t^{*(1)} - \mu_t^*) \\ &= E_{\omega_1}\{(y_t^* - \mu_t^{*(1)})^2\} = v_t^{*(1)} \end{aligned}$$

pois  $E_{\omega_1}\{(y_t^* - \mu_t^{*(1)})\}(\mu_t^{*(1)} - \mu_t^*) = 0$ . Por (2.6) e (2.8) temos que

$$\begin{aligned}
E_{\omega}[U_{\gamma}(\omega)l(\omega)] &= E_{\omega}\{Z^{\top}H[\mathcal{M}(y^* - \mu^*) + (y^{\dagger} - \mu^{\dagger})][(y^* - \mu^*)^{\top}(\Phi\mathcal{M} - I) \\
&\quad + (y^{\dagger} - \mu^{\dagger})^{\top}(\Phi - 2I) + a^{\top}]\}_{\iota} \\
&= Z^{\top}H\{\mathcal{M}E_{\omega}[(y^* - \mu^*)(y^* - \mu^*)^{\top}](\Phi\mathcal{M} - I) \\
&\quad + \mathcal{M}E_{\omega}[(y^* - \mu^*)(y^{\dagger} - \mu^{\dagger})^{\top}](\Phi - 2I) \\
&\quad + E_{\omega}[(y^{\dagger} - \mu^{\dagger})(y^* - \mu^*)^{\top}](\Phi\mathcal{M} - I) \\
&\quad + E_{\omega}[(y^{\dagger} - \mu^{\dagger})(y^{\dagger} - \mu^{\dagger})^{\top}](\Phi - 2I)\}_{\iota} \\
&= Z^{\top}H\{\mathcal{M}V^*(\Phi\mathcal{M} - I) + \mathcal{M}C(\Phi - 2I) + C(\Phi\mathcal{M} - 1) + V^{\dagger}(\Phi - 2I)\}_{\iota}.
\end{aligned}$$

Então,

$$\begin{aligned}
E_{\omega_1}[U_{\gamma}(\omega_1)l(\omega_1)] - E_{\omega_1}[U_{\gamma}(\omega_1)l(\omega)] &= Z^{\top}H^{(1)}\{\mathcal{M}^{(1)}V^{*(1)}(\Phi^{(1)}\mathcal{M}^{(1)} - I) + \mathcal{M}^{(1)}C^{(1)}(\Phi^{(1)} - 2I) \\
&\quad + C^{(1)}(\Phi^{(1)}\mathcal{M}^{(1)} - 1) + V^{\dagger(1)}(\Phi^{(1)} - 2I)\}_{\iota} \\
&\quad - [Z^{\top}H^{(1)}\{\mathcal{M}^{(1)}V^{*(1)}(\Phi\mathcal{M} - I) + \mathcal{M}^{(1)}C^{(1)}(\Phi - 2I) \\
&\quad + C^{(1)}(\Phi\mathcal{M} - I) + V^{\dagger(1)}(\Phi - 2I)\}_{\iota}] \\
&= Z^{\top}H^{(1)}\{\mathcal{M}^{(1)}V^{*(1)}(\Phi^{(1)}\mathcal{M}^{(1)} - \Phi\mathcal{M}) \\
&\quad + \mathcal{M}^{(1)}C^{(1)}(\Phi^{(1)} - \Phi) \\
&\quad + C^{(1)}(\Phi^{(1)}\mathcal{M}^{(1)} - \Phi\mathcal{M}) + V^{\dagger(1)}(\Phi^{(1)} - \Phi)\}_{\iota} \\
&= Z^{\top}H^{(1)}\{(\mathcal{M}^{(1)}V^{*(1)} + C^{(1)})(\Phi^{(1)}\mathcal{M}^{(1)} - \Phi\mathcal{M}) \\
&\quad + (\mathcal{M}^{(1)}C^{(1)} + V^{\dagger(1)})(\Phi^{(1)} - \Phi)\}_{\iota}
\end{aligned}$$

Segue que

$$q = \begin{bmatrix} X^{\top}T^{(1)}\Phi^{(1)}\{V^{*(1)}(\Phi^{(1)}\mathcal{M}^{(1)} - \Phi\mathcal{M}) + C^{(1)}(\Phi^{(1)} - \Phi)\}_{\iota} \\ Z^{\top}H^{(1)}\{(\mathcal{M}^{(1)}V^{*(1)} + C^{(1)})(\Phi^{(1)}\mathcal{M}^{(1)} - \Phi\mathcal{M}) + (\mathcal{M}^{(1)}C^{(1)} + V^{\dagger(1)})(\Phi^{(1)} - \Phi)\}_{\iota} \end{bmatrix},$$

onde  $V^* = \text{diag}\{v_1^*, \dots, v_n^*\}$ ,  $V^{\dagger} = \text{diag}\{v_1^{\dagger}, \dots, v_n^{\dagger}\}$  e  $C = \text{diag}\{c_1^{\dagger}, \dots, c_n^{\dagger}\}$ .

A quantidade (3.2) pode ser escrita como

$$\Upsilon = \begin{bmatrix} E_{\omega_1}[U_{\beta}(\omega_1)U_{\beta}^{\top}(\omega)] & E_{\omega_1}[U_{\beta}(\omega_1)U_{\gamma}(\omega)] \\ E_{\omega_1}[U_{\gamma}(\omega_1)U_{\beta}^{\top}(\omega)] & E_{\omega_1}[U_{\gamma}(\omega_1)U_{\gamma}^{\top}(\omega)] \end{bmatrix}.$$

Por (2.7) temos que

$$\begin{aligned}
E_{\omega_1}\{U_{\beta}(\omega_1)U_{\beta}^{\top}(\omega)\} &= E_{\omega_1}\{X^{\top}T^{(1)}\Phi^{(1)}(y^* - \mu^{*(1)})[X^{\top}T\Phi(y^* - \mu^*)]^{\top}\} \\
&= E_{\omega_1}\{X^{\top}T^{(1)}\Phi^{(1)}(y^* - \mu^{*(1)})(y^* - \mu^*)^{\top}\Phi TX\} \\
&= X^{\top}T^{(1)}\Phi^{(1)}V^{*(1)}\Phi TX.
\end{aligned}$$

Por (2.8) temos que

$$\begin{aligned}
\mathbb{E}_{\omega_1}\{U_\gamma(\omega_1)U_\gamma^\top(\omega)\} &= \mathbb{E}_{\omega_1}\{Z^\top H^{(1)}(\mathcal{M}^{(1)}(y^* - \mu^{*(1)}) + (y^\dagger - \mu^{\dagger(1)})) \\
&\quad (Z^\top H(\mathcal{M}(y^* - \mu^*) + (y^\dagger - \mu^\dagger)))^\top\} \\
&= \mathbb{E}_{\omega_1}\{Z^\top H^{(1)}(\mathcal{M}^{(1)}(y^* - \mu^{*(1)}) + (y^\dagger - \mu^{\dagger(1)})) \\
&\quad ((y^* - \mu^*)^\top \mathcal{M} + (y^\dagger - \mu^\dagger)^\top)HZ\} \\
&= Z^\top H^{(1)}\{\mathcal{M}^{(1)}V^{*(1)}\mathcal{M} + (\mathcal{M}^{(1)} + \mathcal{M})C + V^\dagger\}HZ.
\end{aligned}$$

Por (2.7) e (2.8) temos que

$$\begin{aligned}
\mathbb{E}_{\omega_1}\{U_\beta(\omega_1)U_\gamma^\top(\omega)\} &= \mathbb{E}_{\omega_1}\{X^\top T^{(1)}\Phi^{(1)}(y^* - \mu^{*(1)})[Z^\top H(\mathcal{M}(y^* - \mu^*) + (y^\dagger - \mu^\dagger))]^\top\} \\
&= \mathbb{E}_{\omega_1}\{X^\top T^{(1)}\Phi^{(1)}(y^* - \mu^{*(1)})[((y^* - \mu^*)^\top \mathcal{M} + (y^\dagger - \mu^\dagger)^\top)]HZ\} \\
&= X^\top T^{(1)}\Phi^{(1)}\{\mathcal{M}V^{*(1)} + C^{(1)}\}HZ
\end{aligned}$$

e também

$$\begin{aligned}
\mathbb{E}_{\omega_1}\{U_\gamma(\omega_1)U_\beta^\top(\omega)\} &= \mathbb{E}_{\omega_1}\{Z^\top H^{(1)}[\mathcal{M}^{(1)}(y^* - \mu^{*(1)}) + (y^\dagger - \mu^{\dagger(1)})][X^\top T\Phi(y^* - \mu^*)]^\top\} \\
&= \mathbb{E}_{\omega_1}\{Z^\top H^{(1)}[\mathcal{M}^{(1)}(y^* - \mu^{*(1)}) + (y^\dagger - \mu^{\dagger(1)})](y^* - \mu^*)^\top \Phi TX\} \\
&= Z^\top H^{(1)}\{\mathcal{M}^{(1)}V^{*(1)} + C^{(1)}\}\Phi TX.
\end{aligned}$$

Para o modelo de regressão beta não linear, em que as especificações da média e da precisão são dadas em (2.9) e (2.10) respectivamente, é suficiente observar que as matrizes  $X$ ,  $Z$ ,  $\mathcal{X}$  e  $\mathcal{Z}$  não são funções de variáveis aleatórias  $y$ , sendo consideradas como constantes nos cálculos dos componentes de  $q$  e  $\Upsilon$ . Em cada etapa das deduções anteriores é suficiente substituir  $X$  por  $\mathcal{X}$  e  $Z$  por  $\mathcal{Z}$  já que a função escore do modelo linear é similar à do modelo não linear exceto pelo fato de que as matrizes  $X$  e  $Z$  do modelo linear são substituídas por  $\mathcal{X}$  e  $\mathcal{Z}$ , respectivamente, no modelo não linear.





## Apêndice C

### Programas

Apresentamos neste apêndice dois programas que foram desenvolvidos na linguagem de programação Ox. O primeiro é o de simulação que foi utilizado no Capítulo 4 para avaliação de resultados assintóticos em amostras finitas. O segundo programa foi utilizado para a análise do conjunto de dados do exemplo de Oxidação de Amônia apresentado no Capítulo 5. Foram utilizados dois programas para simulação, um para precisão constante e outro para precisão variável. Os dois programas foram desenvolvidos baseados no programa apresentado na dissertação de mestrado de Oliveira (2004). O programa utilizado para simulação com precisão constante foi aqui omitido, devido à grande semelhança com o de Oliveira (2004). Para a análise do conjunto real de dados foram utilizados três programas, um para testar componentes da média quando a precisão é constante, um para testar componentes da precisão e o terceiro para testar componentes da média quando a precisão é variável. Apresentamos aqui apenas este último já que os outros dois são adaptações mais imediatas dos programas de simulação.

```
/* *****  
 * PROGRAMA: simulação.ox  
 *  
 * USO: Avaliação de resultados assintóticos em amostras finitas no modelo de regressão beta :  
 *       Estimativas dos parâmetros e Tamanho de Testes.  
 *       Modelo:      g(mu)=log(mu/(1-mu))=s_mX*beta,      beta=(beta_1,...,beta_k)  
 *                   h(phi)=log(phi)=s_mZ*gama,          gama=(gama_1,...,gama_m)  
 *       Hipótese nula: (gama_(m-r),...,gama_m)=(gama0_(m-r),...,gama0_m)  
 *  
 * AUTOR: Eliane Cantinho Pinheiro.  
 *  
 * OBS: Este programa foi desenvolvido a partir do programa apresentado na dissertação de  
 *       mestrado de Oliveira(2004).  
 *  
 * DATA: 10/11/2008.  
 *  
 * OBS: O arquivo stat.txt (definido na linha 47) será criado no mesmo diretório onde está este  
 *       programa para simulação. Neste arquivo serão gravadas as estatísticas de cada réplica  
 *       para serem usadas no R para gerar gráficos.  
 * *****  
 #include <oxstd.h>  
 #include <oxprob.h>  
 #import <maximize>  
 #include <oxdraw.h>  
  
 /* Variáveis globais */  
 static decl s_vY; /* Amostra aleatória (vetor coluna) */  
 static decl s_mX; /* Matriz X */  
 static decl s_mXr; /* Matriz X restrita */  
 static decl s_mZ; /* Matriz Z */  
 static decl s_mZr; /* Matriz Z restrita */
```

```

/* Definindo Constantes */
const decl nobs = 20;                                     /* Tamanho da amostra */
const decl iRep = 50000;                                 /* Número de réplicas MC */

const decl k = 4;                                       /* Número de variáveis independentes no modelo para mu*/
const decl m = 4;                                       /* Número de variáveis independentes no modelo para phi*/
const decl r = 3;                                       /* Número de parâmetros que estão sendo testados para phi*/
const decl beta0 = <1,1,5,-4>;                          /* valor de beta em H0 */
// const decl gama0 = <1.609438,0,0,0>;                 /* valor de gama em H0 (incluir m-1 zeros)*/
const decl gama0 = <2.302585,0,0,0>;                     /* valor de gama em H0 (incluir m-1 zeros)*/
// const decl gama0 = <3.401197,0,0,0>;                 /* valor de gama em H0 (incluir m-1 zeros)*/
const decl epsilon = 0.5;                                /* valor para gama em H1 */
// const decl gamal = <1.609438,epsilon,epsilon,epsilon>;
const decl gamal = <2.302585,epsilon,epsilon,epsilon>;
// const decl gamal = <3.401197,epsilon,epsilon,epsilon>;
                                                                    /* valor de gama em H1 (incluir m-1 epsilon) */
const decl TESTE = 1;                                   /*Teste=1:sob H0:(beta0,gama0) , para obter o tamanho do teste
                                                         Teste=2:sob H1:(beta0,gamal) , para obter o poder do teste */

decl stat="stat.txt";
/* Arquivo onde serão guardadas as estatísticas para a produção do gráfico no R */

/* Função de log-verossimilhança */
floglik(const vtheta, const adFunc, const avScore, const amHess)
{
  decl eta = s_mX*vtheta[0:(k-1)];
  decl mu = exp(eta) ./ (1.0 + exp(eta));
  decl delta = s_mZ*vtheta[k:(k+m-1)];
  decl phi = exp(delta);
  decl ystar = log( s_vY ./ (1.0 - s_vY) );
  decl mustar = polygamma(mu.*phi,0)-polygamma((1.0-mu).*phi,0);
  decl ydag = log( 1.0 - s_vY );
  decl mudag = polygamma((1.0 - mu) .* phi, 0) - polygamma(phi, 0) ;
  decl T = diag( exp(eta) ./ (1.0 + exp(eta)) .^2 );
                                                                    /*função de ligação de mu é a logito: g(mu_t)=log(mu_t/(1-mu_t))*/
  decl F = diag( phi );
  decl H = diag( phi );
                                                                    /*função de ligação de phi é h(phi_t)=log(phi_t) então 1/h'(phi_t) = phi_t*/
  decl M = diag(mu);

  adFunc[0] = double(sumc( loggamma(phi) - loggamma(mu .* phi) -
    loggamma((1.0 - mu) .* phi) +
    (mu .* phi - 1.0) .* log(s_vY) +
    ( (1.0 - mu) .* phi - 1.0 ) .* log(1.0 - s_vY) ));
                                                                    /*log-verossimilhança*/

  if(avScore)
  {
    (avScore[0])[0:(k-1)] = s_mX'*T*F*(ystar-mustar);
    (avScore[0])[k:(k+m-1)] = s_mZ'*H*(M*(ystar-mustar)+(ydag-mudag));
  }

  if( isnan(adFunc[0]) || isdotinf(adFunc[0]) )
    return 0;
  else
    return 1; /* 1 indica sucesso */
}

/* Função de log-verossimilhança restrita */
floglikr(const vtheta, const adFunc, const avScore, const amHess)
{
  decl eta = s_mX*vtheta[0:(k-1)];
  decl mu = exp(eta) ./ (1.0 + exp(eta));
  decl delta = s_mZr*vtheta[k:(k+m-r-1)];
  decl phi = exp(delta);
  decl ystar = log( s_vY ./ (1.0 - s_vY) );
  decl mustar = polygamma(mu.*phi,0)-polygamma((1.0-mu).*phi,0);
  decl ydag = log( 1.0 - s_vY );
  decl mudag = polygamma((1.0 - mu) .* phi, 0) - polygamma(phi, 0);

```

```

decl T = diag( mu .* (1 - mu) );          /*função de ligação de mu é a logito: g(mu)=log(mu/(1-mu))
                                          então 1/g'(mu) = mu*(1-mu)*/
decl F = diag( phi );
decl H = diag( phi );          /*função de ligação de phi é h(phi)=log(phi) então 1/h'(phi) = phi*/
decl M = diag(mu);

adFunc[0] = double(sumc(loggamma(phi)-loggamma(mu.*phi)-
                      loggamma((1.0-mu).* phi)+
                      (mu.*phi-1.0).*log(s_vY)+
                      ((1.0- mu).*phi-1.0).*log(1.0-s_vY)));          /*log-verossimilhança*/
if(avScore)
{
    (avScore[0])[0:(k-1)] = s_mX'*T*F*(ystar-mustar);
(avScore[0])[k:(k+m-r-1)] = s_mZr'*H*(M*(ystar-mustar)+(ydag-mudag));
}

if( isnan(adFunc[0]) || isdotinf(adFunc[0]) )
return 0;
else
    return 1; /* 1 indica sucesso */
}

/*Função para obtenção da parte da matriz correspondente ao parâmetro de incômodo para testar
(gama_(m-r+1),...,gama_m), sendo theta=(beta_1,...,beta_k,gama_1,...,gama_m)*/
nuisance(const matriz)
{
    decl ma = matriz;
    ma = ma[:k+m-1-r][:k+m-1-r];
    return ma;
}

/* Início do Programa */
main()
{
    /* Declaração das Variáveis Utilizadas no Programa */

    /* Parâmetros de Sistema */
    decl dExecTime, fail, falha, falhar, controle, ci, i, contador, auxiliar;

    /* Parâmetros do Modelo */
    decl beta, gama, theta, eta, delta, mu, phi ;

    /* Parâmetros Auxiliares sem restrição */
    decl ystar, ydag, betaols, etaols, muols, varols, phiols, gamaols, vtheta0, converge, dfunc1;

    /* Parâmetros Auxiliares com restrição */
    decl nobsr, betaolsr, etaolsr, muolsr, varolsr, phiolsr, gamaolsr, vtheta1, converger, dfunc2;

    /* Parâmetros de Resultados */
    decl vEMV, resul1, resul2, resul3, resul4, cobertura, liminferior, limsuperior, vsig, mStat,
                                                vquantil, cStat;

    /* Inicia o Relógio */
    dExecTime = timer();

    /* Obtendo a matriz X de dimensão nobs x k*/
    s_mX = 1~(ranu(nobs,k-1)-0.5);

    s_mZ = s_mX[:,(m-1)];
    s_mZr = s_mZ[0:(m-r-1)]; /* Matriz Z restrita */

    /* Valores verdadeiros dos parâmetros */
    if(TESTE == 1) /*sob H0*/
    {
        gama = gama0';
    }
    else /*sob H1*/
    {
        gama = gama1';
    }
}

```

```

}
beta = beta0';
theta = beta | gama;
eta = s_mX * beta;
mu = exp(eta) ./ (1.0 + exp(eta)); /* Vetor Coluna nobis x 1 */
delta = s_mZ * gama;
phi = exp(delta);

/* Níveis de significância do teste */
vsig = <0.10; 0.05; 0.01; 0.005;0.001>;

/* Quantis da distribuição Qui-Quadrado (r graus de liberdade) */
vquantil = (quanchi(1-vsig, r))';

/* Inicialização de Algumas Quantidades */

vEMV = zeros(iRep, k+m); /* Inic. do Vetor dos EMV */
cobertura = zeros(1,k+m); /* Inic. do Contador da Taxa de Cobertura */
liminferior = zeros(iRep,k+m); /* Inic. do Contador do L. Inferior */
limsuperior = zeros(iRep,k+m); /* Inic. do Contador do L. Superior */
contador = zeros(1,k+m); /* Inic. do Contador Auxiliar para a Taxa de Cobertura */
falha = 0; /* Inic. do Contador de Amostras de MC que Falharam */
falhar = 0; /* Inic. do Contador de Amostras de MC que Falharam */
fail = 0; /* Inic. do Contador de Amostras de MC que Falharam */
mStat = zeros(iRep,5); /* Inic. do Vetor de Valores das Estatísticas do Teste */
auxiliar = zeros(5, columns(vsig)); /* Inic. do Cont. Auxiliar para os Tamanhos dos Testes */
cStat = zeros(5, columns(vsig)); /* Inic. do Contador dos Tamanhos dos Testes */
decl ocorreu = 0;

/* Escolha do gerador de números aleatórios */
ranseed("GM");

/* Escolha da Semente */
ranseed({1965, 2001});

/* Parâmetro de controle do processo */
controle = (iRep/100);

/* Loop de Monte Carlo */
for(ci = 0; ci < iRep + fail; ci++)
{
    s_vY = zeros(nobs, 1); /* Inicialização do Vetor s_vY */

    for(i = 0; i < nobs; i++)
        /* beta (mu,phi) */
        s_vY[i] = ranbeta(1, 1, mu[i]*phi[i], (1-mu[i])*phi[i]);
    if(s_vY[i]==1.0000)s_vY[i]=0.999999999;
    if(s_vY[i]==0.0000)s_vY[i]=0.0000000001;
}

/* Obtendo o Valor de y* = ystar */
ystar = log( s_vY ./ (1.0-s_vY) );

/* Obtendo o Valor de ydag */
ydag = log( 1.0-s_vY );

/***** Calculo sem restrição *****/
/*Calculando as Estimativas de Momentos sugeridas por Ferrari e Cribari-Neto (2004)*/

if(k > 1){ols2c(ystar, s_mX, &betaols);}
/* Retorna em betaols a Estimativa de Mínimos Quadrados Ordinários : (X'X)^-1 * X'*ystar */
else if(k==1){betaols = meanc(ystar);}

/*Obtendo Algumas Quantidades para o Cálculo do Valor Inicial de gama*/
etaols = s_mX*betaols; /* X(X'X)^-1*X'*ystar */
muols = exp(etaols) ./ (1.0 + exp(etaols)); /* g^-1(X(X'X)^-1*X'*ystar) */
varols = ((ystar-etaols)'*(ystar-etaols))./ ((nobs-k).*((1 ./ (muols.*(1-muols))) .^ (2))));
/* ((ystar-yhat)'*(ystar-yhat))/((nobs-k)*(g'(mu)^(2)))*/*

```

```

// phiols = ( muols .* (1.0-muols) ./ varols ) - 1.0;          /* phi^(0) */
phiols = ( muols .* (1.0-muols) ./ varols );                 /* phi^(0) */
if(m > 1){ols2c(log(phiols), s_mZ, &gamaols);}
/* Retorna em gamaolsr a Estimativa de Mínimos Quadrados Ordinários : (Z'Z)^-1 * Z'*log(phi) */
else if(m==1){gamaols = meanc(phiols);}

/* Valores Iniciais para os Parâmetro (beta, gama) */
vtheta0 = betaols | gamaols;

MaxControl( 50, -1 );

/* Estimação de máxima verossimilhança */
converge = MaxBFGS(floglik, &vtheta0, &dfunc1, 0, 0);

/* Verificando a convergência */
if(converge != MAX_CONV && converge != MAX_WEAK_CONV)
{
    falha++;
    fail++;
    continue; /* sai do loop MC */
}

decl    etahat = s_mX * vtheta0[0:(k-1)];
decl    muhat = exp(etahat) ./ (1+exp(etahat));
decl    T = diag(muhat .* (1-muhat));
decl    deltahat = s_mZ * vtheta0[k:(k+m-1)];
decl    phihat = exp(deltahat);
decl    H = diag(phihat);
decl    s_mYstar = diag(ystar);
decl    s_mYdag = diag(ydag);
decl    M = diag(muhat);
decl    F = diag(phihat);
decl    Mstar = diag(polygamma(muhat .* phihat, 0) - polygamma((1.0 - muhat).* phihat, 0));
                                                    /*média de ystar=log(y/(1-y))*/
decl    Vstar = diag(polygamma(muhat .* phihat, 1) + polygamma((1.0 - muhat) .* phihat, 1));
                                                    /*variância de ystar*/
decl    Mdag = diag(polygamma((1.0 - muhat) .* phihat, 0) - polygamma(phihat, 0));
                                                    /*média de ydag=log(1-y)*/
decl    Vdag = diag(polygamma((1.0 - muhat) .* phihat, 1) - polygamma(phihat, 1));
                                                    /*variância de ydag*/
decl    C = diag(-polygamma((1 - muhat) .* phihat, 1)); /*covariância entre ystar e ydag*/
decl    S = diag((2*muhat-1) ./ ((muhat .* (1-muhat)) .^ 2)); /*g''(mu)*/
decl    Q = diag(-1 ./ (phihat .^ 2)); /*h''(phi)*/
decl    Ibetabeta = s_mX' * F * T * Vstar * T * F * s_mX;
decl    Ibetagama = s_mX' * F * (M * Vstar + C) * T * H * s_mZ;
decl    Igamabeta = Ibetagama';
decl    Igamagama = s_mZ' * H * (M * Vstar * M + (M + M) * C + Vdag) * H * s_mZ;
decl    fisher = (Ibetabeta ~ Ibetagama) | (Igamabeta ~ Igamagama); /*informação de Fisher*/
decl    fisherinv = invert(fisher); /*inversa da informação de Fisher*/
decl    Jbetabeta = s_mX' * (F * T * Vstar + S * T^2 * (s_mYstar - Mstar)) * T * F * s_mX;
decl    Jbetagama = -s_mX' * ((s_mYstar - Mstar) - F * (M * Vstar + C)) * T * H * s_mZ;
decl    Jgamabeta = Jbetagama';
decl    Jgamagama = s_mZ' * (H * (M * Vstar * M + (M + M) * C + Vdag) + (M * (s_mYstar - Mstar) + (s_mYdag - Mdag)) * Q * H^2)
                                                    * H * s_mZ;
decl    infobs = (Jbetabeta ~ Jbetagama) | (Jgamabeta ~ Jgamagama);
decl    infobsinv = invert(infobs);
decl    stderrors = sqrt(diagonal(fisherinv));

/***** Calculo com restrição *****/
/*Calculando as Estimativas de Momentos sugeridas por Ferrari e Cribari-Neto (2004)*/

/*Obtendo Algumas Quantidades para o Cálculo do Valor Inicial de gama*/
// phiolsr = double(meanc((muols.*(1.0-muols))./varols)-1.0); /* phi^(0) */
phiolsr = double(meanc((muols.*(1.0-muols))./varols)); /* phi^(0) */
gamaolsr = log(phiolsr) | zeros((m-1),1);

/* Valores Iniciais para os Parâmetros (beta, gama) */
vtheta1 = betaols | gamaolsr;

```

```

MaxControl( 50, -1 );

/* Estimação de máxima verossimilhança restrita */
converger = MaxBFGS(floglikr, &vtheta1, &dfunc2, 0, 0);

if (dfunc1-dfunc2 < 0){ocorreu++;}

/* Verificando a convergência */
if(converger != MAX_CONV && converger != MAX_WEAK_CONV)
{
  falhar++;
  fail++;
  continue; /* sai do loop MC */
}

decl      etatil = s_mX*(vtheta1[0:(k-1)]);
decl      mutil = exp(etatil) ./ (1+exp(etatil));
decl      Ttil = diag( mutil .* (1-mutil) );
decl      deltatil = s_mZ*(vtheta1[k:(k+m-1-r)] | theta[(k+m-r):(k+m-1)]);
decl      phitil = exp(deltatil);
decl      Htil = diag(phitil);
decl      Mtil = diag(mutil);
decl      Ftil = diag(phitil);
decl      mustartil = polygamma(mutil .* phitil, 0) - polygamma((1.0 - mutil) .* phitil, 0);
                                                    /*média de ystar=log(y/(1-y)*/
decl      Mstartil = diag(mustartil);
decl      mudagtil = polygamma((1.0 - mutil) .* phitil, 0) - polygamma(phitil, 0);
                                                    /*média de ydag=log(1-y)*/
decl      Mdagtil = diag(mudagtil);
decl      Vstartil = diag(polygamma(mutil .* phitil, 1) + polygamma((1.0 - mutil) .* phitil, 1));
                                                    /*variância de ystar*/
decl      Vdagtil = diag(polygamma((1.0 - mutil) .* phitil, 1)- polygamma(phitil, 1));
                                                    /*variância de ydag*/
decl      Ctil = diag(-polygamma((1 - mutil) .* phitil, 1));/*covariância entre ystar e ydag*/
decl      Stil = diag((2*mutil-1) ./ ((mutil .* (1-mutil)).^ 2));          /*g' (mu)*/
decl      Qtil = diag(-1 ./ (phitil .^ 2));          /*h' (phi)*/
decl      Ibetabetatil = s_mX'*Ftil*Ttil*Vstartil*Ttil*Ftil*s_mX;
decl      Ibetagamatil = s_mX'*Ftil*(Mtil*Vstartil+Ctil)*Ttil*Htil*s_mZ;
decl      Igamabetatil = Ibetagamatil';
decl      Igamagamatil = s_mZ'*Htil*(Mtil*Vstartil*Mtil+(Mtil+Mtil)*Ctil+Vdagtil)*Htil*s_mZ;
decl      fishertil = (Ibetabetatil ~ Ibetagamatil) | (Igamabetatil ~ Igamagamatil);
                                                    /*informação de Fisher*/
decl      fisherinvtil = invert(fishertil);          /*inversa da informação de Fisher*/
decl      Jbetabetatil = s_mX'*(Ftil*Ttil*Vstartil + Stil*Ttil^2*(s_mYstar-Mstartil))*Ttil*Ftil*s_mX;
decl      Jbetagamatil = -s_mX'*((s_mYstar-Mstartil)- Ftil*(Mtil*Vstartil+Ctil))*Ttil*Htil*s_mZ;
decl      Jgamabetatil = Jbetagamatil';
decl      Jgamagamatil = s_mZ'*
      (Htil*(Mtil*Vstartil*Mtil+(Mtil+Mtil)*Ctil+Vdagtil)+
      (Mtil*(s_mYstar-Mstartil)+(s_mYdag-Mdagtil))*Qtil*Htil^2)*
      Htil*s_mZ;
decl      infobstil = (Jbetabetatil~Jbetagamatil)| (Jgamabetatil~Jgamagamatil);
                                                    /*informação observada*/
decl      infobsinvtil = invert(infobstil);          /*inversa da informação observada*/
decl      escorebetatil = s_mX'*Ttil*Ftil*(ystar-mustartil);
decl      escoregamatil = s_mZ'*Htil*(Mtil*(ystar-mustartil)+(ydag-mudagtil));
decl      escoretil = escorebetatil | escoregamatil;          /* função escore */

/*Obtendo as quantidades restantes para o cálculo do ajuste de Skovgaard*/

decl      iota = ones(nobs,1);
decl      q1 = s_mX'*T*F*(Vstar*(M*F-Mtil*Ftil)+ C*(F-Ftil))*iota;
decl      q2 = s_mZ'*H*((M*F-Mtil*Ftil)*(M*Vstar+C)+(F-Ftil)*(M*C+Vdag))*iota;
decl      q = q1|q2;
decl      epsilon1 = s_mX'*F*T*Vstar*Ttil*Ftil*s_mX;
decl      epsilon2 = s_mX'*T*F*(Mtil*Vstar+C)*Htil*s_mZ;
decl      epsilon3 = s_mZ'*H*Ftil*(M*Vstar+C)*Ttil*s_mX;
decl      epsilon4 = s_mZ'*H*(M*Vstar*Mtil+(M+Mtil)*C+Vdag)*Htil*s_mZ;

```

```

decl      epsilon = (epsilon1-epsilon2)|(epsilon3-epsilon4);
decl      epsiloninv = invert(epsilon);
decl      nuininfobstl = nuisance(infobstl);
decl      nuisance2 = nuisance(fishertil*epsiloninv*infobs*fisherinv*epsilon);

/* Vetores das EMVs */
vEMV[ci-fail][] = vtheta0';

/*Cálculo dos Limites Inferiores e Superiores do Intervalo de Confiança */
liminferior[ci-fail][] = vtheta0' - 1.96*stderrors;
limsuperior[ci-fail][] = vtheta0' + 1.96*stderrors;

/* Taxa de Cobertura */
contador = (liminferior[ci-fail][] .<= theta') && (theta' .<= limsuperior[ci-fail][]);
cobertura = cobertura + contador;

/* Estatísticas do Teste: RV, Escore, Wald e Skovgaard*/
mStat[ci-fail][0] = 2*(dfuncl-dfunc2); /* Razão de Verossimilhança */

decl nu = fabs(fabs(determinant(fishertil))^(1/2)* fabs(determinant(fisher))^(1/2)*
  fabs(determinant(epsilon))^-1)* fabs(determinant(nuininfobstl))^(1/2)*
  fabs(determinant(nuisance2))^-1/2)*
  fabs(escoretel'*epsiloninv*fisher*infobsinv*epsilon* fisherinvtil*escoretel)^(r/2)/
  fabs((mStat[ci-fail][0])^(r/2)-1)*escoretel'* epsiloninv*q);

mStat[ci-fail][1] = mStat[ci-fail][0]*(1-(log(nu)/mStat[ci-fail][0]))^2; /* Skovgaard 1 */
mStat[ci-fail][2] = mStat[ci-fail][0]-2*log(nu); /* Skovgaard 2 */

if(mStat[ci-fail][0]<0.1) /*forte evidência para não rejeitar a hipótese nula*/
{
  mStat[ci-fail][3] = mStat[ci-fail][0]; /* RV** = Skovgaard 1 com ajuste por RV*/
  mStat[ci-fail][4] = mStat[ci-fail][0]; /* RV* = Skovgaard 2 com ajuste por RV*/
}
else
{
  mStat[ci-fail][3] = mStat[ci-fail][0]*
    (1-(log(nu)/mStat[ci-fail][0]))^2; /* RV** = Skovgaard 1 com ajuste por RV*/
  mStat[ci-fail][4] = mStat[ci-fail][0]-2*log(nu); /* RV* = Skovgaard 2 com ajuste por RV*/
}

/* Nível de significância: 100 * vsig[cq][0] % */
auxiliar = ((mStat[ci-fail][])' .> vquantil);
cStat = cStat + auxiliar;

/* Controle do processo */
if(imod(ci-fail + 1, controle) == 0)
print("\n", ci-fail + 1, " amostras realizadas : ", (ci-fail + 1)*100/iRep, " % das replicas");
} /* Final do Loop de Monte Carlo */

/* ORGANIZAÇÃO DOS RESULTADOS - Estimção de Máxima Verossimilhança */

result1 = zeros(k+m, 9); /* Inicialização da Matriz de Resultados */
decl resultlr = zeros(k+m, 9); /* Inicialização da Matriz de Resultados sob restricao*/

/* Média das Estimativas */
result1[][0] = (meanc(vEMV))';

/* Desvio Padrão das Estimativas */
result1[][1] = sqrt((sumsqrc(vEMV) ./ iRep)' - sqr(result1[][0]));

/* Viés da Média das Estimativas */
result1[][2] = (result1[][0] - theta);

/* Raiz do Erro Quadrático Médio Estimado */

```

```

    resul1[][3] = sqrt((resul1[][1]) .^ 2 + sqr(resul1[][2]));

/* Coeficiente de assimetria */
    resul1[][4] = (sumc((standardize(vEMV)) .^ 3) ./ iRep)';

/* Coeficiente de curtose */
    resul1[][5] = (sumc((standardize(vEMV)) .^ 4) ./ iRep - 3)';

/* Taxa de Cobertura */
    resul1[][6] = ((cobertura ./iRep) .*100)';

/* Limite Inferior do Intervalo de Confiança */
    resul1[][7] = (meanc(liminferior))';

/* Limite Superior */
    resul1[][8] = (meanc(limsuperior))';

/* ORGANIZAÇÃO DOS RESULTADOS - Tamanhos dos Testes */
    resul2 = zeros(columns(vsig), 5);
    resul2 = ((cStat' ./iRep) .*100);

/* Momentos das Estatísticas de Teste */
    resul3 = zeros(4, 5); /* Inicialização da Matriz de Resultados */
    resul3[0][] = meanc(mStat); /* Média */
    resul3[1][] = varc(mStat); /* Variância */
    resul3[2][] = minc(mStat); /* Mínimo */
    resul3[3][] = maxc(mStat); /* Máximo */

/* Quantis Amostrais das Estatísticas de Teste */
    resul4 = zeros(columns(vsig), 6); /* Inicialização da Matriz de Resultados */
    resul4 = quanchi(1-vsig, r) ~ quantilec(mStat, 1 - vsig)';

/* Salvando em arquivo RV, RV*=Skov2.1 e RV**=Skov1.1 para produção de gráficos no R*/
    decl Statgraf = mStat[][0] ~ mStat[][3:4];
    savemat(stat , Statgraf);

/* Impressão dos Resultados: */
    print( "\n PROGRAMA OX: ", oxfilename(0) );
    print( "\n VERSAO OX: ", oxversion() );
    print( "\n DISTRIBUICAO: beta " );
    print( "\n TAMANHO DAS AMOSTRAS: ", nobs );
    print( "\n NUMERO DE REPLICAS MC: ", ci );
    print( "\n NUMERO DE COVARIADAS: ", k+m);
    print( "\n\t NUM. REPLICAS MC COM CONVERGENCIA: ", iRep );
    print( "\n\t NUM. REPLICAS MC SEM CONVERGENCIA: ", fail );
    print( "\n\t\t Sem restricao: ", falha );
    print( "\n\t\t Com restricao: ", falhar);
    print( "\n\t\t Ocorreu: ", ocorreu);

    print("\n\n#####");
    print("\nSkov1 = RV*(1- (1/RV * log(nu))^2);");
    print("\nSkov2 = RV - 2*log(nu)");
    print("\nRV** = RV se RV < 0.1 e RV** = Skov1 c.c");
    print("\nRV* = RV se RV < 0.1 e RV* = Skov2 c.c");
    print("\n#####");

    print( "\n\nValores Verdadeiros dos Parametros: ");
    print( "\n", "%8.3f",
        "%c", {"beta_1", "beta_2","beta_3","beta_4", "beta_5",
            "beta_6", "beta_7", "beta_8", "beta_9", "beta_10"},
        "%r", {"-"}, (theta[(k-1)])');
    print("%8.3f",
        "%c", {"gama_1","gama_2","gama_3","gama_4","gama_5",
            "gama_6","gama_7","gama_8","gama_9","gama_10"},
        "%r", {"-"}, (theta[k:(k+m-1)])');

    print("\n\nNumero de parametros testados = r = ",r);
    print( "\n\nHipotese Nula: Ho: (gama_,m-r+1,...,gama_m)=(,gama0[(m-r):]);");

```



```

/* Resultados da Estimação dos Parâmetros */

print( "\n\nResultados da Estimacao dos Parametros (beta, gama)");
print( "\n", "%8.2f",
      "%c", {"Media", "DP", "Vies", "REQM", "Assim", "Curt.",
            "TCob", "LI", "LS"},
      "%r", {"beta_1", "beta_2", "beta_3", "beta_4", "beta_5",
            "beta_6", "beta_7", "beta_8", "beta_9", "beta_10"},
            resull[0:(k-1)][]);
print( "%8.2f",
      "%r", {"gama_1", "gama_2", "gama_3", "gama_4", "gama_5",
            "gama_6", "gama_7", "gama_8", "gama_9", "gama_10"},
            resull[k:(k+m-1)][]);

if(TESTE == 1)
{
/* Resultados do Tamanho dos Testes */
print( "\n\nResultados do Tamanho dos Testes");
}
else
{
/* Resultados do Poder dos Testes */
print( "\n\nHipotese alternativa: H1: gama = ", gamal);
print( "\n\nResultado do Poder dos Testes");
}
print( "\n", "%10.4f",
      "%c", {"RV", "Skov1", "Skov2", "RV**", "RV*"},
      "%r", {"10%", "5%", "1%", "0,5%", "0,1%"}, resul2);

/* Resultados das medidas resumo das estatísticas de Teste */
print( "\n\nResultados das medidas resumo das Estatisticas de Teste");
print( "\n", "%10.2f",
      "%c", {"RV", "Skov1", "Skov2", "RV**", "RV*"},
      "%r", {"Media", "Variancia", "Minimo", "Maximo"}, resul3);

/* Resultados dos Quantis amostrais das Estatísticas de Teste */
print( "\n\nResultados dos Quantis amostrais das Estatisticas de Teste");
print( "\n\nGraus de liberdade: ", r);
print( "\n", "%9.2f",
      "%c", {"Chi_g1", "RV", "Skov1", "Skov2", "RV**", "RV*"},
      "%r", {"90%", "95%", "99%", "99,5%", "99,9%"}, resul4);

/* Data, Hora e Tempo de Execução */
print( "\n\nDATA: ", date() );
print( "\n\nHORA: ", time(), "\n" );
print( "\n\n\t\t\t TEMPO TOTAL DE EXECUCAO: ", timespan(dExecTime) );
print( "\n" );
}
/*----- Fim do Programa -----*/

```

```

/*****
* PROGRAMA: hetero_browlee_mu.ox
*
* USO: Estimativas dos parâmetros e cálculo de estatísticas para o conjunto de dados Brownlee e
*       Modelo:  g(mu)=log(mu/(1-mu))=s_mX*beta,          beta=(beta_1,...,beta_k)
*               h(phi)=log(phi)=s_mZ*gama,              gama=(gama_1,...,gama_m)
*       Hipótese: beta_(k-r+1)=...=beta_k=0
*
* AUTOR: Eliane Cantinho Pinheiro.
*
* OBS: Este programa foi desenvolvido a partir do programa apresentado na dissertação de
*       mestrado de Oliveira(2004).
*
* DATA: 10/11/2008.
*****/

/* Variáveis globais */
static decl s_vY;      /* Amostra aleatória (vetor coluna) */
static decl s_mX;      /* Matriz X - regressores para mu*/
static decl s_mXr;     /* Matriz X restrita */
static decl s_mZ;      /* Matriz Z - regressores para phi*/
static decl s_mZr;     /* Matriz Z restrita */
static decl nobs;     /* Tamanho da amostra */
static decl k;        /* Número de variáveis independentes no modelo para mu*/
static decl m;        /* Número de variáveis independentes no modelo para phi*/

/* Definindo Constantes */
const decl nome_arquivo = "brownlee.dat";

const decl r = 1;          /* Número de parâmetros que estão sendo testados para phi*/
const decl beta0 = <0>;   /*beta da hipótese nula tem que ter dimensão r)*/

// const decl r = 2;      /* Número de parâmetros que estão sendo testados para phi*/
// const decl beta0 = <0,0>; /*beta da hipótese nula tem que ter dimensão r)*/

/* Função de log-verossimilhança */
floglik(const vtheta, const adFunc, const avScore, const amHess)
{
  decl eta = s_mX*vtheta[0:(k-1)];
  decl mu = exp(eta) ./ (1.0 + exp(eta));
  decl delta = s_mZ*vtheta[k:(k+m-1)];
  decl phi = exp(delta);
  decl ystar = log( s_vY ./ (1.0 - s_vY) );
  decl mustar = polygamma(mu.*phi,0)-polygamma((1.0-mu).*phi,0);
  decl ydag = log( 1.0 - s_vY );
  decl mudag = polygamma((1.0 - mu) .* phi, 0) - polygamma(phi, 0) ;
  decl T = diag( exp(eta) ./ (1.0 + exp(eta)) .^2 );
  /*função de ligação de mu é a logito: g(mu_t)=log(mu_t/(1-mu_t))*/
  decl F = diag( phi );
  decl H = diag( phi );
  /*função de ligação de phi é h(phi_t)=log(phi_t) então 1/h'(phi_t) = phi_t*/
  decl M = diag(mu);

  adFunc[0] = double(sumc( loggamma(phi) - loggamma(mu .* phi) -
    loggamma((1.0 - mu) .* phi) +
    (mu .* phi - 1.0) .* log(s_vY) +
    ( (1.0 - mu) .* phi - 1.0 ) .* log(1.0 - s_vY) )); /*log-verossimilhança*/

  if(avScore)
  {
    (avScore[0])[0:(k-1)] = s_mX'*T*F*(ystar-mustar);
    (avScore[0])[k:(k+m-1)] = s_mZ'*H*(M*(ystar-mustar)+(ydag-mudag));
  }

  if( isnan(adFunc[0]) || isdotinf(adFunc[0]) )
    return 0;
  else
    return 1; /* 1 indica sucesso */
}

```

```

}

/* Função de log-verossimilhança restrita */
floglikr(const vtheta, const adFunc, const avScore, const amHess)
{
  decl eta = s_mX*(vtheta[0:(k-1-r)]|beta0');
  decl delta = s_mZ*(vtheta[(k-r):(k+m-1-r)]);
  decl mu = exp(eta) ./ (1.0 + exp(eta));
  decl phi = exp(delta);
  decl ystar = log( s_vY ./ (1.0 - s_vY) );
  decl mustar = polygamma(mu.*phi,0)-polygamma((1.0-mu).*phi,0);
  decl ydag = log( 1.0 - s_vY );
  decl mudag = polygamma((1.0 - mu) .* phi, 0) - polygamma(phi, 0);
  decl T = diag( mu .* (1 - mu) ); /*função de ligação de mu é a logito: g(mu)=log(mu/(1-mu))
                                     então 1/g'(mu) = mu*(1-mu)*/

  decl F = diag( phi );
  decl H = diag( phi ); /*função de ligação de phi é h(phi)=log(phi) então 1/h'(phi) = phi*/
  decl M = diag(mu);

  adFunc[0] = double(sumc(loggamma(phi)-loggamma(mu.*phi)-
                          loggamma((1.0-mu).* phi)+
                          (mu.*phi-1.0).*log(s_vY)+
                          ((1.0- mu).*phi-1.0).*log(1.0-s_vY))); /*log-verossimilhança*/

  if(avScore)
  {
    (avScore[0])[0:(k-1-r)] = s_mXr'*T*F*(ystar-mustar);
    (avScore[0])[(k-r):(k+m-1-r)] = s_mZ'*H*(M*(ystar-mustar)+(ydag-mudag));
  }

  if( isnan(adFunc[0]) || isdotinf(adFunc[0]) )
  return 0;
  else
    return 1; /* 1 indica sucesso */
}

/*Função para obtenção da parte da matriz correspondente ao parâmetro de incômodo para testar
beta_(k-r+1)=...=beta_k, sendo theta=(beta_1,...,beta_k,gama_1,...,gama_m)*/
nuisance(const matriz)
{
  decl ma = matriz;
  ma = ma[:,k-1-r]~ma[:,k:k-1+m];
  ma = ma[:,k-1-r]||ma[k:k-1+m];
  return ma;
}

/* Início do Programa */
main()
{
  /* Declaração das Variáveis Utilizadas no Programa */

  /* Base de dados */
  decl data;

  /* Parâmetros de Sistema */
  decl dExecTime, fail, falha, falhar, controle, ci, i, contador, auxiliar;

  /* Parâmetros do Modelo */
  decl beta, theta, eta, delta, mu, phi ;

  /* Parâmetros Auxiliares sem restrição */
  decl ystar, ydag, betaols, etaols, muols, varols, phiols, gamaols, vtheta0, converge, dfunc1;

  /* Parâmetros Auxiliares com restrição */
  decl nobsr, betaolsr, etaolsr, muolsr, varolsr, phiolsr,gamaolsr, vtheta1, converger, dfunc2;

  /* Resultados */
  decl RV, Skov1,Skov2;

```

```

/* Inicia o Relógio */
dExecTime = timer();

/* Lendo os dados */
data = loadmat(nome_arquivo);
// s_mX = 1~data[][1];
s_mX = 1~data[][1]~data[][2];
// s_mX = 1~data[][2]~data[][1];
// s_mX = 1~data[][1]~data[][2]~data[][3];
// s_mX = 1~data[][1]~data[][2]~data[][3]~data[][4];
// s_mX = 1~data[][1]~data[][2]~data[][3]~data[][4]~data[][5];
// s_mX = 1~data[][1]~data[][2]~(data[][1].*data[][2])~((data[][1]).^2);
// s_mX = 1~data[][1]~data[][2]~(data[][1].*data[][2]);
print("matriz X      :",s_mX);

// s_mZ = s_mX;
s_mZ = 1~data[][1];
print("matriz Z      :",s_mZ);

s_vY = data[][0];
nobs = rows(data);
k = columns(s_mX);
m = columns(s_mZ);

if(max(s_vY) >= 1.0 || min(s_vY) <= 0.0)
{
    println("\n\nErro: Dados fora do limite (0,1)!\n\n");
    exit(2);
}

if(k >= nobs)
{
    println("\n\nErro: Numero de covariadas nao pode exceder o numero de observacoes!\n\n");
    exit(3);
}

s_mXr = s_mX[][0:(k-1-r)]; /* Matriz Z restrita */
print("matriz Xr     :",s_mXr);

/* Obtendo o Valor de y* = ystar */
ystar = log( s_vY ./ (1.0-s_vY) );

/* Obtendo o Valor de ydag */
ydag = log( 1.0-s_vY );

/*****          Cálculo sem restrição          *****/
/*Calculando as Estimativas de Momentos sugeridas por Ferrari e Cribari-Neto (2004)*/

if(k > 1){ols2c(ystar, s_mX, &betaols);}
/* Retorna em betaols a Estimativa de Mínimos Quadrados Ordinários,
ou seja,  $(X'X)^{-1} * X' * ystar$  */
else if(k==1){betaols = meanc(ystar);}

/*Obtendo Algumas Quantidades para o Cálculo do Valor Inicial de gama*/
etaols = s_mX*betaols; /*  $X(X'X)^{-1} * X' * ystar$  */
muols = exp(etaols) ./ (1.0 + exp(etaols)); /*  $g^{-1}(X(X'X)^{-1} * X' * ystar)$  */
varols = ((ystar-etaols)'*(ystar-etaols))./
((nobs-k).*(1 ./ (muols.*(1-muols))) .^ (2));
/*  $((ystar-yhat)'*(ystar-yhat))/((nobs-k)*(g'(\mu)^{(2)}))$  */
phiols = ( muols .* (1.0-muols) ./ varols ) - 1.0; /*  $\phi^{(0)}$  */

if(m > 1){ols2c(log(phiols), s_mZ, &gamaols);}
/* Retorna em gamaolsr a Estimativa de Mínimos Quadrados Ordinários,
ou seja,  $(Z'Z)^{-1} * Z' * \log(\phi)$  */
else if(m==1){gamaols = meanc(phiols);}

/* Valores Iniciais para os Parâmetro (beta, gama) */
vtheta0 = betaols | gamaols;

```

```

MaxControl( 50, -1 );

/* Estimação de máxima verossimilhança */
converge = MaxBFGS(floglik, &vtheta0, &dfuncl, 0, 0);

/* Verificando a convergência */
if(converge != MAX_CONV && converge != MAX_WEAK_CONV)
{
    falha++;
    fail++;
    continue; /* sai do loop MC */
}

decl    etahat = s_mX * vtheta0[0:(k-1)];
decl    muhat  = exp(etahat) ./ (1+exp(etahat));
decl    T      = diag(muhat .* (1-muhat));
decl    T2     = diag( exp(etahat) ./ (1.0+exp(etahat)) .^2 );
decl    T1     = diag( exp(etahat) ./ ((1.0+exp(etahat)) .^2 ));
decl    deltahat = s_mZ * vtheta0[k:(k+m-1)];
decl    phihat = exp(deltahat);
decl    H      = diag(phihat);
decl    s_mYstar = diag(ystar);
decl    s_mYdag = diag(ydag);
decl    M      = diag(muhat);
decl    F      = diag(phihat);
decl    Mstar  = diag(polygamma(muhat .* phihat, 0) - polygamma((1.0 - muhat).* phihat, 0));
                                                    /*média de ystar=log(y/(1-y))*/
decl    Vstar  = diag(polygamma(muhat .* phihat, 1) + polygamma((1.0 - muhat) .* phihat, 1));
                                                    /*variância de ystar*/
decl    Mdag   = diag(polygamma((1.0 - muhat) .* phihat, 0) - polygamma(phihat, 0));
                                                    /*média de ydag=log(1-y)*/
decl    Vdag   = diag(polygamma((1.0 - muhat) .* phihat, 1) - polygamma(phihat, 1));
                                                    /*variância de ydag*/
decl    C      = diag(-polygamma((1 - muhat) .* phihat, 1)); /*covariância entre ystar e ydag*/
decl    S      = diag((2*muhat-1) ./ ((muhat .* (1-muhat)) .^ 2)); /*g' (mu)*/
decl    Q      = diag(-1 ./ (phihat .^ 2)); /*h' (phi)*/
decl    Ibetabeta = s_mX'*F*T*Vstar*T+F*s_mX;
decl    Ibetagama = s_mX'*F*(M*Vstar+C)*T*H*s_mZ;
decl    Igamabeta = Ibetabeta';
decl    Igamagama = s_mZ'*H*(M*Vstar*M+(M+M)*C+Vdag)*H*s_mZ;
decl    fisher = (Ibetabeta~Ibetagama)|(Igamabeta~Igamagama); /*informação de Fisher*/
decl    fisherinv = invert(fisher); /*inversa da informação de Fisher*/
decl    Jbetabeta = s_mX'*(F*T*Vstar + S*T^2*(s_mYstar-Mstar))*T*F*s_mX;
decl    Jbetagama = -s_mX'*((s_mYstar-Mstar) - F*(M*Vstar+C))*T*H*s_mZ;
decl    Jgamabeta = Jbetabeta';
decl    Jgamagama = s_mZ'*(H*(M*Vstar*M+(M+M)*C+Vdag) + (M*(s_mYstar-Mstar) + (s_mYdag-Mdag))*Q*H^2)
                                                    *H*s_mZ;

decl    infobs = (Jbetabeta~Jbetagama)|(Jgamabeta~Jgamagama);
decl    infobsinv = invert(infobs);
decl    stderrors = sqrt(diagonal(fisherinv));

/***** Cálculo com restrição *****/
/*Calculando as Estimativas de Momentos sugeridas por Ferrari e Cribari-Neto (2004)*/

if(k > 1){ols2c(ystar, s_mXr, &betaolsr);}
/* Retorna em betaols a Estimativa de Mínimos Quadrados Ordinários, ou seja,
(X'X)^-1 * X'*ynew */
else if(k==1){betaolsr = meanc(ystar);}

/* Valores Iniciais para os Parâmetro (beta, phi) */
vtheta1 = betaolsr | gamaols;

MaxControl( 50, -1 );

/* Estimação de máxima verossimilhança restrita */
converger = MaxBFGS(floglikr, &vtheta1, &dfunc2, 0, 0);

```

```

if (dfunc1-dfunc2 < 0){print("\n RV<0");}

/* Verificando a convergência */
if(converger != MAX_CONV && converger != MAX_WEAK_CONV)
{
  print("\n maximização não converge!");
  continue; /* sai do loop MC */
}

decl      etatil = s_mX*(vthetal[0:(k-1-r)]| beta0');
decl      deltatil = s_mZ*(vthetal[(k-r):(k+m-1-r)] );
decl      mutil = exp(etatil) ./ (1+exp(etatil));
decl      Ttil = diag( mutil .* (1-mutil) );
decl      phitil = exp(deltatil);
decl      Htil = diag(phitil);
decl      Mtil = diag(mutil);
decl      Ftil = diag(phitil);
decl      mustartil = polygamma(mutil .* phitil, 0) - polygamma((1.0 - mutil) .* phitil, 0);
                                                    /*média de ystar=log(y/(1-y))*/
decl      Mstartil = diag(mustartil);
decl      mudagtil = polygamma((1.0 - mutil) .* phitil, 0) - polygamma(phitil, 0);
                                                    /*média de ydag=log(1-y)*/
decl      Mdagtil = diag(mudagtil);
decl      Vstartil = diag(polygamma(mutil .* phitil, 1) + polygamma((1.0 - mutil) .* phitil, 1));
                                                    /*variância de ystar*/
decl      Vdagtil = diag(polygamma((1.0 - mutil) .* phitil, 1)- polygamma(phitil, 1));
                                                    /*variância de ydag*/
decl      Ctil = diag(-polygamma((1 - mutil) .* phitil, 1));/*covariância entre ystar e ydag*/
decl      Stil = diag((2*mutil-1) ./ ((mutil .* (1-mutil)).^ 2));          /*g'' (mu)*/
decl      Qtil = diag(-1 ./ (phitil .^ 2));          /*h'' (phi)*/
decl      Ibetabetatil = s_mX'*Ftil*Ttil*Vstartil*Ttil*Ftil*s_mX;
decl      Ibetagamatil = s_mX'*Ftil*(Mtil*Vstartil+Ctil)*Ttil*Htil*s_mZ;
decl      Igamabetatil = Ibetagamatil';
decl      Igamagamatil = s_mZ'*Htil*(Mtil*Vstartil*Mtil+(Mtil+Mtil)*Ctil+Vdagtil)*Htil*s_mZ;
decl      fishertil = (Ibetabetatil ~ Ibetagamatil) | (Igamabetatil ~ Igamagamatil);
                                                    /*informação de Fisher*/
decl      fisherinvtil = invert(fishertil);          /*inversa da informação de Fisher*/
decl      Jbetabetatil = s_mX'*(Ftil*Ttil*Vstartil + Stil*Ttil^2*(s_mYstar-Mstartil))*Ttil*Ftil*s_mX;
decl      Jbetagamatil = -s_mX'*((s_mYstar-Mstartil)- Ftil*(Mtil*Vstartil+Ctil))*Ttil*Htil*s_mZ;
decl      Jgamabetatil = Jbetagamatil';
decl      Jgamagamatil = s_mZ'*
      (Htil*(Mtil*Vstartil*Mtil+(Mtil+Mtil)*Ctil+Vdagtil)+
      (Mtil*(s_mYstar-Mstartil)+(s_mYdag-Mdagtil))*Qtil*Htil^2)*
      Htil*s_mZ;
decl      infobstil = (Jbetabetatil~Jbetagamatil) | (Jgamabetatil~Jgamagamatil);
                                                    /*informação observada*/
decl      infobsinvtil = invert(infobstil);          /*inversa da informação observada*/
decl      escorebetatil = s_mX'*Ttil*Ftil*(ystar-mustartil);
decl      escoregamatil = s_mZ'*Htil*(Mtil*(ystar-mustartil)+(ydag-mudagtil));
decl      escoretil = escorebetatil | escoregamatil;          /* função escore */

/*Obtendo as quantidades restantes para o cálculo do ajuste de Skovgaard*/

decl      iota = ones(nobs,1);
decl      q1 = s_mX'*T*F*(Vstar*(M*F-Mtil*Ftil)+ C*(F-Ftil))*iota;
decl      q2 = s_mZ'*H*((M*F-Mtil*Ftil)*(M*Vstar+C)+(F-Ftil)*(M*C+Vdag))*iota;
decl      q = q1|q2;
decl      upsilon1 = s_mX'*F*T*Vstar*Ttil*Ftil*s_mX;
decl      upsilon2 = s_mX'*T*F*(Mtil*Vstar+C)*Htil*s_mZ;
decl      upsilon3 = s_mZ'*H*Ftil*(M*Vstar+C)*Ttil*s_mX;
decl      upsilon4 = s_mZ'*H*(M*Vstar*Mtil+(M+Mtil)*C+Vdag)*Htil*s_mZ;
decl      upsilon = (upsilon1~upsilon2) | (upsilon3~upsilon4);
decl      upsiloninv = invert(upsilon);
decl      nuifinobstil = nuisance(infobstil);
decl      nuisance2 = nuisance(fishertil*upsiloninv*infobs*fisherinv*upsilon);

/* Estatísticas do Teste: RV (w) e Skovgaard (w* e w**) */
RV = 2*(dfunc1-dfunc2);          /* Razão de Verossimilhanças (w) */

```

```

decl nu = fabs(fabs(determinant(fishertil))^(1/2)* fabs(determinant(fisher))^(1/2)*
  fabs(determinant(upsilon))^-1)* fabs(determinant(nuiinfobstil))^(1/2)*
  fabs(determinant(nuisance2))^-1/2)*
  fabs(escoretil'*upsiloninv*fisher*infobsinv*upsilon* fisherinvtil*escoretil)^(r/2)/
  fabs((RV)^(r/2)-1)*escoretil'* upsiloninv*q));

if(RV<0.1) /*forte evidência para não rejeitar a hipótese nula*/
{
  Skov1 = RV;
  Skov2 = RV;
}
else
{
  Skov1 = RV-2*log(nu);
  Skov2 = RV*(1-(log(nu)/RV))^2;
}

/* Impressão dos Resultados: */
print( "\n PROGRAMA OX: ", oxfilename(0) );
print( "\n VERSAO OX: ", oxversion() );
print( "\n DISTRIBUICAO: beta " );
print( "\n NOME DO ARQUIVO DE DADOS: ", nome_arquivo);
print( "\n TAMANHO DAS AMOSTRAS: ", nobs );
print( "\n NUMERO DE COVARIADAS: ", k+m);
print( "\n\n#####");
print( "\nw* = w se w < 0.1 e w* = w - 2*log(nu) c.c");
print( "\nw** = w se w < 0.1 e w** = w *(1- (1/RV * log(nu)))^2 c.c");
print( "\n#####");
print( "\n\nNumero de parametros no modelo para mu = k = ",k);
print( "\n\nNumero de parametros no modelo para phi = m = ",m);
print( "\n\nNumero de variaveis independentes no modelo para mu = k - 1 = ",k-1);
print( "\n\nNumero de parametros testados = r = ",r);
print( "\n\nHipotese Nula: Ho: beta_(k-r+1)=...=beta_(k)= 0");
print( "\n\nResultados da Estimacao dos Parametros (beta, gama)");
print( "\n", "%8.3f",
  "%r", {"Media", "DP"},
  "%c", {"beta_1","beta_2","beta_3","beta_4","beta_5",
    "beta_6","beta_7","beta_8","beta_9","beta_10"},
  vtheta0[:(k-1)]' | stderrors[:(k-1)]);
print( "%8.3f",
  "%r", {"Media", "DP"},
  "%c", {"gama_1","gama_2","gama_3","gama_4","gama_5",
    "gama_6","gama_7","gama_8","gama_9","gama_10"},
  vtheta0[k:(k+m-1)]' | stderrors[k:(k+m-1)]);

decl resul=zeros(2,3);
resul[0][0]=RV;
resul[0][1]=Skov1;
resul[0][2]=Skov2;
resul[1][0]=1-probchi(RV,r);
resul[1][1]=1-probchi(Skov1,r);
resul[1][2]=1-probchi(Skov2,r);
print( "%10.2f",
  "%c",{"w", "w*", "w**"},
  "%r", {"-", "p-valor"} ,resul);

/* Data, Hora e Tempo de Execução */
print( "\nDATA: ", date() );
print( "\nHORA: ", time(), "\n" );
print( "\n\n\t\t TEMPO TOTAL DE EXECUCAO: ", timespan(dExecTime) );
print( "\n" );
}
/*----- Fim do Programa -----*/

```





## Referências bibliográficas

- Atkinson, A. C. (1985). *Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*. New York: Oxford University Press.
- Barndorff-Nielsen, O.E. (1986). Inference on full or partial parameters, based on the standardized signed log likelihood ratio. *Biometrika*, **73**, 307-322.
- Barndorff-Nielsen, O.E. (1991). Modified signed log likelihood ratio. *Biometrika*, **78**, 557-563.
- Brownlee, K. A. (1965). *Statistical Theory and Methodology in Science and Engineering*. 2nd ed. London: John Wiley & Sons.
- Cordeiro, G. M. (1987). On the corrections to the likelihood ratio statistics. *Biometrika*, **74**, 265-74.
- Cordeiro, G.M., Ferrari, S.L.P. (1991). A modified score test statistic having chi-squared distribution to order  $n^{-1}$ . *Biometrika*, **78**, 573-582.
- Cribari-Neto, F., Ferrari, S.L.P. (1995). Second order asymptotics for score tests in generalised linear models. *Biometrika*, **82**, 426-432.
- Doornik, J. A. (2006). *An Object-Oriented Matrix Language – Ox 4*. 5th ed. London: Timberlake Consultants Press.
- Espinheira, P.L., Ferrari, S.L.P., Cribari-Neto, F. (2008a). Influence diagnostics in beta regression. *Computational Statistics and Data Analysis*, **52**, 4417-4431.
- Espinheira, P.L., Ferrari, S.L.P., Cribari-Neto, F. (2008b). On beta regression residuals. *Journal of Applied Statistics*, **35**, 407-419.
- Ferrari, S.L.P., Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, **31**, 799-815.
- Ferrari, S.L.P., Cysneiros, A.H.M.A. (2009). Skovgaard's adjustment to likelihood ratio tests in exponential family nonlinear models. *Statistics and Probability Letters*. Disponível on line. DOI: 10.1016/j.spl.2008.05.009.
- Ihaka, R., Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of Computational Graphics and Statistics*, **5**, 299-314.
- Lawley, D. (1956). A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika*, **43**, 295-303.
- Lehmann, E.L., Casella, E. (1998). *Theory of Point Estimation*, 2nd ed. New York: Springer-Verlag.
- McCullagh, P., Nelder, J.A. (1989). *Generalized Linear Models*. 2nd ed. London: Chapman and Hall.
- Oliveira, M.S. (2004). *Um Modelo de Regressão Beta: Teoria e Aplicações*. Dissertação de Mestrado, IME - Universidade de São Paulo.
- Severini, T.A. (2000). *Likelihood Methods in Statistics*. Oxford: Oxford University Press.
- Skovgaard, I.M. (2001). Likelihood asymptotics. *Scandinavian Journal of Statistics*, **28**, 3-32.
- Smithson, M., Verkuilen, J. (2006). A better lemon squeezer? Maximum likelihood regression with beta distributed dependent variables. *Psychological Methods*, **11**, 54-71.
- Wei, B.C. (1998). *Exponential Family Nonlinear Models*. Singapore: Springer-Verlag.