

**Aperfeiçoamento de métodos estatísticos  
em modelos de regressão da família  
exponencial**

Alexsandro Bezerra Cavalcanti

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

Programa: Estatística

Orientadora: Profa. Dra. Denise Aparecida Botter

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da  
CAPES/PICDT

São Paulo, Agosto de 2009

# Aperfeiçoamento de métodos estatísticos em modelos de regressão da família exponencial

Este exemplar corresponde à redação  
final da tese devidamente corrigida  
e defendida por Alexsandro Bezerra Cavalcanti  
e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Profa. Dra. Denise Aparecida Botter (orientador) - IME-USP.
- Profa. Dra. Lúcia Pereira Barroso - IME-USP.
- Prof. Dr. Gauss Moutinho Cordeiro - DF-UFRPE.
- Prof. Dr. Klaus P. Vasconcelos - DE-UFPE
- Audrey Helen Mariz de Aquino Cysneiros - DE-UFPE

## Agradecimentos

- A Deus, o autor e consumidor da minha fé.
- À professora Denise A. Botter, por todo o seu empenho, apoio, disponibilidade, confiança e competência, fatores que foram fundamentais para a realização deste trabalho.
- A minha esposa Maria Hosana e meus filhos Tales e Natan, pelo apoio incondicional e espírito de renúncia.
- Ao professor Gauss M. Cordeiro, por suas ideias brilhantes.
- À professora Lúcia P. Barroso, por sua participação direta em todo este trabalho.
- Ao Instituto de Matemática e Estatística da USP, por toda a estrutura oferecida durante todo o curso.
- À Unidade Acadêmica de Matemática e Estatística da UFCG, por todo empenho e esforço realizados para minha liberação.
- Aos amigos Ronald Targino e Artur Lemonte pelos constantes esclarecimentos no uso do programa R.
- À Michelli Karinne Barros da Silva, professora da UFCG, por sua contribuição na parte computacional do Capítulo 4.
- À CAPES, pelo apoio financeiro através do PICDT (Programa Institucional de Capacitação Docente e Técnico).
- A todos que de alguma forma contribuíram para a realização deste trabalho.



## Resumo

Neste trabalho, desenvolvemos três tópicos relacionados a modelos de regressão da família exponencial. No primeiro tópico, obtivemos a matriz de covariância assintótica de ordem  $n^{-2}$ , onde  $n$  é o tamanho da amostra, dos estimadores de máxima verossimilhança corrigidos pelo viés de ordem  $n^{-1}$  em modelos lineares generalizados, considerando o parâmetro de precisão conhecido. No segundo tópico calculamos o coeficiente de assimetria assintótico de ordem  $n^{-1/2}$  para a distribuição dos estimadores de máxima verossimilhança dos parâmetros que modelam a média e dos parâmetros de precisão e dispersão em modelos não-lineares da família exponencial, considerando o parâmetro de dispersão desconhecido, porém o mesmo para todas as observações. Finalmente, obtivemos fatores de correção tipo-Bartlett para o teste escore em modelos não-lineares da família exponencial, considerando covariáveis para modelar o parâmetro de dispersão. Avaliamos os resultados obtidos nos três tópicos desenvolvidos por meio de estudos de simulação de Monte Carlo.



## Abstract

In this work, we develop three topics related to the exponential family nonlinear regression. First, we obtain the asymptotic covariance matrix of order  $n^{-2}$ , where  $n$  is the sample size, for the maximum likelihood estimators corrected by the bias of order  $n^{-1}$  in generalized linear models, considering the precision parameter known. Second, we calculate an asymptotic formula of order  $n^{-1/2}$  for the skewness of the distribution of the maximum likelihood estimators of the mean parameters and of the precision and dispersion parameters in exponential family nonlinear models considering that the dispersion parameter is the same although unknown for all observations. Finally, we obtain Bartlett-type correction factors for the score test in exponential family nonlinear models assuming that the precision parameter is modelled by covariates. Monte Carlo simulation studies are developed to evaluate the results obtained in the three topics.





# Sumário

<b>Lista de Tabelas</b>	<b>ix</b>
<b>1 Introdução</b>	<b>1</b>
<b>2 Matriz de Covariância do EMV corrigido</b>	<b>5</b>
2.1 Introdução . . . . .	5
2.2 Definição do modelo . . . . .	6
2.3 Matriz de covariância assintótica de ordem $n^{-2}$ . . . . .	7
2.4 Testes de Wald modificados . . . . .	10
2.5 Resultados de simulação . . . . .	11
2.6 Conclusões . . . . .	14
<b>3 Coeficiente de assimetria assintótico</b>	<b>15</b>
3.1 Introdução . . . . .	15
3.2 Modelos não-lineares da família exponencial . . . . .	16
3.3 Coeficiente de assimetria assintótico de $\hat{\beta}$ . . . . .	18
3.4 Coeficiente de assimetria assintótico para $\hat{\phi}$ e $\hat{\sigma}^2$ . . . . .	22
3.5 Dois exemplos . . . . .	23
3.6 Resultados de simulações . . . . .	24

---

3.7	Aplicações . . . . .	26
3.7.1	Elasticidade constante de substituição (CES) da função de produção . . . . .	26
3.7.2	Modelo de crescimento do pasto . . . . .	29
3.8	Conclusões . . . . .	30
<b>4</b>	<b>Aperfeiçoamento de testes escore em MNLFs</b>	<b>31</b>
4.1	Introdução . . . . .	31
4.2	Definição do modelo . . . . .	32
4.3	Melhoramento do teste escore . . . . .	35
4.3.1	Testando apenas componentes do vetor de parâmetros $\delta$ . . . . .	40
4.3.2	Testando apenas componentes do vetor de parâmetros $\beta$ . . . . .	43
4.4	Resultados de simulação . . . . .	45
4.5	Aplicação . . . . .	51
4.6	Conclusões . . . . .	51
<b>5</b>	<b>Pesquisas futuras</b>	<b>53</b>
<b>A</b>	<b>Obtenção de alguns cumulantes em MNLFs</b>	<b>55</b>
<b>B</b>	<b>Obtenção dos coeficientes <math>A_1</math>, <math>A_2</math> e <math>A_3</math></b>	<b>63</b>
<b>C</b>	<b>Conjuntos de dados</b>	<b>73</b>
C.1	Elasticidade constante de substituição da função de produção . . . . .	73
C.2	Modelo de crescimento do pasto . . . . .	74
	<b>Referências Bibliográficas</b>	<b>75</b>

## Lista de Tabelas

2.1	Cov( $\hat{\beta}$ ) avaliada em $\hat{\beta}$ , EQM( $\hat{\beta}$ ) e Cov( $\tilde{\beta}$ ) avaliada em $\tilde{\beta}$ . . . . .	13
2.2	Tamanho do teste para as estatísticas $W$ , $W_m$ e $W_c$ . . . . .	14
3.1	Coefficientes de assimetria amostrais e analíticos para os modelos exponencial e gama .	27
3.2	Coefficientes de assimetria amostrais e analíticos para o modelo normal . . . . .	27
3.3	Estimativas dos coeficientes de assimetria da função de produção CES . . . . .	29
3.4	Estimativas dos coeficientes de assimetria para o modelo de crescimento do pasto . . .	29
4.1	Alguns modelos especiais . . . . .	33
4.2	Tamanho dos testes para a hipótese $H_0^1$ com $p = 2, 3$ e $r = 2$ . . . . .	47
4.3	Tamanho dos testes para a hipótese $H_0^1$ com $p = 4, 5$ e $r = 2$ . . . . .	48
4.4	Tamanho dos testes para a hipótese $H_0^2$ com $q = 1, 2$ e $r = 2$ . . . . .	49
4.5	Tamanho dos testes para a hipótese $H_0^2$ com $q = 4, 5$ e $r = 2$ . . . . .	50
C.1	<i>Output</i> agregado em uma certa indústria ( $y_t$ ), trabalho ( $L_t$ ) e capital ( $K_t$ ), numa amostra de 30 observações. . . . .	73
C.2	Taxa de crescimento do pasto ( $y_i$ ) e o tempo decorrido desde o último corte do pasto ( $x_i$ ) para uma amostra de 9 observações. . . . .	74



# Capítulo 1

## Introdução

Os modelos de regressão da família exponencial, tanto os lineares generalizados (Nelder & Wedderburn, 1972) como os não-lineares (Cordeiro & Paula, 1989), são bastante utilizados em diversas áreas do conhecimento. O objetivo principal desta tese é desenvolver técnicas estatísticas que melhorem as propriedades assintóticas dos estimadores de máxima verossimilhança (EMVs) e da estatística do teste escore na classe dos modelos de regressão da família exponencial para pequenas amostras, quando, em geral, estas propriedades não são satisfatórias.

Iremos, inicialmente, destacar algumas das contribuições mais relevantes em termos de teoria assintótica na classe dos modelos lineares generalizados (MLGs). Para pequenas amostras, a aproximação  $\chi^2$  usual para os testes da razão de verossimilhanças e o teste escore não é satisfatória. Visando o aperfeiçoamento desta aproximação, Cordeiro (1983, 1987) obteve correções de Bartlett para a estatística da razão de verossimilhanças, enquanto Cordeiro & Ferrari (1991), Cordeiro *et al.* (1993) e Cribari-Neto & Ferrari (1995) obtiveram correções tipo-Bartlett para a estatística escore. Mais recentemente, Cordeiro *et al.* (2003) compararam a correção tipo-Bartlett para a estatística escore, supondo o parâmetro de dispersão desconhecido e variável, com a correção obtida por Kakizawa (1996) e a obtida por Cordeiro *et al.* (1998). Correções do viés de ordem  $n^{-1}$ , onde  $n$

é o tamanho da amostra, para o EMV foram obtidas por Cordeiro & McCullagh (1991) e Botter & Cordeiro (1998), sendo que o segundo trabalho modelou o parâmetro de dispersão por meio de covariáveis. Uma expressão para o coeficiente de assimetria assintótico de ordem  $n^{-1/2}$  para a distribuição dos EMVs dos parâmetros  $\beta$  que modelam a média e para o parâmetro de precisão foi obtida por Cordeiro & Cordeiro (2001). Em Cordeiro (2004) foi apresentada uma fórmula geral para a matriz de covariância assintótica de ordem  $n^{-2}$  dos EMVs do parâmetro  $\beta$ , considerando o parâmetro de dispersão conhecido. Este resultado foi estendido por Cordeiro *et al.* (2006) considerando o parâmetro de dispersão desconhecido, porém o mesmo para todas as observações.

Os modelos de regressão não-lineares da família exponencial (MNLFEs) são uma extensão dos MLGs e dos modelos de regressão normal não-lineares. Existe uma vasta bibliografia que trata dos modelos de regressão normal não-lineares, por exemplo Ratkowsky (1983, 1990), Gallant (1987), Bates & Watts (1988), McCullagh & Nelder (1989), etc. Por outro lado, para os MNLFEs, existe uma bibliografia bem reduzida tratando do assunto. O livro de Wei (1998) apresenta uma discussão bem detalhada dando ênfase à análise de diagnóstico e medidas de influência. Seguindo a linha dos MLGs, existem alguns resultados para os MNLFEs. Cordeiro & Paula (1989) discutiram melhoramentos na estatística da razão de verossimilhanças através da correção de Bartlett, quando o parâmetro de dispersão é conhecido. Cysneiros & Ferrari (2006) obtiveram uma correção de Bartlett para a estatística do teste da razão de verossimilhanças perfilada modificada considerando os parâmetros que modelam a dispersão variáveis. Ferrari & Cysneiros (2008) utilizaram o ajuste de Skovgaard (Skovgaard, 2001) para a estatística da razão de verossimilhanças. Ferrari & Cordeiro (1996) e Ferrari *et al.* (1997) obtiveram fatores de correção tipo-Bartlett para a estatística score considerando o parâmetro de precisão  $\phi$  constante. Correções de viés de ordem  $n^{-1}$  para EMVs foram obtidas por Paula (1992) e Cordeiro & Udo (2008), este último considerando covariáveis para modelar o parâmetro de dispersão. Em Cordeiro & Santana (2008) foi apresentada uma fórmula geral para a

matriz de covariância assintótica de ordem  $n^{-2}$  dos EMVs, generalizando o resultado de Cordeiro (2004).

Neste trabalho abordamos três tópicos relacionados aos modelos de regressão da família exponencial:

1. Obtenção da matriz de covariância assintótica de ordem  $n^{-2}$  dos EMVs corrigidos pelo viés de ordem  $n^{-1}$  em MLGs, considerando o parâmetro de precisão conhecido;
2. Cálculo do coeficiente de assimetria assintótico de ordem  $n^{-1/2}$  para a distribuição dos EMVs dos parâmetros que modelam a média e dos parâmetros de precisão e dispersão em MNLFs, considerando o parâmetro de dispersão desconhecido, porém o mesmo para todas as observações;
3. Obtenção de fatores de correção tipo-Bartlett para o teste score em MNLFs, considerando covariáveis para modelar o parâmetro de dispersão.

Esta tese está organizada em 5 capítulos e 3 apêndices. No Capítulo 2 apresentamos a matriz de covariâncias de ordem  $n^{-2}$  do EMV do parâmetro que modela a média corrigido pelo viés de ordem  $n^{-1}$  para MLGs, considerando o parâmetro de dispersão conhecido. Mostramos, através de estudos de simulação, que é possível melhorar a estimação da variância do estimador corrigido e assim obter estimativas intervalares mais precisas, principalmente quando o tamanho da amostra é pequeno ou moderado.

No Capítulo 3 apresentamos uma fórmula para o coeficiente de assimetria assintótico de ordem  $n^{-1/2}$  da distribuição dos EMVs dos parâmetros que modelam a média, a precisão e dispersão em modelos não-lineares da família exponencial quando o parâmetro de dispersão é desconhecido, porém o mesmo para todas as observações. O resultado segue da expressão geral do coeficiente de assimetria assintótico para o EMV dada por Bowman & Shenton (1998). Este trabalho generaliza o resultado

obtido por Cordeiro & Cordeiro (2001) para os MLGs. O cálculo do coeficiente de assimetria possibilita investigar a não adequação da aproximação normal para a distribuição do EMV em pequenas amostras.

No Capítulo 4 obtemos fatores de correção tipo-Bartlett para o teste escore dos parâmetros que modelam a média e a dispersão, em modelos não-lineares da família exponencial, com dispersão variável. Os resultados obtidos se aplicam apenas às distribuições normal e normal inversa, visto que o modelo gama não admite reparametrização que torne os parâmetros ortogonais no sentido abordado em Cox & Reid (1987). Estudos de simulação mostram que a estatística escore é bastante conservativa e, portanto tende à não rejeição da hipótese nula quando a mesma é verdadeira. Através de estudos de simulação mostramos que a estatística corrigida é razoavelmente melhor do que a estatística escore usual.

Finalmente, no Capítulo 5 apresentamos algumas considerações finais e alguns tópicos que poderão ser desenvolvidos em pesquisas futuras.

Os Capítulos 2, 3 e 4 foram escritos de maneira independente. Assim, algumas notações e definições são apresentadas mais de uma vez com o objetivo de que cada capítulo possa ser lido em qualquer ordem.



## Capítulo 2

# Matriz de covariância do estimador de máxima verossimilhança para o parâmetro $\beta$ corrigido pelo viés em modelos lineares generalizados

### 2.1 Introdução

Os métodos para análise de um modelo linear generalizado (MLG) dependem fortemente de propriedades assintóticas dos estimadores de máxima verossimilhança (EMVs) quando o tamanho  $n$  da amostra cresce. Muitas pesquisas têm sido realizadas com o intuito de desenvolver uma teoria assintótica de segunda ordem para os MLGs, ou seja melhorar a inferência por verossimilhança. Expressões para o EMV do parâmetro que modela a média corrigido pelo viés de ordem  $n^{-1}$  em modelos da família exponencial uniparamétrica foram obtidas por Ferrari *et al.* (1996). Nesse trabalho, os autores obtiveram também a variância e o erro quadrático médio do EMV corrigido. Em Cordeiro (2004) foi apresentada uma fórmula geral até ordem  $O(n^{-2})$  para a matriz de covariância dos EMVs para o vetor de parâmetros  $\beta$  que modelam a média em MLGs, considerando o parâmetro de precisão  $\phi$  conhecido. Nosso objetivo neste capítulo é obter a matriz de covariância assintótica de ordem  $n^{-2}$

dos EMVs para o vetor de parâmetros  $\beta$  corrigidos pelo viés de ordem  $n^{-1}$  em MLGs supondo  $\phi$  conhecido. Será mostrado que essa matriz é o resultado da soma de parcelas que envolvem a matriz de covariância assintótica de ordem  $n^{-2}$  e o vício de ordem  $n^{-1}$  do EMV de  $\beta$  nos MLGs. Esperamos, com o cálculo desta matriz, obter estimativas mais precisas para as variâncias dos EMVs corrigidos. Estas estimativas podem ser usadas para melhorar a precisão do teste de Wald, pois é muito comum corrigir o viés do estimador e utilizar a matriz de informação de Fisher no cálculo da estatística do teste de Wald. Uma outra aplicação é, simplesmente, calcular intervalos de confiança mais precisos para o parâmetro  $\beta$ .

## 2.2 Definição do modelo

Suponha as variáveis  $Y_1, \dots, Y_n$  independentes com cada  $Y_\ell$  tendo função densidade de probabilidade ou função de probabilidade na forma

$$\pi(y; \theta_\ell, \phi) = \exp\{\phi[y\theta_\ell - b(\theta_\ell)] + c(y, \phi)\}, \quad (2.1)$$

onde  $b(\cdot)$  e  $c(\cdot, \cdot)$  são funções conhecidas,  $\mu_\ell = E(Y_\ell) = db(\theta_\ell)/d\theta_\ell$  é a média de  $Y_\ell$ ,  $\theta_\ell = q(\mu_\ell)$  varia num subconjunto da reta e o parâmetro  $\phi$  varia num subconjunto de  $R_+$ , sendo constante sobre todas as observações. A variância de  $Y_\ell$  é dada por  $\text{Var}(Y_\ell) = \phi^{-1}V_\ell$ , sendo  $V_\ell = d\mu_\ell/d\theta_\ell$  denominada função de variância.

Os MLGs são definidos por (2.1) e pela componente sistemática parametrizada como

$$t(\mu) = \eta = X\beta,$$

onde  $X$  é a matriz de planejamento de dimensão  $n \times p$  de posto completo e  $\beta = (\beta_1, \dots, \beta_p)^\top$  é um vetor de parâmetros desconhecido a ser estimado. Assumimos que  $t(\cdot)$  é uma função monótona,

contínua, conhecida e duas vezes diferenciável. O EMV  $\hat{\beta}$  de  $\beta$ , assumido disponível neste trabalho, pode ser obtido pelo método de Newton-Raphson. Por simplicidade, assumimos que o parâmetro de precisão  $\phi$  em (2.1) é conhecido ou pode ser substituído por uma estimativa consistente  $\hat{\phi}$  de modo que a função em (2.1) pertença à família exponencial de distribuições com parâmetro natural  $\theta_\ell$ .

### 2.3 Matriz de covariância assintótica de ordem $n^{-2}$

Nesta seção nosso objetivo é obter a matriz de covariância assintótica de ordem  $n^{-2}$  dos EMVs corrigidos pelo viés de ordem  $n^{-1}$  nos MLGs.

Vamos denotar o logaritmo da função de verossimilhança para  $\beta$  por  $\mathcal{L} = \mathcal{L}(\beta)$  e os cumulantes conjuntos das derivadas do logaritmo da função de verossimilhança por  $\kappa_{rs} = E(\partial^2 \mathcal{L} / \partial \beta_r \partial \beta_s)$ ,  $\kappa_{r,s} = E(\partial \mathcal{L} / \partial \beta_r \partial \mathcal{L} / \partial \beta_s)$ ,  $\kappa_{rst} = E(\partial^3 \mathcal{L} / \partial \beta_r \partial \beta_s \partial \beta_t)$ ,  $\kappa_{r,st} = E(\partial \mathcal{L} / \partial \beta_r \partial^2 \mathcal{L} / \partial \beta_s \partial \beta_t)$ , etc, todos os índices variam pelos inteiros  $1, \dots, p$ . Todos os  $\kappa$ 's referem-se a um total sobre a amostra e são, em geral, de ordem  $O(n)$ . A matriz de informação total de Fisher de  $\beta$ ,  $K_\beta$ , tem elementos  $\kappa_{r,s} = -\kappa_{rs}$  e seja  $\kappa^{r,s} = -\kappa^{rs}$  os elementos correspondentes de sua matriz inversa  $K_\beta^{-1}$ . Assumimos que, quando  $n$  cresce, o EMV  $\hat{\beta}$  de  $\beta$  converge para seu verdadeiro valor e que sua distribuição assintótica é normal multivariada  $N_p(\beta, \phi^{-1}(X^\top W X)^{-1})$ , onde  $W = \text{diag}\{(d\mu/d\eta)^2 V^{-1}\}$ .

Seja  $\tilde{\beta} = \hat{\beta} - d(\hat{\beta})$ , o EMV corrigido pelo viés de ordem  $n^{-1}$ , onde  $d(\hat{\beta})$  é o viés de ordem  $n^{-1}$  de  $\beta$  avaliado em  $\hat{\beta}$ . Considere  $\tilde{\beta}_r$  o  $r$ -ésimo elemento de  $\tilde{\beta}$ . Temos então  $\tilde{\beta}_r = \hat{\beta}_r - d^r(\hat{\beta})$ , onde  $d^r(\hat{\beta})$  é o  $r$ -ésimo elemento de  $d(\hat{\beta})$ .

De Pace & Salvan (1997, p. 360) vem que

$$d^r(\hat{\beta}) = d^r(\beta) + \sum_v d_v^r(\hat{\beta}_v - \beta_v) + O_p(n^{-2}),$$

onde

$$\begin{aligned}
d_v^r &= \frac{\partial d^r}{\partial \beta_v} \\
&= \sum_{w,s,y,t,u} \{ \kappa^{rw} \kappa^{sy} \kappa^{tu} (\kappa_{stu} + 2\kappa_{st,u}) (\kappa_{vwy} + \kappa_{v,wy}) \\
&\quad + \frac{1}{2} \kappa^{rs} \kappa^{tu} (\kappa_{stuv} + \kappa_{stu,v} + 2\kappa_{stv,u} + 2(\kappa_{st,uv} + \nu_{st,u,v})) \} \\
&= O(n^{-1})
\end{aligned}$$

e

$$d^r(\beta) = \frac{1}{2} \sum_{s,t,u} \kappa^{rs} \kappa^{tu} (\kappa_{stu} + 2\kappa_{st,u}), \quad r = 1, \dots, p.$$

Queremos encontrar uma expressão de ordem  $n^{-2}$  para o elemento  $rs$  da matriz  $\text{Cov}(\tilde{\beta})$ , ou seja,

$$[\text{Cov}(\tilde{\beta})]_{rs} = E[(\tilde{\beta}_r - \beta_r)(\tilde{\beta}_s - \beta_s)].$$

Temos que

$$\begin{aligned}
[\text{Cov}(\tilde{\beta})]_{rs} &= E[(\hat{\beta}_r - d^r(\hat{\beta}) - \beta_r)(\hat{\beta}_s - d^s(\hat{\beta}) - \beta_s)] \\
&= E\{[(\hat{\beta}_r - \beta_r) - d^r(\hat{\beta})][(\hat{\beta}_s - \beta_s) - d^s(\hat{\beta})]\} \\
&= E[(\hat{\beta}_r - \beta_r)(\hat{\beta}_s - \beta_s) - d^s(\hat{\beta})(\hat{\beta}_r - \beta_r) - d^r(\hat{\beta})(\hat{\beta}_s - \beta_s) + d^r(\hat{\beta})d^s(\hat{\beta})]. \quad (2.2)
\end{aligned}$$

A primeira parcela da expressão (2.2) é a covariância de ordem  $n^{-2}$  de  $\hat{\beta}$  ( $\phi$  conhecido) obtida por Cordeiro (2004), que é dada por

$$\phi^{-1}(X^\top W X)^{-1} + \phi^{-2} P \Lambda P^\top + \phi^{-2} (X^\top W X)^{-1} \Delta (X^\top W X)^{-1},$$

onde  $\Lambda = HZ_d + \frac{3}{2}FZ^{(2)}F + GZ^{(2)}F - GZ^{(2)}G$ ,  $P = (X^\top WX)^{-1}X^\top$ ,  $Z = XP$ ,  $H = \text{diag}\{h_1, \dots, h_n\}$ ,  $h_\ell = -\mu'_\ell \mu''_\ell / V_\ell - \mu''_\ell \mu''_\ell V_\ell^{(1)} / V_\ell^2 + \mu''_\ell V_\ell^{(1)2} / V_\ell^3$ ,  $\mu'_\ell = d\mu_\ell / d\eta_\ell$ ,  $\mu''_\ell = d^2\mu_\ell / d\eta_\ell^2$ ,  $\mu'''_\ell = d^3\mu_\ell / d\eta_\ell^3$ ,  $V_\ell^{(1)} = dV_\ell / d\mu_\ell$ ,  $F = \text{diag}\{f_1, \dots, f_n\}$ ,  $f_\ell = V_\ell^{-1} \mu'_\ell \mu''_\ell$ ,  $G = \text{diag}\{g_1, \dots, g_n\}$ ,  $g_\ell = V_\ell^{-1} \mu'_\ell \mu''_\ell - V_\ell^{-2} V_\ell^{(1)} \mu''_\ell^3$  e  $Z^{(2)} = Z \odot Z$ , onde  $\odot$  denota o produto de Hadamard (Rao, 1973, p. 30) entre matrizes. Além disso, o sub-índice  $d$  indica que uma matriz diagonal foi obtida da matriz original. A matriz  $\Delta$  é definida como

$$\Delta = \sum_{\ell=1}^n \Delta_\ell c_\ell,$$

em que  $\Delta_\ell = (f_\ell + g_\ell) x_\ell x_\ell^\top$  e  $c_\ell = \delta_\ell^\top Z_\beta Z_{\beta_d} F \mathbf{1}$ , sendo  $x_\ell^\top = (x_{\ell 1}, \dots, x_{\ell p})$  a  $\ell$ -ésima linha da matriz de covariadas  $X$ ,  $Z_\beta = X(X^\top WX)^{-1}X^\top$ ,  $\delta_\ell$  um vetor de dimensão  $(n \times 1)$  com 1 na  $\ell$ -ésima posição e zero nas demais e  $\mathbf{1}$  é um vetor de dimensão  $(n \times 1)$  de uns.

A segunda e terceira parcelas da expressão (2.2) vem do cálculo de  $E[d^s(\hat{\beta})(\hat{\beta}_r - \beta_r)]$ , sendo dadas por

$$\begin{aligned} E[d^s(\hat{\beta})(\hat{\beta}_r - \beta_r)] &= E \left\{ (\hat{\beta}_r - \beta_r) \left[ d^s(\beta) + \sum_v d_v^s (\hat{\beta} - \beta)_v + O_p(n^{-2}) \right] \right\} \\ &= d^s(\beta) E[(\hat{\beta}_r - \beta_r)] + o(n^{-2}), \end{aligned}$$

podendo ser expressas na seguinte forma matricial

$$\frac{1}{\phi^2} P(F + G)ZZ_dFP^\top + \frac{1}{2}\phi PDZ_dP^\top,$$

em que  $D = \text{diag}\{V^{-2}V^{(1)}\mu''^2\mu'' - V^{-1}\mu'\mu''' - 5V^{-1}\mu''^2\}$ .

Finalmente, a última parcela da expressão (2.2) vem da expressão  $E[d^r(\hat{\beta})d^s(\hat{\beta})]$  calculada como

$$\begin{aligned} E[d^r(\hat{\beta})d^s(\hat{\beta})] &= E \left\{ \left[ d^r(\beta) + \sum_v d_v^r(\hat{\beta}_v - \beta_v) + O_p(n^{-2}) \right] \left[ d^s(\beta) + \sum_w d_w^s(\hat{\beta}_w - \beta_w) + O_p(n^{-2}) \right] \right\} \\ &= d^r(\beta)d^s(\beta) + o(n^{-2}). \end{aligned}$$

O viés  $d(\beta)$  em notação matricial foi obtido por Cordeiro & McCullagh (1991) sendo dado por

$$d(\beta) = \phi^{-1}(X^\top W X)^{-1} X^\top Z_d F \mathbf{1}.$$

Assim, a matriz de covariância assintótica de ordem  $n^{-2}$  do EMV de  $\beta$  corrigido pelo viés de ordem  $n^{-1}$  é dada por:

$$\begin{aligned} \text{Cov}(\tilde{\beta}) &= \phi^{-1}(X^\top W X)^{-1} + \phi^{-2} P \Lambda P^\top + \phi^{-2}(X^\top W X)^{-1} \Delta (X^\top W X)^{-1} \\ &\quad - (2\phi)^{-2} P Z_d F \mathbf{1} \mathbf{1}^\top F Z_d P^\top \\ &\quad - \phi^{-2} P D Z_d P^\top + 2\phi^{-2} P (F + G) Z Z_d F P^\top. \end{aligned} \tag{2.3}$$

## 2.4 Testes de Wald modificados

Suponhamos que queremos testar  $H_0 : \beta = \beta^{(0)}$  versus  $H_1 : \beta \neq \beta^{(0)}$ , em que o vetor  $\beta$  tem dimensão  $p$ , ou seja, estamos testando todo o vetor de parâmetros. Uma estatística bastante simples para testar a hipótese  $H_0$  é a estatística de Wald, que neste caso é dada pela expressão

$$W = (\hat{\beta} - \beta^{(0)})^\top K_\beta(\hat{\beta})(\hat{\beta} - \beta^{(0)}), \tag{2.4}$$

em que  $K(\beta) = \phi(X^\top W X)$  é avaliada na estimativa de máxima verossimilhança de  $\beta$ .

Podemos modificar esta estatística simplesmente substituindo o estimador  $\hat{\beta}$  pelo estimador corrigido pelo viés de ordem  $n^{-1}$ ,  $\tilde{\beta}$ . Desta forma temos

$$W_m = (\tilde{\beta} - \beta^{(0)})^\top K_\beta(\tilde{\beta})(\tilde{\beta} - \beta^{(0)}). \quad (2.5)$$

Uma outra modificação na estatística de Wald resulta em substituir simultaneamente o EMV  $\hat{\beta}$  pelo estimador corrigido  $\tilde{\beta}$  e a matriz de informação de Fisher pela matriz de covariâncias do estimador corrigido de ordem  $n^{-2}$  dada em (2.3), avaliada em  $\tilde{\beta}$ . Assim, obtemos a estatística

$$W_c = (\tilde{\beta} - \beta^{(0)})^\top \left\{ \text{Cov}(\tilde{\beta}) \right\}^{-1} (\tilde{\beta} - \beta^{(0)}). \quad (2.6)$$

Vamos analisar o desempenho de cada uma destas estatísticas por meio de alguns estudos de simulação.

## 2.5 Resultados de simulação

Consideramos um modelo gama com ligação log, ou seja,  $\log(\mu_\ell) = \beta_0 + \beta_1 x_{1\ell} + \beta_2 x_{2\ell}$ ,  $\ell = 1, \dots, n$ . Os valores verdadeiros para os parâmetros foram fixados em  $\beta_0 = 1$ ,  $\beta_1 = 1$ ,  $\beta_2 = -1$  e  $\phi = 2$ . As covariáveis  $x_1$  e  $x_2$  foram obtidas da distribuição uniforme no intervalo  $(0, 1)$  e, para cada  $n$ , foram mantidas constantes em todas as simulações. Desenvolvemos dois estudos de simulação. No primeiro comparamos a matriz de covariância de ordem  $n^{-2}$  do EMV corrigido pelo viés de ordem  $n^{-1}$  e a inversa da matriz de informação de Fisher com a matriz de covariâncias observadas dos EMVs dos parâmetros  $\beta_0$ ,  $\beta_1$  e  $\beta_2$ , calculada utilizando os valores verdadeiros dos parâmetros. O número de réplicas foi fixado em 10.000 e todas as simulações foram realizadas através do programa computacional R. Os resultados encontram-se na Tabela 2.1, em que  $\text{Cov}(\hat{\beta})$  representa a média dos 10.000 valores obtidos para a matriz de informação de Fisher,  $\text{EQM}(\hat{\beta})$  representa a média dos 10.000

valores obtidos para o erro quadrático médio do EMV em relação aos valores verdadeiros e  $\text{Cov}(\tilde{\beta})$  é a matriz cujas entradas são dadas pela média das 10.000 entradas da matriz de covariâncias do estimador corrigido dada pela expressão (2.3). Na segunda simulação, comparamos as estatísticas (2.4), (2.5) e (2.6) por meio do tamanho empírico do teste de Wald da hipótese  $H_0 : \beta_0^{(0)} = 1, \beta_1^{(0)} = 1$  e  $\beta_2^{(0)} = -1$  contra  $H_1$  : pelo menos uma das igualdades em  $H_0$  não se verifica. Assumindo  $H_0$  verdadeira, o tamanho empírico do teste de Wald em 10.000 replicações é calculado como a proporção do número de vezes em que  $P(\chi_r^2 > w) < \alpha$ ,  $P(\chi_r^2 > w_m) < \alpha$  e  $P(\chi_r^2 > w_c) < \alpha$ , em que  $r$  é o número de parâmetros testados em  $H_0$ ,  $\alpha$  é o nível nominal do teste e  $w$ ,  $w_m$  e  $w_c$  são, respectivamente, os valores das estatísticas  $W$ ,  $W_m$  e  $W_c$  avaliados em cada amostra. Foram utilizados os seguintes níveis nominais:  $\alpha = 1\%$ ,  $5\%$  e  $10\%$ . Os resultados encontram-se na Tabela 2.2. Os tamanhos empíricos encontram-se todos em porcentagens. No primeiro estudo de simulação variamos o tamanho da amostra em  $n = 10, 20, 30$  e  $40$  e no segundo variamos o tamanho da amostra em  $n = 10, 20, \dots$  e  $100$ .

Podemos observar na Tabela 2.1 que os elementos da matriz de covariâncias de ordem  $n^{-2}$  dos EMVs corrigidos pelo viés de ordem  $n^{-1}$  estão bem mais próximos dos valores da matriz de erro quadrático médio em relação aos parâmetros verdadeiros do que os elementos da matriz inversa da informação de Fisher. Além disso, se considerarmos estes valores em valor absoluto, podemos notar que os mesmos são sempre maiores que os respectivos elementos da matriz inversa da informação de Fisher. Isto mostra que, em geral ao se corrigir o viés do estimador e utilizar os elementos da diagonal da matriz inversa da informação de Fisher como estimadores da variância do estimador corrigido, estamos subestimando esta variância. Este fato se reflete claramente nos testes de Wald, como podemos observar na Tabela 2.2, onde vemos que a estatística de teste  $W_c$  obteve um desempenho melhor do que as demais nos três níveis nominais. Estes resultados mostram que as estatísticas Wald e  $W_m$  tendem a rejeitar mais do que deveriam, principalmente quando o tamanho da amostra é pequeno. Como já esperávamos, à medida que o tamanho da amostra aumenta, os tamanhos



empíricos dos testes baseados nas estatísticas  $W$ ,  $W_m$  e  $W_c$  se aproximam do valor nominal, sendo que o da estatística  $W_c$  é sempre mais próximo do nível nominal do que o das outras estatísticas. Além disso, os valores das estatísticas  $W$  e  $W_m$  se aproximam do valor de  $W_c$  à medida que o tamanho da amostra cresce.

Tabela 2.1:  $\text{Cov}(\hat{\beta})$  avaliada em  $\hat{\beta}$ ,  $\text{EQM}(\hat{\beta})$  e  $\text{Cov}(\tilde{\beta})$  avaliada em  $\tilde{\beta}$ 

	$n = 10$			$n = 20$		
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
$\hat{\beta}_0$	0,63166	-0,48998	-0,69830	0,27125	-0,25231	-0,22918
	0,70838	-0,54588	-0,78027	0,29176	-0,26860	-0,24963
	0,70199	-0,55431	-0,78011	0,29294	-0,27414	-0,25052
$\hat{\beta}_1$		0,68931	0,29104		0,35014	0,14293
		0,76220	0,32982		0,36884	0,15764
		0,77835	0,33054		0,37606	0,15943
$\hat{\beta}_2$			1,15772			0,30545
			1,29040			0,33013
			1,28545			0,33062
	$n = 30$			$n = 40$		
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
$\hat{\beta}_0$	0,13183	-0,10180	-0,12912	0,09770	-0,07414	-0,09273
	0,14195	-0,10864	-0,13982	0,10308	-0,07684	-0,09812
	0,13763	-0,10758	-0,13541	0,10055	-0,07659	-0,09610
$\hat{\beta}_1$		0,23933	-0,01453		0,13910	0,00690
		0,24785	-0,00902		0,14365	0,00764
		0,25167	-0,01437		0,14361	0,00729
$\hat{\beta}_2$			0,25562			0,17390
			0,27003			0,18252
			0,26736			0,18006

Tabela 2.2: Tamanho do teste para as estatísticas  $W$ ,  $W_m$  e  $W_c$ .

$n$	$\alpha(\%)$	$W$	$W_m$	$W_c$	$n$	$\alpha(\%)$	$W$	$W_m$	$W_c$
10	1,0	3,09	2,39	1,99	60	1,0	1,19	1,11	1,05
	5,0	8,82	7,63	6,42		5,0	5,70	5,41	5,20
	10,0	14,71	13,05	11,40		10,0	10,96	10,73	10,35
20	1,0	1,87	1,54	1,21	70	1,0	1,22	1,12	1,08
	5,0	7,03	6,19	5,55		5,0	5,66	5,37	5,19
	10,0	12,45	11,96	11,03		10,0	11,15	10,98	10,69
30	1,0	1,72	1,51	1,25	80	1,0	1,17	1,09	1,01
	5,0	6,50	5,99	5,43		5,0	5,54	5,41	5,22
	10,0	11,93	11,33	10,50		10,0	11,02	10,73	10,46
40	1,0	1,45	1,29	1,21	90	1,0	1,29	1,19	1,15
	5,0	6,38	5,98	5,65		5,0	5,58	5,46	5,35
	10,0	11,51	11,13	10,63		10,0	10,78	10,54	10,30
50	1,0	1,42	1,19	1,07	100	1,0	1,24	1,11	1,09
	5,0	5,77	5,56	5,31		5,0	5,42	5,39	5,22
	10,0	11,04	10,99	10,38		10,0	10,16	10,15	9,97

## 2.6 Conclusões

Através do estudo de simulação pudemos observar que ao estimarmos as covariâncias do EMV pelos elementos da matriz inversa da informação de Fisher, estamos subestimando estes valores, ao passo que, se utilizarmos os elementos da matriz de covariâncias do estimador corrigido, as estimativas são bem mais precisas. O efeito desta precisão foi verificado no melhoramento da estatística do teste de Wald para o parâmetro  $\beta$  em MLGs com dispersão conhecida, simplesmente substituindo a estimativa da matriz de informação de Fisher na estatística de teste usual pela estimativa da inversa da matriz de covariância assintótica de ordem  $n^{-2}$  do estimador corrigido.

## Capítulo 3

# Coeficiente de assimetria assintótico da distribuição do estimador de máxima verossimilhança em modelos não-lineares da família exponencial

### 3.1 Introdução

Em muitas distribuições de probabilidades, a propriedade de simetria representa uma suposição muito importante. Uma das medidas de assimetria mais utilizadas é o terceiro cumulante de Pearson padronizado, definido por  $\gamma_1 = \kappa_3/\kappa_2^{3/2}$ , em que  $\kappa_r$  é o  $r$ -ésimo cumulante da distribuição. Se a distribuição é simétrica,  $\gamma_1$  é igual a zero e, portanto, seu valor dará alguma indicação do grau de afastamento da simetria.

Quando  $\gamma_1 > 0$  ( $\gamma_1 < 0$ ) a distribuição é positivamente (negativamente) assimétrica e terá uma cauda longa (curta) à direita e curta (longa) à esquerda. O valor do índice  $\gamma_1$  também pode ser usado como uma possível medida de não-normalidade da distribuição, já que, para a distribuição normal,  $\gamma_1 = 0$ .

Nosso objetivo neste capítulo é calcular o coeficiente de assimetria assintótico de ordem  $n^{-1/2}$

da distribuição do estimador de máxima verossimilhança (EMV) em modelos não lineares da família exponencial (MNLFEs), estendendo os resultados de Cordeiro & Cordeiro (2001). Estimativas deste coeficiente podem ser usadas como uma medida de afastamento da distribuição assintótica, uma vez que a distribuição do EMV é assintoticamente normal. Assim, espera-se que seu coeficiente de assimetria se aproxime de zero à medida em que se aumenta o tamanho da amostra.

### 3.2 Modelos não-lineares da família exponencial

Os MNLFEs foram definidos inicialmente por Cordeiro & Paula (1989). Estes modelos se baseiam nos modelos exponenciais de dispersão, discutidos com bastante detalhe por Jørgensen (1987) e são uma extensão dos MLGs (Nelder & Wedderburn, 1972) e do modelo normal não-linear discutidos por Ratkowsky (1983) e Seber & Wild (1989).

Suponha uma amostra de  $n$  variáveis  $Y_1, \dots, Y_n$  independentes com cada  $Y_\ell$  tendo função densidade de probabilidade ou função de probabilidade na forma

$$\pi(y; \theta_\ell, \phi) = \exp[\phi\{y\theta_\ell - b(\theta_\ell)\} + c(y, \phi)], \quad (3.1)$$

em que  $b(\cdot)$  e  $c(\cdot)$  são funções conhecidas e  $\phi > 0$  é o parâmetro de precisão. O parâmetro  $\sigma^2 = \phi^{-1}$  é chamado parâmetro de dispersão. Além disso,  $E(Y_\ell) = \mu_\ell = b'(\theta_\ell) = db(\theta_\ell)/d\theta_\ell$ ,  $Var(Y_\ell) = \phi^{-1}V_\ell$  sendo que  $V_\ell = d\mu_\ell/d\theta_\ell$  é chamada de função de variância. Suponha também que o parâmetro de dispersão  $\phi$  seja desconhecido, porém o mesmo para todas as observações.

Os MNLFE's são definidos por (3.1) e pela componente sistemática

$$t(\mu_\ell) = \eta_\ell = f(x_\ell; \beta), \quad (3.2)$$

em que  $t(\cdot)$  é uma função conhecida monótona e diferenciável, chamada função de ligação,  $\beta =$

$(\beta_1, \dots, \beta_p)^T$ ,  $p < n$ , é o conjunto de parâmetros desconhecidos a serem estimados,  $f(\cdot; \cdot)$  é uma função contínua e diferenciável e  $x_\ell = (x_{\ell 1}, \dots, x_{\ell q})^T$  é um vetor de valores conhecidos associado à resposta observada  $y_\ell$ . Se considerarmos os MLGs, temos  $f(x_\ell; \beta) = x_\ell^\top \beta$ , ( $q = p$ ). Portanto, esta classe de modelos estende a classe dos MLGs com dispersão constante.

A classe de distribuições (3.1) inclui muitas distribuições contínuas importantes tais como a normal, gama, normal inversa, exponencial, Hermite, Neyman tipo A, secante hiperbólica generalizada, log zeta e a distribuição Tweedie com função de variância potência  $V(\mu_\ell) = \mu_\ell^\delta$  para  $\delta \leq 0$  e  $\delta \geq 2$ , e distribuições discretas como a binomial, Poisson e a binomial negativa.

Vamos supor identificabilidade no sentido de que diferentes  $\beta$ 's impliquem em diferentes  $\eta$ 's, em que  $\eta = (\eta_1, \dots, \eta_m)^\top$ . Esta suposição fará com que a matriz de derivadas  $X^* = X^*(\beta) = \partial\eta/\partial\beta$  tenha posto  $p$ , para todo  $\beta$ .

Assumimos também, as suposições usuais de regularidade para a função de verossimilhança (Cox & Hinkley, 1974) obtida a partir de (3.1) e (3.2). Lehmann & Casella (1998) mostraram que, sob estas condições, o EMV  $\hat{\beta}$  do vetor de parâmetros  $\beta$  tem boas propriedades assintóticas como consistência, suficiência e normalidade.

Denotamos o logaritmo da função de verossimilhança total por  $\mathcal{L}(\beta)$  e por  $\hat{\beta}$ ,  $\hat{\eta}$  e  $\hat{\mu}$  as estimativas de máxima verossimilhança de  $\beta$ ,  $\eta$  e  $\mu$ , respectivamente, em que  $\mu = (\mu_1, \dots, \mu_n)^T$ .

A função escore para  $\beta$  é dada por  $U(\beta) = \phi X^{*\top} W H (y - \mu)$ , sendo  $W = \text{diag}\{d\mu_\ell/d\eta_\ell\}^2/V_\ell$ ,  $H = \text{diag}\{d\eta_\ell/d\mu_\ell\}$  e  $y = (y_1, \dots, y_n)^T$ .

A matriz de informação de Fisher para  $\beta$  é dada por  $E\{U(\beta)U(\beta)^\top\} = \phi K_\beta = \phi X^{*\top} W X^*$ . O parâmetro de dispersão  $\phi$  é ortogonal a  $\beta$  no sentido de Cox & Reid (1987), ou seja,  $E(-\partial^2 \mathcal{L}/\partial\beta\partial\phi) = 0$ . A matriz de covariâncias de ordem  $n^{-1}$  de  $\hat{\beta}$  é  $\phi^{-1} K_\beta^{-1} = \phi^{-1} (X^{*\top} W X^*)^{-1}$ .

A função  $c(y, \phi)$  em (3.1) é fundamental para a estimação de  $\phi$ , uma vez que o EMV de  $\beta$

não depende desta função. Quando (3.1) é uma distribuição da família exponencial bi-paramétrica completa com parâmetros canônicos  $\phi$  e  $\phi\theta$ , a decomposição

$$c(y, \phi) = \phi a_1(y) + m(\phi) + a_2(y), \quad (3.3)$$

é válida e o estimador de  $\phi$  é facilmente obtido por máxima verossimilhança. No entanto, para alguns modelos exponenciais de dispersão, a estimação deste parâmetro se torna bastante complicada. A expressão (3.3) é válida para os modelos normal, gama e normal inversa mas nem todos os modelos da família exponencial possuem esta propriedade.

### 3.3 Coeficiente de assimetria assintótico para a distribuição de $\hat{\beta}$

Adotamos a seguinte notação para as derivadas do logaritmo da função de verossimilhança total, sendo que todos os índices variam pelos inteiros  $1, \dots, p$ :

$$\kappa_{rs} = E(\partial^2 \mathcal{L} / \partial \beta_r \partial \beta_s), \kappa_{r,s} = E(\partial \mathcal{L} / \partial \beta_r \partial \mathcal{L} / \partial \beta_s), \kappa_{rst} = E(\partial^3 \mathcal{L} / \partial \beta_r \partial \beta_s \partial \beta_t),$$

$$\kappa_{r,s,t} = E(\partial \mathcal{L} / \partial \beta_r \partial \mathcal{L} / \partial \beta_s \partial \mathcal{L} / \partial \beta_t), \kappa_{r,st} = E(\partial \mathcal{L} / \partial \beta_r \partial^2 \mathcal{L} / \partial \beta_s \partial \beta_t),$$

$\kappa_{rs}^{(t)} = \partial \kappa_{rs} / \partial \beta_t$ , e assim por diante. Os  $\kappa$ 's correspondem aos cumulantes de  $\mathcal{L}(\beta)$  e todos são de ordem  $n$ . A matriz de informação total de Fisher tem elementos  $\kappa_{r,s} = -\kappa_{rs}$  e considere  $\kappa^{r,s} = -\kappa^{rs}$  seu elemento correspondente na matriz inversa. Além disso, definimos

$$(r)_\ell = \frac{\partial \eta_\ell}{\partial \beta_r}, (rs)_\ell = \frac{\partial^2 \eta_\ell}{\partial \beta_r \partial \beta_s}, (r,s)_\ell = \frac{\partial \eta_\ell}{\partial \beta_r} \frac{\partial \eta_\ell}{\partial \beta_s},$$

e assim por diante.

Seja  $\kappa_3(\hat{\beta}_a) = E\{(\hat{\beta}_a - \beta_a)^3\}$  o terceiro cumulante do EMV  $\hat{\beta}_a$  de  $\beta_a$ , para  $a = 1, \dots, p$ .

Pela expressão geral para o terceiro cumulante do EMV desenvolvida por Bowman & Shenton (1998), podemos escrever até ordem  $n^{-2}$  (ou seja, desprezando termos de ordem menor que  $n^{-2}$ )

$$\kappa_3(\hat{\beta}_a) = \sum_{r,s,t \in \Gamma} \kappa^{a,r} \kappa^{a,s} \kappa^{a,t} (\kappa_{r,s,t} + 3\kappa_{rst} + 6\kappa_{rs,t}), \quad (3.4)$$

em que  $\Gamma = \{\beta, \phi\}$  é o conjunto com todos os  $p + 1$  parâmetros  $\beta_1, \dots, \beta_p$  e  $\phi$ . Um fato que facilitará o cálculo de  $\kappa_3(\hat{\beta})$  é a invariância dos  $\kappa$ 's sob permutação dos parâmetros  $\beta$  e a ortogonalidade entre  $\phi$  e  $\beta$ . Alguns dos termos da equação (3.4) podem ser calculados de maneira mais simples através da relação

$$\kappa_{r,s,t} = 2\kappa_{rst} - \kappa_{rs}^{(t)} - \kappa_{st}^{(r)} - \kappa_{rt}^{(s)}. \quad (3.5)$$

Portanto, substituindo a expressão (3.5) em (3.4) obtemos

$$\kappa_3(\hat{\beta}_a) = \sum_{r,s,t \in \Gamma} \kappa^{a,r} \kappa^{a,s} \kappa^{a,t} (5\kappa_{rst} - \kappa_{st}^{(r)} - \kappa_{rt}^{(s)} + 6\kappa_{rs,t}). \quad (3.6)$$

Para o modelo definido por (3.1) e (3.2) obtemos os cumulantes de terceira ordem como

$$\kappa_{rst} = -\phi \sum_{\ell=1}^n [(f + 2g)(r, s, t) + w\{(rs, t) + (st, r) + (rt, s)\}]_{\ell}, \quad (3.7)$$

$$\kappa_{rs,t} = \phi \sum_{\ell=1}^n [g(r, s, t) + w(rs, t)]_{\ell} \quad (3.8)$$

e

$$\kappa_{rs}^{(t)} = -\phi \sum_{\ell=1}^n [(f + g)(r, s, t) + w\{(st, r) + (rt, s)\}]_{\ell}, \quad (3.9)$$

em que

$$f_\ell = \left( \frac{d\theta_\ell}{d\eta_\ell} \right) \frac{d^2\mu_\ell}{d\eta_\ell^2} = V_\ell^{-1} \left( \frac{d\mu_\ell}{d\eta_\ell} \right) \frac{d^2\mu_\ell}{d\eta_\ell^2} \quad (3.10)$$

e

$$g_\ell = \left( \frac{d\mu_\ell}{d\eta_\ell} \right) \frac{d^2\theta_\ell}{d\eta_\ell^2} = V_\ell^{-1} \left( \frac{d\mu_\ell}{d\eta_\ell} \right) \frac{d^2\mu_\ell}{d\eta_\ell^2} - V_\ell^{-2} \left( \frac{d\mu_\ell}{d\eta_\ell} \right)^3 \frac{dV_\ell}{d\mu_\ell}. \quad (3.11)$$

Assim, substituindo as expressões (3.7), (3.8) e (3.9) em (3.6) obtemos

$$\begin{aligned} \kappa_3(\widehat{\beta}_a) &= -\phi \left\{ \sum_{\ell=1}^n (2f + g)_\ell \left[ \sum_{r,s,t \in \Gamma} \kappa^{a,r} \kappa^{a,s} \kappa^{a,t}(r, s, t) \right]_\ell - 3 \sum_{i=1}^n w_\ell \sum_{r,s,t \in \Gamma} \kappa^{a,r} \kappa^{a,s} \kappa^{a,t} \{ (rs, t) \right. \\ &\quad \left. - (rt, s) - (st, r) \} \right\}. \end{aligned}$$

Como

$$\sum_{r,s,t \in \Gamma} \kappa^{a,r} \kappa^{a,s} \kappa^{a,t}(r, s, t) = \left[ \sum_{r=1}^p \kappa^{a,r}(r) \right]^3$$

e

$$\sum_{r,s,t \in \Gamma} \kappa^{a,r} \kappa^{a,s} \kappa^{a,t} \{ (rs, t) - (rt, s) - (st, r) \} = \sum_{t=1}^p \kappa^{a,t}(t) \sum_{r=1}^p \sum_{s=1}^p \kappa^{a,r} \kappa^{a,s}(rs),$$

temos que

$$\kappa_3(\widehat{\beta}_a) = -\phi \sum_{\ell=1}^n \left\{ (2f + g) \left[ \sum_{r=1}^p \kappa^{a,r}(r) \right]^3 + 3 \sum_{i=1}^n w \sum_{t=1}^p \kappa^{a,t}(t) \sum_{r=1}^p \sum_{s=1}^p \kappa^{a,r} \kappa^{a,s}(rs) \right\}_\ell. \quad (3.12)$$

Claramente, a expressão (3.12) mostra que o terceiro cumulante  $\kappa_3(\widehat{\beta}_a)$  do estimador  $\widehat{\beta}_a$  é de ordem



$O(p^3)$  e, portanto, cresce quando o número  $p$  de variáveis explicativas do modelo cresce. Definimos agora as seguintes matrizes de dimensão  $p \times p$ ,  $\tilde{X}_\ell = (\partial^2 \eta_\ell / \partial \beta_r \partial \beta_s)$ ,  $S_\ell = \text{diag}\{K_\beta^{-1} \tilde{X}_\ell K_\beta^{-1}\}$  e a matriz  $N$  de dimensão  $p \times n$  dada por  $N = [s_1 \ s_2 \ \dots \ s_n]$ , em que  $s_\ell = S_\ell \mathbf{1}$ ,  $\mathbf{1}$  é um vetor  $p \times 1$  de uns,  $\ell = 1, \dots, n$ . Além disso, sejam as matrizes  $M = (X^{*\top} W X^*)^{-1} X^{*T}$  e  $T = \text{diag}\{N W M^\top\}$  de dimensões  $p \times n$  e  $p \times p$  respectivamente.

Da expressão (3.12) podemos escrever, até ordem  $n^{-2}$ , o vetor de terceiros cumulantes em notação matricial da seguinte forma:

$$\kappa_3(\hat{\beta}) = -\frac{1}{\phi^2} \left\{ M^{(3)}(2f + g) + 3T\mathbf{1} \right\}, \quad (3.13)$$

em que  $f = (f_1, \dots, f_n)^T$  e  $g = (g_1, \dots, g_n)^T$  são vetores  $n \times 1$ , com  $f_\ell$  e  $g_\ell$  definidos anteriormente pelas expressões (3.10) e (3.11), respectivamente e  $M^{(3)} = M \odot M \odot M$ , em que  $\odot$  denota o produto de Hadamard (Rao, 1973, p. 30) entre matrizes. O vetor  $\kappa_3(\hat{\beta})$  é ponderado pelo inverso do quadrado do parâmetro de precisão  $\phi$  e depende da matriz  $X^*$  de primeiras derivadas associada ao modelo, das duas primeiras derivadas da função de ligação, mas apenas da primeira derivada da função de variância. Claramente a aproximação normal da distribuição de  $\hat{\beta}$  se deteriora quando  $\phi$  decresce. A expressão (3.13) é válida apenas para o EMV de  $\hat{\beta}$  e não se aplica ao estimador corrigido pelo viés de primeira ordem obtido por Paula (1992).

A expressão (3.13) generaliza o resultado obtido por Cordeiro & Cordeiro (2001) para modelos lineares generalizados. Esta expressão é de fácil implementação em programas que permitem executar operações matriciais, como `Ox`, `MAPLE`, `MATHEMATICA`, `R`, etc. Uma forma de interpretar esta expressão é considerar a primeira parcela como a contribuição da parte linear e a segunda parcela como a contribuição da parte não-linear, uma vez que esta segunda parcela é o que se acrescenta à expressão obtida por Cordeiro & Cordeiro (2001) para o caso linear. Estimativas consistentes de  $\mu$  e  $\phi$ , digamos  $\hat{\mu}$  e  $\hat{\phi}$ ,

podem ser inseridas na expressão (3.13) a fim de obtermos estimativas do coeficiente de assimetria da distribuição de  $\hat{\beta}$ . Utilizando a matriz de covariância assintótica  $\phi^{-1}(X^{*\top}WX^*)^{-1}$  de  $\hat{\beta}$  e o vetor  $\kappa_3(\hat{\beta})$ , obtemos o coeficiente de assimetria do EMV estimado,  $\hat{\gamma}_1(\hat{\beta}_a) = \kappa_3(\hat{\beta}_a)/\text{Var}(\hat{\beta}_a)^{3/2}$ , de ordem  $n^{-1/2}$ , avaliado em  $\hat{\beta}$ . Sabendo que a distribuição do EMV é assintoticamente normal, a magnitude de  $\hat{\gamma}_1(\hat{\beta}_a)$  pode ser usada como um indicador da qualidade da aproximação pela distribuição normal, embora o fato de esta estimativa ir para zero não necessariamente indique que a distribuição seja normal. Ou seja, este critério deve ser usado com uma certa cautela. Quando o valor de  $\hat{\gamma}_1(\hat{\beta}_a)$  é pequeno não haverá, neste contexto, grandes prejuízos em assumir a aproximação normal para a distribuição do EMV. Porém, quando este valor é grande, deve-se olhar com cuidado esta aproximação. Por exemplo, podemos considerar a aproximação normal razoável se  $|\hat{\gamma}_1(\hat{\beta}_a)| < 1/10$ .

Calculando-se  $\kappa_3(\hat{\beta}_a)$  através da expressão (3.13), podemos obter uma expansão de Edgeworth para a função densidade de probabilidade do estimador  $\hat{\beta}_a$  dada por

$$f_{\hat{\beta}_a}(x) = \phi(x) \left\{ 1 + \frac{\kappa_3(\hat{\beta}_a)}{6} H_3(x) + \frac{\kappa_3(\hat{\beta}_a)^2}{72} H_6(x) \right\},$$

em que  $\phi(x)$  é a função densidade de probabilidade da distribuição normal padrão,  $H_3(x) = x^3 - 3x$  e  $H_6(x) = x^6 - 15x^4 + 45x^2 - 15$  são polinômios de Hermite.

### 3.4 Coeficiente de assimetria assintótico das distribuições dos estimadores de máxima verossimilhança $\hat{\phi}$ e $\hat{\sigma}^2$

Os coeficientes de assimetria assintóticos das distribuições dos EMVs dos parâmetros de precisão  $\phi$  e dispersão  $\sigma^2$  podem ser bem complicados, para o caso em que a função  $c(y, \phi)$  é geral. Porém, quando a decomposição (3.3) é verificada, estes coeficientes coincidem com os do caso linear obtidos por Cordeiro & Cordeiro (2001), pois as derivadas do logaritmo da função de verossimilhança total em relação a  $\phi$  e  $\sigma^2$  não dependem do parâmetro  $\beta$ . Assim, para cada caso, temos

$$\gamma_1(\hat{\phi}) = \frac{2m^{(3)}(\phi)}{\sqrt{n}\{-m^{(2)}(\phi)\}^{3/2}}$$

e

$$\gamma_1(\hat{\sigma}^2) = 2\sqrt{\frac{\sigma^2}{n}} \frac{\{\sigma^2 e^{(3)}(\sigma^2) + 3e^{(2)}(\sigma^2)\}}{\{-2e^{(1)}(\sigma^2) - \sigma^2 e^{(2)}(\sigma^2)\}^{3/2}},$$

sendo  $e(\sigma^2) = m(\sigma^{-2})$ ,  $m^{(r)}(\phi) = d^r m(\phi)/d\phi^r$  e  $e^{(r)}(\sigma^2) = d^r e(\sigma^2)/d\sigma^{2r}$ .

### 3.5 Dois exemplos

Em primeiro lugar consideramos o modelo não-linear uniparamétrico da família exponencial, para o qual temos  $\eta'_\ell = rmd\eta_\ell/d\beta$ ,  $\eta''_\ell = d^2\eta_\ell/d\beta^2$ ,  $\tilde{X}_\ell = \eta''_\ell$  e  $s_\ell = \phi^{-2}\eta''_\ell(\sum_{\ell=1}^n w_\ell\eta_\ell'^2)^{-2}$  para  $\ell = 1, \dots, n$ ,  $K_\beta^{-1} = (\sum_{\ell=1}^n w_\ell\eta_\ell'^2)^{-1}$ ,  $T = (\sum_{\ell=1}^n w_\ell\eta_\ell'^2)^{-3} \sum_{\ell=1}^n \eta'_\ell\eta_\ell''w_\ell$ , e

$$M^{(3)}(2f + g) = \left( \sum_{\ell=1}^n w_\ell\eta_\ell'^2 \right)^{-3} \sum_{\ell=1}^n \eta_\ell'^3(2f_\ell + g_\ell).$$

Assim, substituindo estas quantidades na expressão (3.13), obtemos

$$\kappa_3(\hat{\beta}) = -\frac{1}{\phi^2} \left[ \sum_{i=1}^n \omega_i\eta_i'^2 \right]^{-3} \left\{ \sum_{i=1}^n \left[ \eta_i'^3(2f_i + g_i) + 3\eta_i'\eta_i''\omega_i \right] \right\}.$$

Em segundo lugar consideramos o modelo com dois parâmetros na componente sistemática (3.2), digamos  $\beta = (\delta, \gamma)$ , e alguma distribuição na família (3.1).

Podemos expressar o coeficiente de assimetria assintótico para os EMV's de  $\delta$  e  $\gamma$ , respectivamente por

$$\begin{aligned} \kappa_3(\widehat{\delta}) &= -\phi \left\{ \sum_{i=1}^n \left[ \left( \text{Var}(\widehat{\delta}) \frac{\partial \eta_\ell}{\partial \delta} + \text{Cov}(\widehat{\delta}, \widehat{\gamma}) \frac{\partial \eta_\ell}{\partial \gamma} \right)^3 (2f_\ell + g_\ell) \right] + 3\text{Var}(\widehat{\delta}) \sum_{i=1}^n \frac{\partial \eta_\ell}{\partial \delta} D_{11}^i \omega_\ell \right. \\ &\quad \left. + \text{Cov}(\widehat{\delta}, \widehat{\gamma}) \sum_{i=1}^n \frac{\partial \eta_\ell}{\partial \gamma} D_{11}^i \omega_\ell \right\} \end{aligned}$$

e

$$\begin{aligned} \kappa_3(\widehat{\gamma}) &= -\phi \left\{ \sum_{i=1}^n \left[ \left( \text{Var}(\widehat{\gamma}) \frac{\partial \eta_\ell}{\partial \gamma} + \text{Cov}(\widehat{\delta}, \widehat{\gamma}) \frac{\partial \eta_\ell}{\partial \delta} \right)^3 (2f_\ell + g_\ell) \right] + 3\text{Var}(\widehat{\gamma}) \sum_{i=1}^n \frac{\partial \eta_\ell}{\partial \gamma} D_{22}^i \omega_\ell \right. \\ &\quad \left. + \text{Cov}(\widehat{\delta}, \widehat{\gamma}) \sum_{i=1}^n \frac{\partial \eta_\ell}{\partial \delta} D_{22}^i \omega_\ell \right\}, \end{aligned}$$

em que

$$D_{11}^i = (\text{Var}(\widehat{\delta}))^2 \frac{\partial^2 \eta_\ell}{\partial \delta^2} + 2\text{Cov}(\widehat{\delta}, \widehat{\gamma}) \frac{\partial^2 \eta_\ell}{\partial \delta \partial \gamma} \text{Var}(\widehat{\delta}) + \frac{\partial^2 \eta_\ell}{\partial \gamma^2} [\text{Cov}(\widehat{\delta}, \widehat{\gamma})]^2$$

e

$$D_{22}^i = (\text{Var}(\widehat{\gamma}))^2 \frac{\partial^2 \eta_\ell}{\partial \gamma^2} + 2\text{Cov}(\widehat{\delta}, \widehat{\gamma}) \frac{\partial^2 \eta_\ell}{\partial \delta \partial \gamma} \text{Var}(\widehat{\gamma}) + \frac{\partial^2 \eta_\ell}{\partial \delta^2} [\text{Cov}(\widehat{\delta}, \widehat{\gamma})]^2.$$

### 3.6 Resultados de simulações

Nosso objetivo aqui é comparar o coeficiente de assimetria analítico até ordem  $n^{-1/2}$  da distribuição do EMV em MNLFEs, dado por  $\gamma_1 = \kappa_3(\widehat{\beta})/\kappa_2(\widehat{\beta})^3/2$ , em que  $\kappa_3(\widehat{\beta})$  é dado pela expressão

(3.13) e  $\kappa_2(\hat{\beta})$  é obtido da inversa da matriz de informação de Fisher, com o coeficiente de assimetria amostral, definido pela estatística da razão de momentos, que definimos logo em seguida. Para isso, consideramos dois estudos de simulação de Monte Carlo variando a distribuição da resposta, a distribuição da covariável e o número de observações. No primeiro estudo consideramos as distribuições gama e exponencial para a variável resposta. A estrutura não-linear para todos os modelos foi definida por  $\log(\mu_\ell) = \alpha + \exp(\beta x_\ell)$ ,  $\ell = 1, \dots, n$ . Os valores verdadeiros dos parâmetros foram fixados em  $\alpha = 3$ ,  $\beta = 2$  e  $\phi = 1$  e  $2$  correspondentes aos modelos exponencial e gama, respectivamente. A covariável  $x$  foi gerada da distribuição uniforme  $U(0, 1)$ .

No segundo estudo, consideramos a distribuição normal para a resposta e as distribuições uniforme  $U(0, 1)$  e qui-quadrado com 2 graus de liberdade para as covariáveis. Os valores verdadeiros dos parâmetros foram fixados em  $\alpha = 3$ ,  $\beta = -1$  e  $\phi = 2$ . Em ambos estudos de simulação, o tamanho da amostra,  $n$ , variou de  $n = 10, \dots, 100$ . Para cada valor de  $n$ , os valores das covariáveis foram mantidos constantes em todas as replicações do experimento.

Para cada caso, foram gerados 10.000 vetores ( $y$ ) de observações e em cada replicação ajustamos o modelo a fim de calcular as estimativas  $\hat{\alpha}$ ,  $\hat{\beta}$  e os seus respectivos valores ajustados  $\hat{\mu}_1, \dots, \hat{\mu}_n$ . Assim, as estatísticas da razão de momentos amostrais das estimativas  $\hat{\alpha}$  e  $\hat{\beta}$ , foram calculadas como

$$g_1(\hat{\alpha}) = \frac{m_3(\hat{\alpha})}{m_2(\hat{\alpha})^{3/2}}$$

e

$$g_1(\hat{\beta}) = \frac{m_3(\hat{\beta})}{m_2(\hat{\beta})^{3/2}},$$

em que

$$m_r(\hat{\alpha}) = \frac{\sum_{i=1}^{10000} (\hat{\alpha}_i - \bar{\hat{\alpha}})^r}{10000}, \quad m_r(\hat{\beta}) = \frac{\sum_{i=1}^{10000} (\hat{\beta}_i - \bar{\hat{\beta}})^r}{10000}, \quad \bar{\hat{\alpha}} = \frac{\sum_{i=1}^{10000} \hat{\alpha}_i}{10000} \quad \text{e} \quad \bar{\hat{\beta}} = \frac{\sum_{i=1}^{10000} \hat{\beta}_i}{10000}.$$

Então,  $g_1(\hat{\alpha})$  e  $g_1(\hat{\beta})$  são os valores amostrais dos coeficientes de assimetria da distribuição de  $\hat{\alpha}$  e  $\hat{\beta}$  baseados nas estimativas  $\hat{\alpha}_1, \dots, \hat{\alpha}_{10000}$  e  $\hat{\beta}_1, \dots, \hat{\beta}_{10000}$  obtidas em todos os 10.000 experimentos. Todos os cálculos foram desenvolvidos no programa computacional R.

As Tabelas 3.1 e 3.2 apresentam as medidas dos coeficientes de assimetria amostrais  $g_1(\hat{\alpha})$  e  $g_1(\hat{\beta})$  com seus respectivos coeficientes de assimetria analíticos  $\gamma_1(\hat{\alpha})$  e  $\gamma_1(\hat{\beta})$  obtidos de (3.13). Alguns comentários podem ser feitos aqui. Primeiro, os coeficientes de assimetria analíticos das distribuições dos estimadores  $\hat{\alpha}$  e  $\hat{\beta}$  decrescem numericamente à medida que o tamanho da amostra aumenta, o que já era esperado, de acordo com a teoria assintótica de segunda ordem. Segundo, a diferença entre os valores dos coeficientes de assimetria analítico e amostral é mais acentuada quando o tamanho da amostra é pequeno. Terceiro, está claro que os coeficientes de assimetria analíticos e amostrais da distribuição dos EMVs são afetados pela distribuição da resposta e, para  $n$  fixado, eles são quase sempre maiores no modelo exponencial ( $\gamma_1 = 2$ ) que no modelo gama ( $\gamma_1 = \sqrt{2}$ ). Quarto, observamos também que, para  $n$  fixado, a distribuição do estimador do parâmetro não-linear  $\beta$  é mais assimétrica do que a distribuição do estimador do parâmetro linear  $\alpha$ . Finalmente, os valores da Tabela 3.2 indicam que a distribuição das covariáveis não tem uma influência tão significativa na assimetria da distribuição dos EMVs como a influência da distribuição da variável resposta.

## 3.7 Aplicações

### 3.7.1 Elasticidade constante de substituição (CES) da função de produção

Suponha que estamos interessados em estimar a relação entre um *output* agregado em uma certa indústria,  $y$ , e dois *inputs*, trabalho ( $L$ ) e capital ( $K$ ). A função de produção de Cobb-Douglas é

Tabela 3.1: Coeficientes de assimetria amostrais e analíticos para os modelos exponencial e gama

n	Modelo Gama				Modelo Exponencial			
	$g_1(\hat{\alpha})$	$\gamma_1(\hat{\alpha})$	$g_1(\hat{\beta})$	$\gamma_1(\hat{\beta})$	$g_1(\hat{\alpha})$	$\gamma_1(\hat{\alpha})$	$g_1(\hat{\beta})$	$\gamma_1(\hat{\beta})$
10	-0.1391	-0.1190	-1.1154	-0.5099	-0.2041	-0.3485	-0.6842	-0.4638
20	-0.1178	-0.1107	-0.5203	-0.4576	-0.1532	-0.2060	-0.6903	-0.5610
30	-0.1177	-0.1113	-0.2954	-0.2487	-0.1288	-0.1382	-0.6231	-0.5664
40	-0.1342	-0.1107	-0.2069	-0.1787	-0.1724	-0.1555	-0.3460	-0.3442
50	-0.0752	-0.0870	-0.2589	-0.1888	-0.1893	-0.1336	-0.3116	-0.3365
60	-0.0730	-0.0782	-0.2340	-0.2003	-0.1112	-0.1090	-0.3959	-0.3434
70	-0.1017	-0.0691	-0.1667	-0.1813	-0.1005	-0.1133	-0.2927	-0.3089
80	-0.0350	-0.0716	-0.1866	-0.1707	-0.1322	-0.1089	-0.2227	-0.2463
90	-0.0784	-0.0675	-0.1706	-0.1435	-0.1029	-0.0994	-0.2560	-0.2561
100	-0.04740	-0.0582	-0.1480	-0.1630	-0.0823	-0.0965	-0.2662	-0.2427

Tabela 3.2: Coeficientes de assimetria amostrais e analíticos para o modelo normal

n	Covariável Uniforme				Covariável Qui-Quadrado			
	$g_1(\hat{\alpha})$	$\gamma_1(\hat{\alpha})$	$g_1(\hat{\beta})$	$\gamma_1(\hat{\beta})$	$g_1(\hat{\alpha})$	$\gamma_1(\hat{\alpha})$	$g_1(\hat{\beta})$	$\gamma_1(\hat{\beta})$
10	0.0267	0.1629	-0.1441	-1.4901	-0.0706	-0.6995	-0.1157	-0.9005
20	-0.0072	0.1168	-0.0827	-1.0739	-0.0101	-0.5048	-0.1253	-0.6443
30	0.0080	0.0957	-0.0624	-0.8805	0.0236	-0.4159	-0.1327	-0.5286
40	0.0333	0.0829	-0.1084	-0.7633	-0.0158	-0.3609	-0.0632	-0.4576
50	-0.0240	0.0741	-0.0241	-0.6837	0.0036	-0.3235	-0.0848	-0.4096
60	0.0137	0.0677	-0.0856	-0.6249	-0.0190	-0.2958	-0.0584	-0.3744
70	-0.0130	0.0627	-0.0389	-0.5788	-0.0306	-0.2738	-0.0780	-0.3467
80	-0.0094	0.0587	-0.0539	-0.5414	-0.0208	-0.2563	-0.0211	-0.3250
90	-0.0005	0.0554	-0.0354	-0.5112	0.0055	-0.2419	-0.0760	-0.3057
100	-0.0043	0.0526	-0.0632	-0.4849	-0.0034	-0.2292	-0.0501	-0.2895

descrita pela relação

$$y = \alpha L^{\beta_2} K^{\beta_3} \exp\{\epsilon\},$$

em que  $\epsilon$  representa o erro aleatório.

A elasticidade constante de substituição da função de produção é uma função mais geral que a

de Cobb-Douglas. Ela é especificada por

$$y = \alpha[\delta L^{-\rho} + (1 - \delta)K^{-\rho}]^{-\tau/\rho} \exp\{\epsilon\}, \quad (3.14)$$

em que  $\alpha > 0$  é o parâmetro de eficiência,  $\tau > 0$  é o parâmetro de escala,  $\rho > -1$  é o parâmetro de substituição,  $0 < \delta < 1$  é o parâmetro de distribuição e  $\epsilon$  o erro aleatório.

O modelo de Cobb-Douglas é um caso particular do modelo de produção CES quando  $\rho \rightarrow 0$ .

Griffiths *et al.* (1993, p. 722) apresentaram um exemplo ilustrativo da função de produção CES. Tomando logaritmo em ambos os lados de (3.14), obtiveram o modelo

$$\log(y_t) = \beta - \frac{\tau}{\rho} \log[\delta L_t^{-\rho} + (1 - \delta)K_t^{-\rho}] + \epsilon_t, \quad t = 1, \dots, 30,$$

em que  $\beta = \log \alpha$ . Eles assumiram que  $\epsilon_t$  são variáveis aleatórias independentes e identicamente distribuídas de média zero e variância constante  $\sigma^2$ . Se considerarmos os erros normalmente distribuídos, o EMV coincide com o de mínimos quadrados.

A Tabela 3.3 apresenta os valores ajustados e seus respectivos coeficientes de assimetria estimados. Esta tabela mostra que o coeficiente de assimetria estimado da distribuição de  $\hat{\delta}$  é muito elevado, indicando um afastamento acentuado (à direita) da simetria, embora o tamanho da amostra seja moderado ( $n = 30$ ). O afastamento da simetria pode estar relacionado ao tipo de não-linearidade do modelo, uma vez que os resultados de simulações para o modelo normal, mostraram uma certa simetria para pequenos valores amostrais ( $n < 30$ ).



Tabela 3.3: Estimativas dos coeficientes de assimetria da função de produção CES

Parâmetros	EMV	$\hat{\gamma}_1$
$\beta$	0,1245	-0,0015
$\tau$	1,0126	0,1401
$\delta$	0,3367	-2,0715
$\rho$	3,0109	0,2856

### 3.7.2 Modelo de crescimento do pasto

Consideramos aqui um conjunto de dados apresentado em Ratkowsky (1983), em que a variável resposta ( $Y_\ell$ ) é a taxa de crescimento do pasto e a covariável ( $x_\ell$ ) é o tempo decorrido desde o último corte do pasto. O modelo proposto é escrito (para  $\ell = 1, \dots, 9$ ) como

$$Y_\ell = \beta_1 - \beta_2 \exp\{-\exp[\beta_3 + \beta_4 \log(x_\ell)]\} + \epsilon_\ell,$$

em que os erros  $\epsilon_\ell$ 's são variáveis aleatórias independentes e identicamente distribuídas de média zero e variância constante  $\sigma^2$ .

A Tabela 3.4 apresenta os EMVs para os parâmetros  $\beta$  e seus respectivos coeficientes de assimetria. Neste exemplo, os coeficientes de assimetria dos termos lineares do modelo,  $\hat{\beta}_1$  e  $\hat{\beta}_2$ , sugerem que a aproximação normal não é adequada, enquanto que, para os estimadores dos parâmetros não-lineares, esta aproximação parece ser razoável.

Tabela 3.4: Estimativas dos coeficientes de assimetria para o modelo de crescimento do pasto

Parametros	EMV	$\hat{\gamma}_1$
$\beta_1$	69,9501	0,6610
$\beta_2$	61,6803	0,5601
$\beta_3$	-9,2090	-0,1064
$\beta_4$	2,3785	0,0702

### 3.8 Conclusões

Obtivemos uma fórmula bastante simples para o coeficiente de assimetria assintótico de ordem  $n^{-1/2}$  da distribuição dos EMVs de  $\beta$  e dos parâmetros de precisão e dispersão para os MNLFEs com dispersão desconhecida, porém a mesma para todas as observações, sendo possível notar que o mesmo é ponderado pelo inverso do quadrado do parâmetro de precisão. Além disso, pudemos observar através do estudo de simulação que existe uma certa influência da variável resposta na assimetria da distribuição do EMV, enquanto que para as covariáveis não observamos nenhuma influência significativa.

## Capítulo 4

# Aperfeiçoamento de testes escore em modelos não-lineares da família exponencial

### 4.1 Introdução

Os testes escore são bastante utilizados em Estatística e Econometria como uma alternativa aos testes da razão de verossimilhanças, principalmente quando a estimação segundo a hipótese alternativa é mais complicada do que segundo a hipótese nula. Nestes casos, o teste escore é mais simples pois requer apenas a estimação dos parâmetros segundo a hipótese nula. Iremos considerar, neste capítulo, o teste escore para a classe de modelos não-lineares da família exponencial (MNLFEs) com parâmetro de dispersão variável. O modelo considerado aqui é uma generalização do modelo linear generalizado com dispersão variável, definido por Smyth (1989). A importância de considerar a dispersão variável é o fato de que as distribuições contínuas mais utilizadas da família exponencial, como a normal, normal inversa e a gama, têm dispersão desconhecida e, além disso, é bastante comum incluir nos modelos de regressão uma estrutura de variação indicando que a dispersão não é constante ao longo das observações.

Nosso objetivo é calcular os coeficientes do fator de correção tipo-Bartlett (Cordeiro & Ferrari,

1991) para a estatística do teste escore.

## 4.2 Definição do modelo

Suponha as variáveis  $Y_1, \dots, Y_n$  independentes com cada  $Y_\ell$  tendo função densidade de probabilidade ou função de probabilidade na família exponencial da forma

$$\pi(y_\ell; \theta_\ell, \phi_\ell) = \exp \left\{ \phi_\ell [y_\ell \theta_\ell - b(\theta_\ell)] - \frac{1}{2} e(y_\ell, \phi_\ell) \right\}, \quad \ell = 1, \dots, n, \quad (4.1)$$

em que  $b(\cdot)$ ,  $c(\cdot)$  e  $e(\cdot, \cdot)$  são funções conhecidas e  $\theta_\ell$  e  $\phi_\ell > 0$  são chamados de parâmetros canônico e de precisão, respectivamente. Assumimos que  $\phi_\ell$  é desconhecido e  $\phi_\ell^{-1}$  é um parâmetro de dispersão. Além disso, consideramos que  $\phi_\ell^{-1} = \sigma^2 m_\ell$ , sendo  $m_\ell = m(z_\ell, \delta) > 0$  o  $\ell$ -ésimo elemento da matriz diagonal  $M$  de dimensão  $n \times n$ ,  $z_\ell^\top$  a  $\ell$ -ésima linha da matriz  $Z$  de dimensão  $n \times s$  de covariadas usadas para modelar a estrutura do parâmetro de precisão,  $\sigma^2$  uma constante desconhecida finita e estritamente positiva e  $\delta$  um vetor de dimensão  $q \times 1$  de parâmetros desconhecidos.

Nesta classe de modelos valem as relações:  $E(Y_\ell) = \mu_\ell = b'(\theta_\ell) = db(\theta_\ell)/d\theta_\ell$  e  $Var(Y_\ell) = \phi^{-1} V_\ell$ , sendo  $V_\ell = d\mu_\ell/d\theta_\ell$  denominada função de variância. Os MNLFs são definidos por (4.1) e pela componente sistemática

$$h(\mu_\ell) = \eta_\ell = f(x_\ell; \beta), \quad (4.2)$$

em que  $h(\cdot)$  é uma função conhecida monótona e diferenciável, chamada função de ligação,  $\beta = (\beta_1, \dots, \beta_p)^\top$ ,  $p < n$ , é o conjunto de parâmetros desconhecidos a serem estimados,  $f(\cdot; \cdot)$  é uma função contínua e diferenciável e  $x_\ell = (x_{\ell 1}, \dots, x_{\ell t})^\top$  é um vetor de valores conhecidos associado à resposta observada  $y_\ell$ .

Vamos supor identificabilidade no sentido de que diferentes  $\beta$ 's impliquem em diferentes  $\eta$ 's, em

que  $\eta = (\eta_1, \dots, \eta_n)^\top$ . Esta suposição fará com que a matriz de derivadas  $X^* = X^*(\beta) = \partial\eta/\partial\beta^\top$  tenha posto  $p$ , para todo  $\beta$ .

Assumimos válidas também as suposições usuais de regularidade (Cox & Hinkley, 1974) para a função de verossimilhança obtida a partir de (4.1) e (4.2).

Supondo em (4.1) que  $e(y_\ell, \phi_\ell) = s(\phi_\ell) + t(y_\ell)$ , podemos reescrever (4.1) como

$$\pi(y_\ell; \theta_\ell, \phi_\ell) = \exp\left(-\frac{1}{2}\{\phi_\ell d(y_\ell) + s(\phi_\ell) + t(y_\ell)\}\right), \quad \ell = 1, \dots, n, \quad (4.3)$$

sendo  $d(y_\ell) = d_\ell = -2[y_\ell\theta_\ell - b(\theta_\ell) - c(y_\ell)]$ , que corresponde a um modelo da família exponencial de distribuições na forma natural com parâmetros canônicos  $\phi_\ell$  e  $\phi_\ell\theta_\ell$ . Admite-se que  $s(\phi_\ell)$  possui as quatro primeiras derivadas. Na Tabela 4.1 temos as funções  $s(\phi_\ell)$  e  $d_\ell$  para as distribuições normal, normal inversa e gama.

Tabela 4.1: Alguns modelos especiais

Modelo	$s(\phi_\ell)$	$d_\ell$
Normal	$-\log(\phi_\ell)$	$(y_\ell - \mu_\ell)^2$
Normal inversa	$-\log(\phi_\ell)$	$\frac{(y_\ell - \mu_\ell)^2}{\mu_\ell^2 y_\ell}$
Gama	$-2\{\phi_\ell \log(\phi_\ell) - \log \Gamma(\phi_\ell)\}$	$2\left\{\frac{y_\ell}{\mu_\ell} - \log\left(\frac{y_\ell}{\mu_\ell}\right)\right\}$

O logaritmo da função de verossimilhança do vetor de parâmetros  $(\beta^\top, \delta^\top, \sigma^2)$ , dado o vetor de observações  $(y_1, \dots, y_n)$ , do modelo definido por (4.3) é

$$\mathcal{L}(\beta^\top, \delta^\top, \sigma^2) = -\frac{1}{2} \sum_{\ell=1}^n \left\{ \frac{1}{\sigma^2 m_\ell} d_\ell + t(y_\ell) + s(\phi_\ell) \right\}, \quad \ell = 1, \dots, n.$$

Para esta parametrização temos que  $(\delta^\top, \sigma^2)^\top$  é ortogonal a  $\beta$ , porém  $\delta$  e  $\sigma^2$  não são ortogonais.

Uma transformação que torna os parâmetros ortogonais para as distribuições normal e normal inversa é

$$\sigma^2 = \frac{\gamma}{(\prod_{\ell=1}^n m_\ell)^{1/n}}.$$

Para a distribuição gama não é possível encontrar uma reparametrização que torne os parâmetros  $\delta$  e  $\sigma^2$  ortogonais.

O logaritmo da função de verossimilhança para o modelo reparametrizado é dado por

$$\mathcal{L}(\beta, \delta, \gamma) = -\frac{1}{2} \sum_{\ell=1}^n \left\{ \frac{q_\ell}{\gamma} d_\ell + t(y_\ell) + \log \left( \frac{\gamma}{q_\ell} \right) \right\}, \ell = 1, \dots, n, \quad (4.4)$$

em que

$$q_\ell = q_\ell(\delta) = \frac{(\prod_{s=1}^n m_s)^{1/n}}{m_\ell}.$$

A função escore total  $U = U(\beta, \delta, \gamma) = (U_\beta^\top, U_\delta^\top, U_\gamma^\top)^\top$  tem componentes dadas por:

$$U_\beta = \partial \mathcal{L}(\beta, \delta, \gamma) / \partial \beta = \tilde{X}^\top \Phi T V^{-1} (y - \mu),$$

$$U_\delta = \partial \mathcal{L}(\beta, \delta, \gamma) / \partial \delta = -\frac{1}{2\gamma} \dot{Q}^\top d$$

e

$$U_\gamma = \partial \mathcal{L}(\beta, \delta, \gamma) / \partial \gamma = \frac{\mathbf{1}^\top}{2\gamma} (\Phi d - \mathbf{1}),$$

em que  $\Phi$ ,  $T$  e  $V$  são matrizes diagonais de dimensão  $n \times n$  cujos respectivos elementos são dados por  $\phi_\ell = q_\ell/\gamma$ ,  $T_\ell = d\mu_\ell/d\eta_\ell$  e  $V_\ell = d\mu_\ell/d\theta_\ell$ ,  $\dot{Q}$  é uma matriz  $n \times q$  com a  $\ell$ -ésima linha dada por  $\partial q_\ell / \partial \delta^\top$ ,  $\ell = 1, \dots, n$ ,  $d = (d_1, \dots, d_n)^\top$ ,  $\mu = (\mu_1, \dots, \mu_n)^\top$  e  $\mathbf{1}$  é um vetor de uns de dimensão  $n$ .

Pode ser mostrado facilmente que:

1.  $E(d_\ell) = -\dot{s}\left(\frac{q_\ell}{\gamma}\right) = \frac{\gamma}{q_\ell}$ ;
2.  $\text{Var}(d_\ell) = 2\ddot{s}\left(\frac{q_\ell}{\gamma}\right) = \frac{2\gamma^2}{q_\ell^2}$ ;
3.  $E(Y_\ell d_\ell) = \mu_\ell \frac{\gamma}{q_\ell}$ .

Estas propriedades são úteis para o cálculo de alguns cumulantes de derivadas do logaritmo da função de verossimilhança total.

Considerando o caso em que  $m_\ell = \exp\{z_\ell^\top \delta\}$ , obtemos  $q_\ell = \exp\{-(z_\ell - \bar{z})^\top \delta\}$ . Assim, a matriz de informação total de Fisher para este caso é dada por

$$K = -E\left(\frac{\partial^2 \mathcal{L}}{\partial \psi \partial \psi^\top}\right) = \begin{pmatrix} X^{*\top} W \Phi X^* & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2}(Z - \bar{Z})(Z - \bar{Z})^\top & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{n}{2\gamma^2} \end{pmatrix},$$

em que  $\psi = (\beta, \delta, \gamma)^\top$  é o vetor de parâmetros e  $\mathbf{0}$  são matrizes nulas de dimensões apropriadas.

### 4.3 Melhoramento do teste escore

Para o caso em que  $m_\ell = \exp\{z_\ell^\top \delta\}$ , vamos considerar o problema de testar hipóteses do tipo  $H_0 : \beta_1 = \beta_1^{(0)}, \delta_1 = \delta_1^{(0)}$  contra a hipótese alternativa  $H_1$  : violação de pelo menos uma das igualdades, em que  $\beta_1^{(0)}$  e  $\delta_1^{(0)}$  são vetores especificados de dimensões  $p_1$  e  $q_1$  respectivamente. Assumimos que  $0 < p_1 \leq p$  e  $0 < q_1 \leq q$ . Considerando  $p_1 < p$  e  $q_1 < q$  e seguindo a partição induzida por  $H_0$ , sejam  $X^* = (X_1^*, X_2^*)$  e  $Z = (Z_1, Z_2)$  as matrizes do modelo correspondente a esta partição, em que  $X_1^*, X_2^*, Z_1$  e  $Z_2$  são matrizes de posto completo  $n \times p_1$ ,  $n \times (p - p_1)$ ,  $n \times q_1$  e  $n \times (q - q_1)$ , respectivamente. Se  $p_1 = p$  definimos  $X_1^* = X^*$  e analogamente se  $q_1 = q$  definimos  $Z_1 = Z$ . Denotamos os EMVs irrestritos de  $\beta$  e  $\delta$  por  $\hat{\beta}$  e  $\hat{\delta}$ , enquanto os EMVs restritos dos parâmetros  $\beta_2$  e  $\delta_2$  sob  $H_0$  são denotados por  $\tilde{\beta}_2$  e  $\tilde{\delta}_2$ . Similarmente, todas as quantidades calculadas nos EMVs

irrestritos serão denotadas pela adição de um circunflexo, enquanto todas as quantidades calculadas nos EMVs restritos serão denotadas pela adição de um til.

A estatística escore para testar  $H_0$  é dada por  $S_R = \tilde{U}^\top \tilde{K}^{-1} \tilde{U}$ . Seguindo Cordeiro & Ferrari (1991), podemos melhorar a estatística escore  $S_R$  por meio da estatística escore modificada dada por

$$S_R^* = S_R \{1 - (c + bS_R + aS_R^2)\}, \quad (4.5)$$

em que os coeficientes  $a$ ,  $b$ , e  $c$  são dados por

$$a = \frac{A_3}{12\nu(\nu+2)(\nu+4)}, \quad b = \frac{(A_2 - 2A_3)}{12\nu(\nu+2)} \quad \text{e} \quad c = \frac{(A_1 - A_2 + A_3)}{12\nu}, \quad (4.6)$$

sendo  $\nu$  o número de graus de liberdade da distribuição aproximada de  $S_R$  sob  $H_0$  e os coeficientes  $A_1$ ,  $A_2$  e  $A_3$  obtidos de Harris (1985).

Consideremos as matrizes  $Z_\beta = X^*(X^{*\top}W\Phi X^*)^{-1}X^{*\top}$ ,  $Z_{\beta_2} = X_2^*(X_2^{*\top}W\Phi X_2^*)^{-1}X_2^{*\top}$ , (para  $q \leq p$ ) e  $P_2 = \text{diag}\{p_1, \dots, p_n\}$  com  $p_\ell = \text{tr}\{X_2^*(X_2^{*\top}W\Phi X_2^*)^{-1}D_{22}^{(\ell)}\}$ , em que

$$D^{(\ell)} = \left\{ \frac{\partial^2 \eta_\ell}{\partial \beta_i \partial \beta_j} \right\} = \begin{pmatrix} D_{11}^{(\ell)} & D_{12}^{(\ell)} \\ D_{21}^{(\ell)} & D_{22}^{(\ell)} \end{pmatrix}.$$

Além disso, definimos  $C_2 = \{c_{\ell m}\}$  e  $J_2 = \{j_{\ell m}\}$ , com

$$c_{\ell m} = x_{2m}^* (X_2^{*\top}W\Phi X_2^*)^{-1} D_{22}^{(\ell)} (X_2^{*\top}W\Phi X_2^*)^{-1} x_{2m}^{*\top}$$

e

$$j_{\ell m} = \text{tr}\{D_{22}^{(\ell)} (X_2^{*\top}W\Phi X_2^*)^{-1} D_{22}^{(m)} (X_2^{*\top}W\Phi X_2^*)^{-1}\},$$



em que  $x_{2m}^*$  denota a  $m$ -ésima linha de  $X_2^*$ . Sejam  $F = \text{diag}\{f_1, \dots, f_n\}$ ,  $G = \text{diag}\{g_1, \dots, g_n\}$ ,  $B = \text{diag}\{b_1, \dots, b_n\}$  e  $H = \text{diag}\{h_1, \dots, h_n\}$  com  $f_\ell$ ,  $g_\ell$ ,  $b_\ell$  e  $h_\ell$  definidos pelas expressões

$$f_\ell = \left( \frac{d\theta_\ell}{d\eta_\ell} \right) \frac{d^2\mu_\ell}{d\eta_\ell^2} = V_\ell^{-1} \left( \frac{d\mu_\ell}{d\eta_\ell} \right) \frac{d^2\mu_\ell}{d\eta_\ell^2}, \quad (4.7)$$

$$g_\ell = \left( \frac{d\mu_\ell}{d\eta_\ell} \right) \frac{d^2\theta_\ell}{d\eta_\ell^2} = V_\ell^{-1} \left( \frac{d\mu_\ell}{d\eta_\ell} \right) \frac{d^2\mu_\ell}{d\eta_\ell^2} - V_\ell^{-2} \left( \frac{d\mu_\ell}{d\eta_\ell} \right)^3 \frac{dV_\ell}{d\mu_\ell}, \quad (4.8)$$

$$b_\ell = V_\ell^{-3} \left( \frac{d\mu_\ell}{d\eta_\ell} \right)^4 \left\{ \left( \frac{dV_\ell}{d\mu_\ell} \right)^2 + V_\ell \frac{d^2V_\ell}{d\mu_\ell^2} \right\} \quad (4.9)$$

e

$$h_\ell = V_\ell^{-2} \frac{dV_\ell}{d\mu_\ell} \left( \frac{d\mu_\ell}{d\eta_\ell} \right)^2 \frac{d^2\mu_\ell}{d\eta_\ell^2} + V_\ell^{-2} \frac{d^2V_\ell}{d\mu_\ell^2} \left( \frac{d\mu_\ell}{d\eta_\ell} \right)^4. \quad (4.10)$$

Como  $m_\ell = \exp\{z_\ell^\top \delta\}$ , obtemos  $q_\ell = \exp\{-(z_\ell - \bar{z})^\top \delta\}$ . Daí, podemos definir as matrizes  $Z_\delta = 2(Z - \bar{Z})[(Z - \bar{Z})^\top (Z - \bar{Z})]^{-1} (Z - \bar{Z})^\top$  e  $Z_{\delta 2} = 2(Z_2 - \bar{Z}_2)[(Z_2 - \bar{Z}_2)^\top (Z_2 - \bar{Z}_2)]^{-1} (Z_2 - \bar{Z}_2)^\top$ , em que  $\Phi = (1/\gamma)Q$ , sendo  $Q = \text{diag}\{q_1, \dots, q_n\}$ .

Denotamos  $Z^{(2)} = Z \odot Z$ ,  $Z^{(3)} = Z \odot Z \odot Z$ , em que  $\odot$  é o produto direto de matrizes. O sub-índice  $d$  indicará que uma matriz diagonal foi obtida da matriz original.

Portanto, podemos escrever os  $A$ 's da seguinte forma:

$$A_1 = A_{11} + A_{12} + A_{13} + A_{14},$$

$$A_2 = A_{21} + A_{22} + A_{23} + A_{24}$$

e

$$A_3 = A_{31} + A_{32},$$

em que

$$\begin{aligned} A_{11} &= 3\mathbf{1}^\top \Phi F Z_{\beta_2 d} (Z_\beta - Z_{\beta_2}) Z_{\beta_2 d} F \Phi \mathbf{1} + 6\mathbf{1}^\top \Phi W P_2 (Z_\beta - Z_{\beta_2}) Z_{\beta_2 d} F \Phi \mathbf{1} \\ &+ 3\mathbf{1}^\top \Phi W P_2 (Z_\beta - Z_{\beta_2}) P_2 W \Phi \mathbf{1} + 3\mathbf{1}^\top \Phi W Z_{\beta_2 d} (Z_\delta - Z_{\delta_2}) Z_{\delta_2 d} \mathbf{1} \\ &+ 3\mathbf{1}^\top \Phi W Z_{\beta_2 d} (Z_\delta - Z_{\delta_2}) Z_{\beta_2 d} W \Phi \mathbf{1} + \frac{3}{4} \mathbf{1}^\top Z_{\delta_2 d} (Z_\delta - Z_{\delta_2}) Z_{\delta_2 d} \mathbf{1}, \end{aligned}$$

$$\begin{aligned} A_{12} &= 6\mathbf{1}^\top \Phi F Z_{\beta_2 d} Z_{\beta_2} (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1} + 6\mathbf{1}^\top \Phi W P_2 Z_{\beta_2} (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1} \\ &+ 6\mathbf{1}^\top \Phi W Z_{\beta_2 d} Z_{\delta_2} (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} + 6\mathbf{1}^\top \Phi W Z_{\beta_2 d} Z_{\delta_2} (Z_\delta - Z_{\delta_2})_d \mathbf{1} \\ &+ \frac{12}{n} \mathbf{1}^\top \Phi W Z_{\beta_2 d} (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} + \frac{12}{n} \mathbf{1}^\top \Phi W Z_{\beta_2 d} (Z_\delta - Z_{\delta_2})_d \mathbf{1} \\ &+ 3\mathbf{1}^\top Z_{\delta_2 d} Z_{\delta_2} (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} + 3\mathbf{1}^\top Z_{\delta_2 d} Z_{\delta_2} (Z_\delta - Z_{\delta_2})_d \mathbf{1} \\ &+ \frac{6}{n} \mathbf{1}^\top Z_{\delta_2 d} (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} + \frac{6}{n} \mathbf{1}^\top Z_{\delta_2 d} (Z_\delta - Z_{\delta_2})_d \mathbf{1}, \end{aligned}$$

$$\begin{aligned} A_{13} &= -6\mathbf{1}^\top \Phi (2G - F) \left[ Z_{\beta_2}^{(2)} \odot (Z_\beta - Z_{\beta_2}) \right] F \Phi \mathbf{1} - 6\mathbf{1}^\top \Phi W \left[ (Z_\beta - Z_{\beta_2}) \odot J_2 \right] W \Phi \mathbf{1} \\ &- 6\mathbf{1}^\top \Phi W \left[ (Z_\beta - Z_{\beta_2}) \odot C_2^\top \right] F \Phi \mathbf{1} + \frac{9}{2} \mathbf{1}^\top \left[ Z_{\delta_2}^{(2)} \odot (Z_\delta - Z_{\delta_2}) \right] \mathbf{1} + \frac{18}{n} \mathbf{1}^\top \left[ Z_{\delta_2} \odot (Z_\delta - Z_{\delta_2}) \right] \mathbf{1} \\ &- 6\mathbf{1}^\top \Phi (2G - F) \left[ (Z_\beta - Z_{\beta_2}) \odot C_2 \right] W \Phi \mathbf{1} \\ &+ 6\mathbf{1}^\top \Phi W \left[ (Z_\delta - Z_{\delta_2}) \odot Z_{\beta_2}^{(2)} \right] W \Phi \mathbf{1} + 12\mathbf{1}^\top \Phi W \left[ (Z_\beta - Z_{\beta_2}) \odot Z_{\beta_2} Z_{\delta_2} \right] W \Phi \mathbf{1}, \end{aligned}$$

$$\begin{aligned} A_{14} &= -12\mathbf{1}^\top \Phi W Z_{\beta_2 d} (Z_\delta - Z_{\delta_2})_d \mathbf{1} - 6\mathbf{1}^\top \Phi H (Z_\beta - Z_{\beta_2})_d Z_{\beta_2 d} \mathbf{1} \\ &- 12\mathbf{1}^\top (Z_\delta - Z_{\delta_2})_d Z_{\delta_2 d} \mathbf{1} - 6\mathbf{1}^\top \Phi W (Z_\beta - Z_{\beta_2})_d Z_{\delta_2 d} \mathbf{1} \end{aligned}$$

$$\begin{aligned}
& - 6\mathbf{1}^\top \Phi(F - G)P_2(Z_\beta - Z_{\beta_2})_d \mathbf{1} - \frac{12}{n} \mathbf{1}^\top (Z_\delta - Z_{\delta_2})_d \mathbf{1}, \\
A_{21} & = -3\mathbf{1}^\top \Phi(F - G)(Z_\beta - Z_{\beta_2})_d Z_{\beta_2} (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1} - \frac{6}{n} \mathbf{1}^\top [(Z_\delta - Z_{\delta_2})_d]^2 \mathbf{1} \\
& - 3\mathbf{1}^\top \Phi W (Z_\beta - Z_{\beta_2})_d Z_{\delta_2} (Z_\delta - Z_{\delta_2})_d \mathbf{1} - 3\mathbf{1}^\top (Z_\delta - Z_{\delta_2})_d Z_{\delta_2} (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} \\
& - \frac{12}{n} \mathbf{1}^\top \Phi W (Z_\beta - Z_{\beta_2})_d (Z_\delta - Z_{\delta_2})_d \mathbf{1} - 3\mathbf{1}^\top (Z_\delta - Z_{\delta_2})_d Z_{\delta_2} (Z_\delta - Z_{\delta_2})_d \mathbf{1} \\
& - 3\mathbf{1}^\top \Phi W (Z_\beta - Z_{\beta_2})_d Z_{\delta_2} (Z_\beta - Z_{\beta_2})_d \mathbf{1} - \frac{6}{n} \mathbf{1}^\top [\Phi W (Z_\beta - Z_{\beta_2})_d]^2 \mathbf{1}, \\
A_{22} & = -6\mathbf{1}^\top \Phi F Z_{\beta_2 d} (Z_\beta - Z_{\beta_2}) (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1} \\
& - 6\mathbf{1}^\top \Phi W Z_{\beta_2 d} (Z_\delta - Z_{\delta_2}) (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} + 6\mathbf{1}^\top \Phi W Z_{\beta_2 d} (Z_\delta - Z_{\delta_2}) (Z_\delta - Z_{\delta_2})_d \mathbf{1} \\
& - 3\mathbf{1}^\top \Phi W Z_{\delta_2 d} (Z_\delta - Z_{\delta_2}) (Z_\beta - Z_{\beta_2})_d \mathbf{1} - 3\mathbf{1}^\top Z_{\delta_2 d} (Z_\delta - Z_{\delta_2}) (Z_\delta - Z_{\delta_2})_d \mathbf{1} \\
& - 6\mathbf{1}^\top \Phi W P_2(Z_\beta - Z_{\beta_2}) (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1}, \\
A_{23} & = -6\mathbf{1}^\top \Phi(F - G) \left[ Z_{\beta_2} \odot (Z_\beta - Z_{\beta_2})^{(2)} \right] (F - G) \Phi \mathbf{1} - \frac{12}{n} \mathbf{1}^\top \left[ (Z_\delta - Z_{\delta_2})^{(2)} \right] \mathbf{1} \\
& - 6\mathbf{1}^\top \Phi W \left[ Z_{\delta_2} \odot (Z_\beta - Z_{\beta_2})^{(2)} \right] W \Phi \mathbf{1} - 6\mathbf{1}^\top \left[ Z_{\delta_2} \odot (Z_\delta - Z_{\delta_2})^{(2)} \right] \mathbf{1} \\
& - 12\mathbf{1}^\top \Phi W [Z_{\beta_2} \odot (Z_\beta - Z_{\beta_2}) \odot (Z_\delta - Z_{\delta_2})] W \Phi \mathbf{1} - \frac{12}{n} \mathbf{1}^\top \Phi W \left[ (Z_\beta - Z_{\beta_2})^{(2)} \right] W \Phi \mathbf{1}, \\
A_{24} & = 3\mathbf{1}^\top \Phi B (Z_\beta - Z_{\beta_2})_d^2 \mathbf{1} + 12\mathbf{1}^\top \Phi W (Z_\delta - Z_{\delta_2})_d (Z_\beta - Z_{\beta_2})_d \mathbf{1} + 9\mathbf{1}^\top (Z_\delta - Z_{\delta_2})_d^2 \mathbf{1}, \\
A_{31} & = 3\mathbf{1}^\top \Phi(F - G)(Z_\beta - Z_{\beta_2})_d (Z_\beta - Z_{\beta_2}) (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1} \\
& + 3\mathbf{1}^\top (Z_\delta - Z_{\delta_2})_d (Z_\delta - Z_{\delta_2}) (Z_\delta - Z_{\delta_2})_d \mathbf{1} + 6\mathbf{1}^\top \Phi W (Z_\beta - Z_{\beta_2})_d (Z_\delta - Z_{\delta_2}) (Z_\delta - Z_{\delta_2})_d \mathbf{1} \\
& + 3\mathbf{1}^\top \Phi W (Z_\beta - Z_{\beta_2})_d (Z_\delta - Z_{\delta_2}) (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1},
\end{aligned}$$

$$\begin{aligned}
 A_{32} &= 2\mathbf{1}^\top \Phi(F - G) \left[ (Z_\beta - Z_{\beta_2})^{(3)} \right] (F - G)\Phi\mathbf{1} + 6\mathbf{1}^\top \Phi W \left[ (Z_\delta - Z_{\delta_2}) \odot (Z_\beta - Z_{\beta_2})^{(2)} \right] W\Phi\mathbf{1} \\
 &+ 2\mathbf{1}^\top \left[ (Z_\delta - Z_{\delta_2})^{(3)} \right] \mathbf{1}.
 \end{aligned}$$

Maiores detalhes sobre a dedução destas expressões se encontram no Apêndice B.

Podemos observar que as expressões dadas acima dependem basicamente das matrizes  $Z_{\beta_2}$ ,  $Z_\beta - Z_{\beta_2}$ ,  $Z_{\delta_2}$ ,  $Z_\delta - Z_{\delta_2}$  e das funções de ligação e de variância com suas duas primeiras derivadas. Caso dependam também de parâmetros desconhecidos, estes devem ser substituídos por suas respectivas estimativas de máxima verossimilhança sob  $H_0$ . Uma vez obtidas estas quantidades, o cálculo de  $A_1$ ,  $A_2$  e  $A_3$  é conduzido sem dificuldade, pois envolve apenas operações simples com matrizes e vetores. Embora este cálculo seja simples, a interpretação de  $A_1$ ,  $A_2$  e  $A_3$  é difícil. As expressões acima foram obtidas para o caso mais geral possível, em termos das hipóteses  $H_0$  e  $H_1$ . Desta forma elas podem ser usadas para obter correções em amostras finitas para os testes escore em muitos casos particulares como veremos a seguir.

#### 4.3.1 Testando apenas componentes do vetor de parâmetros $\delta$

Suponha que estamos interessados em testar  $H_0 : \delta_1 = \delta_1^{(0)}$ , contra a hipótese alternativa  $H_1 : \delta_1 \neq \delta_1^{(0)}$ . Para este caso, basta considerar  $Z_\beta = Z_{\beta_2}$ . Assim, os coeficientes da estatística escore corrigida são dados por:

$$\begin{aligned}
 A_{11} &= 3\mathbf{1}^\top \Phi W Z_{\beta_d} (Z_\delta - Z_{\delta_2}) Z_{\delta_{2d}} \mathbf{1} + 3\mathbf{1}^\top \Phi W Z_{\beta_d} (Z_\delta - Z_{\delta_2}) Z_{\beta_d} W \Phi \mathbf{1} \\
 &+ \frac{3}{4} \mathbf{1}^\top Z_{\delta_{2d}} (Z_\delta - Z_{\delta_2}) Z_{\delta_{2d}} \mathbf{1},
 \end{aligned}$$

$$\begin{aligned}
A_{12} &= 6\mathbf{1}^\top \Phi W Z_{\beta_d} Z_{\delta_2} (Z_\delta - Z_{\delta_2})_d \mathbf{1} + \frac{12}{n} \mathbf{1}^\top \Phi W Z_{\beta_d} (Z_\delta - Z_{\delta_2})_d \mathbf{1} \\
&+ 3\mathbf{1}^\top Z_{\delta_2 d} Z_{\delta_2} (Z_\delta - Z_{\delta_2})_d \mathbf{1} + \frac{6}{n} \mathbf{1}^\top Z_{\delta_2 d} (Z_\delta - Z_{\delta_2})_d \mathbf{1},
\end{aligned}$$

$$\begin{aligned}
A_{13} &= \frac{9}{2} \mathbf{1}^\top \left[ Z_{\delta_2}^{(2)} \odot (Z_\delta - Z_{\delta_2}) \right] \mathbf{1} + \frac{18}{n} \mathbf{1}^\top \left[ Z_{\delta_2} \odot (Z_\delta - Z_{\delta_2}) \right] \mathbf{1} \\
&+ 6\mathbf{1}^\top \Phi W \left[ (Z_\delta - Z_{\delta_2}) \odot Z_{\beta}^{(2)} \right] W \Phi \mathbf{1},
\end{aligned}$$

$$\begin{aligned}
A_{14} &= -12\mathbf{1}^\top \Phi W Z_{\beta_d} (Z_\delta - Z_{\delta_2})_d \mathbf{1} - 12\mathbf{1}^\top (Z_\delta - Z_{\delta_2})_d Z_{\delta_2 d} \mathbf{1} \\
&- \frac{12}{n} \mathbf{1}^\top (Z_\delta - Z_{\delta_2})_d \mathbf{1},
\end{aligned}$$

$$A_{21} = -\frac{6}{n} \mathbf{1}^\top \left[ (Z_\delta - Z_{\delta_2})_d \right]^2 \mathbf{1} - 3\mathbf{1}^\top (Z_\delta - Z_{\delta_2})_d Z_{\delta_2} (Z_\delta - Z_{\delta_2})_d \mathbf{1},$$

$$A_{22} = 6\mathbf{1}^\top \Phi W Z_{\beta_d} (Z_\delta - Z_{\delta_2}) (Z_\delta - Z_{\delta_2})_d \mathbf{1} - 3\mathbf{1}^\top Z_{\delta_2 d} (Z_\delta - Z_{\delta_2}) (Z_\delta - Z_{\delta_2})_d \mathbf{1},$$

$$A_{23} = -\frac{12}{n} \mathbf{1}^\top \left[ (Z_\delta - Z_{\delta_2})^{(2)} \right] \mathbf{1} - 6\mathbf{1}^\top \left[ Z_{\delta_2} \odot (Z_\delta - Z_{\delta_2})^{(2)} \right] \mathbf{1},$$

$$A_{24} = 9\mathbf{1}^\top (Z_\delta - Z_{\delta_2})_d^2 \mathbf{1},$$

$$A_{31} = 3\mathbf{1}^\top (Z_\delta - Z_{\delta_2})_d (Z_\delta - Z_{\delta_2}) (Z_\delta - Z_{\delta_2})_d \mathbf{1},$$

$$A_{32} = 2\mathbf{1}^\top \left[ (Z_\delta - Z_{\delta_2})^{(3)} \right] \mathbf{1}.$$

Além disso, se nosso interesse é testar todos os componentes do vetor de parâmetros  $\delta$ , basta considerar  $Z_{\delta 2} = 0$ . Desta forma, obtemos

$$A_{11} = 3\mathbf{1}^\top \Phi W Z_{\beta_d} Z_\delta Z_{\beta_d} W \Phi \mathbf{1},$$

$$A_{12} = \frac{12}{n} \mathbf{1}^\top \Phi W Z_{\beta_d} Z_{\delta d} \mathbf{1},$$

$$A_{13} = 6\mathbf{1}^\top \Phi W \left[ Z_\delta \odot Z_{\beta_d}^{(2)} \right] W \Phi \mathbf{1},$$

$$A_{14} = -12\mathbf{1}^\top \Phi W Z_{\beta_d} Z_{\delta d} \mathbf{1} - \frac{12}{n} \mathbf{1}^\top Z_{\delta d} \mathbf{1},$$

$$A_{21} = -\frac{6}{n} \mathbf{1}^\top [Z_{\delta d}]^2 \mathbf{1},$$

$$A_{22} = 6\mathbf{1}^\top \Phi W Z_{\beta_d} Z_\delta Z_{\delta d} \mathbf{1},$$

$$A_{23} = -\frac{12}{n} \mathbf{1}^\top \left[ Z_\delta^{(2)} \right] \mathbf{1},$$

$$A_{24} = 9\mathbf{1}^\top Z_{\delta d}^2 \mathbf{1},$$

$$A_{31} = 3\mathbf{1}^\top Z_{\delta d} Z_\delta Z_{\delta d} \mathbf{1},$$

$$A_{32} = 2\mathbf{1}^\top \left[ Z_\delta^{(3)} \right] \mathbf{1}.$$

Um teste bastante interessante aqui nesta classe é o teste de homoscedasticidade (para o caso normal) ou dispersão constante,  $H_0 : \delta = 0$  contra  $H_1 : \delta \neq 0$ .

#### 4.3.2 Testando apenas componentes do vetor de parâmetros $\beta$

Suponha que estamos interessados em testar  $H_0 : \beta_1 = \beta_1^{(0)}$ , contra a hipótese alternativa  $H_1 : \beta_1 \neq \beta_1^{(0)}$ . Para este caso, basta considerar  $Z_\delta = Z_{\delta_2}$ . Assim, os coeficientes da estatística escore corrigida são dados por:

$$\begin{aligned} A_{11} &= 3\mathbf{1}^\top \Phi F Z_{\beta_2 d} (Z_\beta - Z_{\beta_2}) Z_{\beta_2 d} F \Phi \mathbf{1} + 6\mathbf{1}^\top \Phi W P_2 (Z_\beta - Z_{\beta_2}) Z_{\beta_2 d} F \Phi \mathbf{1} \\ &+ 3\mathbf{1}^\top \Phi W P_2 (Z_\beta - Z_{\beta_2}) P_2 W \Phi \mathbf{1}, \end{aligned}$$

$$\begin{aligned} A_{12} &= 6\mathbf{1}^\top \Phi F Z_{\beta_2 d} Z_{\beta_2} (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1} + 6\mathbf{1}^\top \Phi W P_2 Z_{\beta_2} (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1} \\ &+ 6\mathbf{1}^\top \Phi W Z_{\beta_2 d} Z_\delta (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} + \frac{12}{n} \mathbf{1}^\top \Phi W Z_{\beta_2 d} (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} \\ &+ 3\mathbf{1}^\top Z_{\delta d} Z_\delta (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} + \frac{6}{n} \mathbf{1}^\top Z_{\delta d} (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1}, \end{aligned}$$

$$\begin{aligned} A_{13} &= -6\mathbf{1}^\top \Phi (2G - F) [Z_{\beta_2}^{(2)} \odot (Z_\beta - Z_{\beta_2})] F \Phi \mathbf{1} - 6\mathbf{1}^\top \Phi W [(Z_\beta - Z_{\beta_2}) \odot J_2] W \Phi \mathbf{1} \\ &- 6\mathbf{1}^\top \Phi W [(Z_\beta - Z_{\beta_2}) \odot C_2^\top] F \Phi \mathbf{1} - 6\mathbf{1}^\top \Phi (2G - F) [(Z_\beta - Z_{\beta_2}) \odot C_2] W \Phi \mathbf{1} \\ &+ 12\mathbf{1}^\top \Phi W [(Z_\beta - Z_{\beta_2}) \odot Z_{\beta_2} Z_\delta] W \Phi \mathbf{1}, \end{aligned}$$

$$\begin{aligned} A_{14} &= -6\mathbf{1}^\top \Phi H (Z_\beta - Z_{\beta_2})_d Z_{\beta_2 d} \mathbf{1} - 6\mathbf{1}^\top \Phi W (Z_\beta - Z_{\beta_2})_d Z_{\delta d} \mathbf{1} \\ &- 6\mathbf{1}^\top \Phi (F - G) P_2 (Z_\beta - Z_{\beta_2})_d \mathbf{1}, \end{aligned}$$

$$A_{21} = -3\mathbf{1}^\top \Phi(F - G)(Z_\beta - Z_{\beta_2})_d Z_{\beta_2} (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1} - 3\mathbf{1}^\top \Phi W (Z_\beta - Z_{\beta_2})_d Z_\delta (Z_\beta - Z_{\beta_2})_d \mathbf{1} \\ - \frac{6}{n} \mathbf{1}^\top [\Phi W (Z_\beta - Z_{\beta_2})_d]^2 \mathbf{1},$$

$$A_{22} = -6\mathbf{1}^\top \Phi F Z_{\beta_2 d} (Z_\beta - Z_{\beta_2}) (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1} \\ - 6\mathbf{1}^\top \Phi W P_2 (Z_\beta - Z_{\beta_2}) (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1},$$

$$A_{23} = -6\mathbf{1}^\top \Phi (F - G) \left[ Z_{\beta_2} \odot (Z_\beta - Z_{\beta_2})^{(2)} \right] (F - G) \Phi \mathbf{1} - 6\mathbf{1}^\top \Phi W [Z_\delta \odot (Z_\beta - Z_{\beta_2})^{(2)}] W \Phi \mathbf{1} \\ - \frac{12}{n} \mathbf{1}^\top \Phi W \left[ (Z_\beta - Z_{\beta_2})^{(2)} \right] W \Phi \mathbf{1},$$

$$A_{24} = 3\mathbf{1}^\top \Phi B (Z_\beta - Z_{\beta_2})_d^2 \mathbf{1},$$

$$A_{31} = 3\mathbf{1}^\top \Phi (F - G) (Z_\beta - Z_{\beta_2})_d (Z_\beta - Z_{\beta_2}) (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1},$$

$$A_{32} = 2\mathbf{1}^\top \Phi (F - G) \left[ (Z_\beta - Z_{\beta_2})^{(3)} \right] (F - G) \Phi \mathbf{1}.$$

Além disso, se nosso interesse é testar todos os componentes do vetor de parâmetros  $\beta$ , basta considerar  $Z_{\beta_2} = P_2 = C_2 = J_2 = 0$ . Desta forma, obtemos

$$A_{11} = 0,$$

$$A_{12} = 3\mathbf{1}^\top Z_{\delta d} Z_\delta Z_{\beta d} W \Phi \mathbf{1} + \frac{6}{n} \mathbf{1}^\top Z_{\delta d} Z_{\beta d} W \Phi \mathbf{1},$$



$$A_{13} = 0,$$

$$A_{14} = -6\mathbf{1}^\top \Phi W Z_{\beta_d} Z_{\delta_d} \mathbf{1},$$

$$A_{21} = -3\mathbf{1}^\top \Phi W Z_{\beta_d} Z_{\delta} Z_{\beta_d} \mathbf{1} - \frac{6}{n} \mathbf{1}^\top [\Phi W Z_{\beta_d}]^2 \mathbf{1},$$

$$A_{22} = 0,$$

$$A_{23} = -6\mathbf{1}^\top \Phi W \left[ Z_{\delta} \odot Z_{\beta}^{(2)} \right] W \Phi \mathbf{1} - \frac{12}{n} \mathbf{1}^\top \Phi W [Z_{\beta}^{(2)}] W \Phi \mathbf{1},$$

$$A_{24} = 3\mathbf{1}^\top \Phi B Z_{\beta_d}^2 \mathbf{1},$$

$$A_{31} = 3\mathbf{1}^\top \Phi (F - G) Z_{\beta_d} Z_{\beta} Z_{\beta_d} (F - G) \Phi \mathbf{1},$$

$$A_{32} = 2\mathbf{1}^\top \Phi (F - G) Z_{\beta}^{(3)} (F - G) \Phi \mathbf{1}.$$

#### 4.4 Resultados de simulação

Apresentamos agora alguns resultados de simulação comparando o desempenho da estatística  $S_R$  do teste escore usual com o da estatística  $S_R^*$  do teste escore corrigido. Consideramos o modelo de

regressão normal não-linear

$$y_\ell = \beta_1 + \exp\{\beta_2 x_{\ell 2}\} + \sum_{i=3}^p \beta_i x_{\ell i} + u_\ell, \ell = 1, \dots, n,$$

em que  $u_\ell \sim N(0, \sigma^2 \exp\{\delta_1 z_{\ell 1} + \dots + \delta_q z_{\ell q}\})$  e  $\text{Cov}(u_\ell, u_m) = 0, \forall \ell \neq m$ . Consideramos as seguintes hipóteses nulas:  $H_0^1 : \delta_1 = \delta_2 = 0$ , para  $p = 2, 3, 4, 5$  e  $H_0^2 : \beta_1 = \beta_2 = 1$ , para  $q = 1, 2, 4, 5$ . Para gerar a variável resposta  $y_\ell$ , fixamos, para testar  $H_0^1$ ,  $\beta_1 = \beta_2 = \dots = \beta_p = 1, \sigma^2 = 1$  e  $\delta_1 = \delta_2 = \dots = \delta_q = 0$ , e, para testar  $H_0^2$ , fixamos  $\beta_1 = \beta_2 = \dots = \beta_p = 1, \sigma^2 = 1$  e  $\delta_1 = \delta_2 = \dots = \delta_q = 1$ . As covariadas  $x_2, \dots, x_p$  foram geradas da distribuição uniforme  $U(0, 1)$ . A matriz  $Z$  foi obtida por meio das colunas  $2, \dots, q$  da matriz  $X$ . Foram geradas 10.000 réplicas para cada tamanho de amostra  $n$  fixado,  $n = 10, 20, \dots, 100$ . Para cada valor de  $n$  e sob cada hipótese nula, calculamos as taxas de rejeição de  $S_R$  e de  $S_R^*$ , isto é, a proporção de vezes (em 10.000) em que  $S_R$  e  $S_R^*$  excederam os quantis  $x_\alpha$  de ordem  $(1 - \alpha)$  da distribuição  $\chi_r^2$ , em que  $r$  é o número de parâmetros testados nas hipóteses  $H_0^1$  e  $H_0^2$ , no caso  $r = 2$ , e  $\alpha = 1\%, 5\%$  e  $10\%$  (nível nominal). Para cada tamanho de amostra e cada nível nominal considerado, verificamos, sob cada hipótese nula, a proximidade das taxas de rejeição aos respectivos níveis nominais fixados. Os resultados encontram-se nas Tabelas 4.2, 4.3, 4.4 e 4.5, todos em porcentagem.

Podemos observar, em geral, que a taxa de rejeição da estatística escore corrigida está mais próxima dos níveis nominais do que a taxa de rejeição da estatística escore usual. A estatística escore usual tende a ser mais conservativa do que a estatística corrigida, ou seja as hipóteses  $H_0^1$  e  $H_0^2$  são rejeitadas bem menos do que seria esperado. Considerando o teste de  $H_0^1$ , há aparentemente uma pequena influência do valor de  $p$  sobre a taxa de rejeição de  $S_R^*$ , visto que para  $n \leq 40$ , esta taxa aumenta conforme  $p$  aumenta, ficando mais distante dos níveis nominais considerados. Esta situação, no entanto, não ocorre no teste de  $H_0^2$ , uma vez que a taxa de rejeição de  $S_R^*$  não parece alterar-se com o valor adotado de  $q$ . Como era esperado, as taxas de rejeição das estatísticas escore

usual e escore corrigida se aproximam uma da outra à medida que o tamanho da amostra cresce.

Tabela 4.2: Tamanho dos testes para a hipótese  $H_0^1$  com  $p = 2, 3$  e  $r = 2$

$n$	$\alpha$	$p = 2,$		$p = 3$	
		$P_r[S_R \geq c]$	$P_r[S_R^* \geq c]$	$P_r[S_R \geq c]$	$P_r[S_R^* \geq c]$
10	10	6,11	11,15	5,39	11,08
	5	2,79	5,88	1,89	5,94
	1	0,30	0,44	0,04	1,03
20	10	7,96	10,33	8,37	10,48
	5	3,42	5,32	3,73	5,39
	1	0,60	0,78	0,57	1,04
30	10	8,17	10,13	8,11	10,25
	5	4,26	5,19	4,06	5,15
	1	0,85	1,08	0,81	0,97
40	10	8,61	10,07	8,40	10,16
	5	4,31	5,06	4,22	5,17
	1	0,85	0,96	0,90	1,05
50	10	8,92	10,04	8,79	10,04
	5	4,43	5,09	4,36	5,06
	1	1,02	0,91	0,89	0,98
60	10	9,24	10,12	8,79	10,02
	5	4,47	5,05	4,45	5,08
	1	0,89	0,94	0,97	1,02
70	10	8,82	10,52	8,77	10,12
	5	4,48	5,14	4,29	5,16
	1	0,97	0,91	0,96	1,05
80	10	9,47	10,12	9,06	10,08
	5	4,78	5,16	4,68	5,10
	1	1,16	1,03	1,23	1,13
90	10	9,43	10,07	9,26	10,11
	5	4,60	5,02	4,69	5,07
	1	0,82	0,95	0,85	0,97
100	10	9,38	10,03	9,42	10,09
	5	4,72	5,02	4,61	5,04
	1	0,89	0,95	0,09	0,98

Tabela 4.3: Tamanho dos testes para a hipótese  $H_0^1$  com  $p = 4, 5$  e  $r = 2$

$n$	$\alpha$	$p = 4,$		$p = 5$	
		$P_r[S_R \geq c]$	$P_r[S_R^* \geq c]$	$P_r[S_R \geq c]$	$P_r[S_R^* \geq c]$
10	10	4,30	12,23	4,49	12,87
	5	1,24	6,99	1,21	7,62
	1	0,01	1,70	0,01	2,03
20	10	5,96	11,06	8,66	11,30
	5	3,94	5,52	3,81	5,89
	1	0,56	1,32	0,54	1,45
30	10	8,75	10,44	7,87	10,62
	5	4,17	5,12	3,79	5,38
	1	0,75	1,00	0,58	1,20
40	10	8,96	10,34	8,32	10,37
	5	3,75	5,06	4,03	5,17
	1	0,70	1,11	0,82	1,15
50	10	8,49	10,07	9,12	10,14
	5	4,32	5,03	4,35	5,07
	1	0,95	1,08	1,07	1,04
60	10	9,55	10,03	9,66	10,09
	5	4,77	5,06	4,98	5,07
	1	0,89	1,01	0,88	1,04
70	10	9,02	10,02	8,73	10,06
	5	4,35	5,01	4,15	5,03
	1	0,76	0,97	0,80	1,02
80	10	8,90	10,01	9,67	10,02
	5	4,47	5,04	4,84	5,07
	1	1,05	1,02	0,97	1,03
90	10	8,99	10,14	9,10	10,12
	5	4,57	5,08	4,39	5,09
	1	1,01	1,09	0,70	1,00
100	10	9,59	10,08	9,32	10,10
	5	4,79	5,05	4,67	5,07
	1	0,97	1,01	0,89	1,02

Tabela 4.4: Tamanho dos testes para a hipótese  $H_0^2$  com  $q = 1, 2$  e  $r = 2$ 

$n$	$\alpha$	$q = 1,$		$q = 2$	
		$P_r[S_R \geq c]$	$P_r[S_R^* \geq c]$	$P_r[S_R \geq c]$	$P_r[S_R^* \geq c]$
10	10	6,84	10,83	5,99	11,15
	5	1,76	4,77	2,98	5,18
	1	0,00	0,50	0,10	1,43
20	10	9,01	10,09	7,52	10,58
	5	3,63	5,09	2,91	4,83
	1	0,23	0,72	0,22	0,89
30	10	8,82	10,33	7,31	9,88
	5	3,71	4,89	2,86	4,62
	1	0,40	0,89	0,38	0,82
40	10	9,43	10,17	8,51	10,48
	5	4,34	5,29	4,08	5,33
	1	0,54	0,96	0,48	1,05
50	10	8,99	9,81	8,28	9,80
	5	4,42	5,18	3,69	4,66
	1	0,70	1,00	0,57	0,96
60	10	9,38	10,11	8,63	9,53
	5	4,23	4,91	4,18	4,82
	1	0,74	0,96	0,60	0,90
70	10	9,29	9,92	8,96	9,75
	5	4,15	4,81	4,21	4,94
	1	0,75	0,96	0,81	1,06
80	10	9,83	10,35	8,99	9,86
	5	4,65	5,18	4,27	4,95
	1	0,88	1,07	0,83	1,08
90	10	9,60	10,17	8,96	9,74
	5	4,29	4,88	4,02	4,67
	1	0,78	0,99	0,78	1,00
100	10	9,82	10,07	9,66	9,73
	5	4,45	4,92	4,74	4,97
	1	0,87	1,02	0,84	0,93

Tabela 4.5: Tamanho dos testes para a hipótese  $H_0^2$  com  $q = 4, 5$  e  $r = 2$

$n$	$\alpha$	$q = 4,$		$q = 5$	
		$P_r[S_R \geq c]$	$P_r[S_R^* \geq c]$	$P_r[S_R \geq c]$	$P_r[S_R^* \geq c]$
10	10	5,99	12,07	4,60	13,86
	5	1,35	6,09	1,26	7,09
	1	0,00	0,77	0,00	1,13
20	10	8,03	11,17	6,89	11,27
	5	3,26	5,33	2,68	5,89
	1	0,16	1,02	0,16	1,06
30	10	8,35	10,57	7,88	10,49
	5	3,40	5,49	3,28	5,56
	1	0,36	1,32	0,34	1,04
40	10	9,41	10,26	9,17	10,17
	5	4,05	5,25	4,02	5,41
	1	0,43	1,12	0,40	1,16
50	10	9,02	10,13	8,85	10,49
	5	4,03	5,18	3,93	5,23
	1	0,60	1,17	0,56	1,19
60	10	9,37	10,05	8,97	10,13
	5	4,37	5,04	4,00	5,18
	1	0,68	1,12	0,48	1,17
70	10	9,29	10,08	9,82	9,97
	5	4,22	5,09	4,39	5,14
	1	0,52	1,04	0,80	1,08
80	10	9,86	10,00	9,14	9,94
	5	4,54	5,02	4,24	5,13
	1	0,65	0,92	0,55	0,88
90	10	9,32	10,19	9,08	10,14
	5	4,43	5,11	4,01	5,21
	1	0,77	1,01	0,78	1,14
100	10	9,32	10,13	9,25	10,01
	5	4,10	5,08	4,37	5,13
	1	0,88	1,07	0,81	1,12

## 4.5 Aplicação

Na Seção 3.7 introduzimos a elasticidade constante de substituição da função de produção, em que o logaritmo da variável resposta  $y$  tem, por hipótese, variância constante. Vamos considerar os dados da Seção 3.7.1 e testar a hipótese de que a variância é constante para todas as observações. Supomos que o erro aleatório  $\epsilon_\ell$ ,  $\ell = 1, \dots, 30$  tem distribuição normal com média zero e variância  $\sigma^2 \exp\{K\delta_1 + L\delta_2\}$ , em que  $L$  e  $K$  são as variáveis trabalho e capital, respectivamente. A Figura 4.1 mostra o gráfico do quadrado do resíduo  $(y_\ell - \hat{\mu}_\ell)^2$ , em função de cada observação. Este gráfico sugere que a variância não é constante para todas as observações. Nosso objetivo é testar a hipótese  $H_0 : \delta_1 = \delta_2 = 0$  contra a hipótese alternativa  $H_1$  : pelo menos uma das igualdades em  $H_0$  não se verifica. O teste escore usual rejeitou a hipótese de igualdade das variâncias, para o qual obtivemos a estatística de teste  $SR = 12,06$  levando a um nível descritivo de 0,25%. A estatística corrigida também rejeitou a hipótese  $H_0$ , sendo que encontramos  $SR^* = 21,36$  levando a um nível descritivo de 0,01%.

Ambas as estatísticas confirmaram as evidências de que a variância não é constante ao longo das observações. A estatística corrigida aumentou esta evidência.

## 4.6 Conclusões

Por meio do cálculo dos fatores de correção tipo-Bartlett para a estatística escore nos MNLFEs com dispersão variável e desconhecida, obtivemos um melhoramento na aproximação da distribuição da estatística associada aos testes sobre os parâmetros que modelam a média e a dispersão. Pudemos observar através do estudo de simulação que a estatística escore usual é muito conservativa no sentido de rejeitar a hipótese nula menos do que deveria. O tamanho empírico do teste escore corrigido ficou bem mais próximo do valor nominal, além disso, pudemos notar que a estatística escore usual sofreu mais influência do número de parâmetros de perturbação do que a estatística escore corrigida.

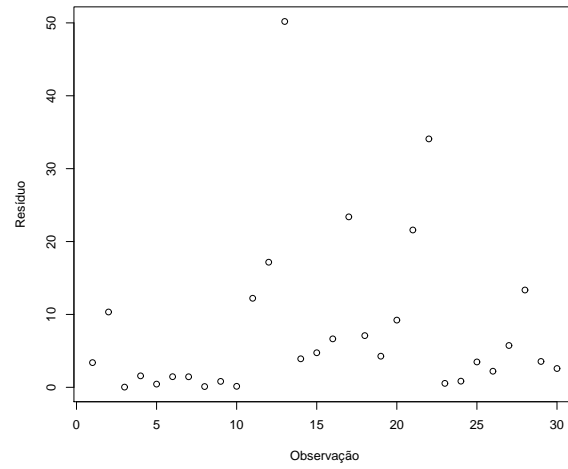


Figura 4.1: Gráfico do quadrado do resíduo em função da  $\ell$ -ésima observação



## Capítulo 5

### Pesquisas futuras

Basicamente, os resultados principais desta tese estão resumidos em três pontos:

- Desenvolvimento da matriz de covariância de ordem  $n^{-2}$  do EMV de  $\beta$  corrigido pelo viés de ordem  $n^{-1}$  para os MLGs com dispersão conhecida. Mostramos que é possível melhorar o teste de Wald simplesmente substituindo a estimativa da matriz de informação de Fisher na estatística de teste usual pela estimativa da inversa da matriz assintótica de ordem  $n^{-2}$  do estimador corrigido. Além disso, se substituirmos também o EMV pelo corrigido, os resultados são mais precisos ainda;
- Cálculo do coeficiente de assimetria assintótico de ordem  $n^{-1/2}$  da distribuição dos EMVs de  $\beta$  e dos parâmetros de precisão e dispersão para os MNLFs com dispersão desconhecida, porém a mesma para todas as observações. Obtivemos uma fórmula bastante simples para o cálculo deste coeficiente sendo possível notar que o mesmo é ponderado pelo inverso do quadrado do parâmetro de precisão. Este coeficiente pode servir como uma medida de afastamento da distribuição normal;
- Cálculo dos fatores de correção tipo-Bartlett para a estatística score nos MNLFs com dis-

persão variável e desconhecida. Através desta correção obtivemos um melhoramento na aproximação da estatística associada aos testes sobre os parâmetros que modelam a média e a dispersão.

Vários trabalhos poderão ser desenvolvidos a partir das idéias desta tese. Dentre estes, podemos citar:

1. Extensão do cálculo da matriz de covariância de ordem  $n^{-1}$  do EMV de  $\beta$  corrigido pelo viés de ordem  $n^{-1}$  para o caso em que o coeficiente de dispersão é desconhecido, mas constante sobre todas as observações ou pode ser modelado por meio de covariáveis, na classe dos MLGs;
2. Cálculo da matriz de covariâncias assintótica de ordem  $n^{-2}$  do estimador de  $\beta$  corrigido pelo viés para os MNLFEs com dispersão conhecida ou desconhecida, porém a mesma para todas as observações;
3. Cálculo do coeficiente de curtose assintótico para a distribuição do EMV de  $\beta$  em MNLFEs;
4. Extensão do cálculo do coeficiente de assimetria da distribuição do EMV de  $\beta$  para os estimadores corrigidos pelo viés;
5. Estudo comparativo das estatísticas de Wald modificadas com outras estatísticas como a estatística de razão de verossimilhanças e/ou a estatística score associadas aos testes sobre o parâmetro  $\beta$  em MLGs com dispersão conhecida;
6. Desenvolvimento de rotinas computacionais para o cálculo das quantidades desenvolvidas nesta tese que possam ser disponibilizadas para usuários do programa R.

Esperamos, através dos resultados deste trabalho, ter contribuído de forma significativa para o desenvolvimento de métodos assintóticos na classe dos modelos de regressão da família exponencial.

## Apêndice A

# Obtenção de alguns cumulantes de derivadas do logaritmo da função de verossimilhança total em modelos não-lineares com dispersão variável

Apresentamos aqui a obtenção de alguns cumulantes conjuntos de derivadas do logaritmo da função de verossimilhança de modelos não-lineares da família exponencial com dispersão variável reparametrizados. Estes cumulantes são necessários para o cálculo dos  $A$ 's na obtenção do fator de correção tipo-Bartlett para a estatística escore.

O logaritmo da função de verossimilhança do vetor de parâmetros  $\omega = (\beta^\top, \delta^\top, \gamma)^\top$  para o modelo definido por (4.1) e (4.2) reparametrizado é dado pela expressão (4.4).

Muitos dos cumulantes das derivadas da expressão (4.4) em relação aos parâmetros podem ser encontrados na literatura como, por exemplo, em Ferrari *et al.* (2004). Apresentamos aqui apenas alguns cumulantes que não se encontram na literatura em geral e que foram importantes para o cálculo dos  $A$ 's.

Adotamos a seguinte notação para os cumulantes das derivadas do logaritmo da função de veros-

semilhança total,  $\mathcal{L} = \mathcal{L}(\beta, \delta, \gamma)$ , com respeito aos componentes do vetor  $\beta$ , sendo que todos são de ordem  $n$ :

$$\kappa_{rs} = E(\partial^2 \mathcal{L} / \partial \beta_r \partial \beta_s), \kappa_{r,s} = E(\partial \mathcal{L} / \partial \beta_r \partial \mathcal{L} / \partial \beta_s), \kappa_{rst} = E(\partial^3 \mathcal{L} / \partial \beta_r \partial \beta_s \partial \beta_t),$$

$$\kappa_{r,s,t} = E(\partial \mathcal{L} / \partial \beta_r \partial \mathcal{L} / \partial \beta_s \partial \mathcal{L} / \partial \beta_t), \kappa_{r,st} = E(\partial \mathcal{L} / \partial \beta_r \partial^2 \mathcal{L} / \partial \beta_s \partial \beta_t),$$

$\kappa_{rs}^{(t)} = \partial \kappa_{rs} / \partial \beta_t$ , e assim por diante.

Os cumulantes correspondentes com respeito a  $\delta$  são definidos de maneira similar usando-se os índices  $R, S, T, U, V$ . Cumulantes mistos com respeito a  $\beta, \delta$  e  $\gamma$  são denotados por  $\kappa_{rS\gamma} = E(\partial^3 \mathcal{L} / \partial \beta_r \partial \delta_s \partial \gamma)$  e assim sucessivamente.

Diferenciando o logaritmo da função de verossimilhança total  $\mathcal{L}$ , usando algumas condições de regularidade, a ortogonalidade entre os parâmetros  $\beta, \delta$  e  $\gamma$  e algumas identidades de Bartlett, obtemos os seguintes cumulantes de segunda ordem:

$$\begin{aligned} \kappa_{rs} &= -\frac{1}{\gamma} \sum_{\ell=1}^n x_{\ell r}^* q_{\ell} w_{\ell} x_{\ell s}^*, \\ \kappa_{RS} &= -\frac{1}{2} \sum_{\ell=1}^n \frac{q_{\ell RS}}{q_{\ell}}, \\ \kappa_{\gamma\gamma} &= -\frac{n}{2\gamma^2}, \\ \kappa_{rS} &= \kappa_{r\gamma} = \kappa_{R\gamma} = 0, \end{aligned}$$

em que  $x_{\ell r}^* = \partial \eta_{\ell} / \partial \beta^r$  e  $q_{\ell RS} = \partial^2 q_{\ell} / \partial \delta^R \partial \delta^S$ , para  $r = 1, \dots, p$  e  $R, S = 1, \dots, q$ .

Os cumulantes de terceira ordem são dados por:

$$\begin{aligned}
\kappa_{rst} &= -\frac{1}{\gamma} \sum_{\ell=1}^n \{q_{\ell}(f_{\ell} + 2g_{\ell})x_{\ell r}^* x_{\ell s}^* x_{\ell t}^* + q_{\ell} w_{\ell}(x_{\ell r s}^{**} x_{\ell t}^* + x_{\ell r t}^{**} x_{\ell s}^* + x_{\ell s t}^{**} x_{\ell r}^*)\}, \\
\kappa_{RST} &= -\frac{1}{2} \sum_{\ell=1}^n \frac{q_{\ell RST}}{q_{\ell}}, \\
\kappa_{\gamma\gamma\gamma} &= \frac{2n}{\gamma^3}, \\
\kappa_{rsT} &= -\frac{1}{\gamma} \sum_{\ell=1}^n x_{\ell r}^* q_{\ell T} w_{\ell} x_{\ell s}^*, \\
\kappa_{rs\gamma} &= \frac{1}{\gamma^2} \sum_{\ell=1}^n x_{\ell r}^* q_{\ell} w_{\ell} x_{\ell s}^*, \\
\kappa_{RS\gamma} &= \frac{1}{2\gamma} \sum_{\ell=1}^n \frac{q_{\ell RS}}{q_{\ell}}, \\
\kappa_{R\gamma\gamma} &= \kappa_{r\gamma\gamma} = \kappa_{RSt} = \kappa_{rS\gamma} = 0, \\
\kappa_{r,s,t} &= \frac{1}{\gamma} \sum_{\ell=1}^n q_{\ell}(f_{\ell} - g_{\ell})x_{\ell r}^* x_{\ell s}^* x_{\ell t}^*, \\
\kappa_{R,S,T} &= \frac{1}{2} \sum_{\ell=1}^n \left\{ \frac{q_{\ell RST}}{q_{\ell}} - \frac{q_{\ell RS} q_{\ell T}}{q_{\ell}^2} - \frac{q_{\ell RT} q_{\ell S}}{q_{\ell}^2} - \frac{q_{\ell ST} q_{\ell R}}{q_{\ell}^2} \right\}, \\
\kappa_{\gamma,\gamma,\gamma} &= \frac{n}{\gamma^3}, \\
\kappa_{r,s,T} &= -\frac{1}{\gamma} \sum_{\ell=1}^n q_{\ell T} x_{\ell r}^* x_{\ell s}^* w_{\ell}, \\
\kappa_{r,s,\gamma} &= \frac{1}{\gamma^2} \sum_{\ell=1}^n q_{\ell} x_{\ell r}^* x_{\ell s}^* w_{\ell}, \\
\kappa_{R,S,\gamma} &= \frac{1}{\gamma} \sum_{\ell=1}^n \frac{q_{\ell RS}}{q_{\ell}}, \\
\kappa_{r,S,\gamma} &= \kappa_{r,\gamma,\gamma} = \kappa_{r,S,T} = \kappa_{R,\gamma,\gamma} = 0, \\
\kappa_{r,st} &= \frac{1}{\gamma} \sum_{\ell=1}^n q_{\ell}(g_{\ell} x_{\ell r}^* x_{\ell s}^* x_{\ell t}^* + w_{\ell} x_{\ell r}^* x_{\ell s}^{**}),
\end{aligned}$$

$$\begin{aligned}
\kappa_{R,ST} &= \frac{1}{2} \sum_{\ell=1}^n \frac{q_{\ell R} q_{\ell ST}}{q_{\ell}^2}, \\
\kappa_{\gamma,\gamma\gamma} &= -\frac{n}{\gamma^3}, \\
\kappa_{r,sT} &= \frac{1}{\gamma} \sum_{\ell=1}^n x_{\ell r}^* q_{\ell T} w_{\ell} x_{\ell s}^*, \\
\kappa_{r,s\gamma} &= -\frac{1}{\gamma^2} \sum_{\ell=1}^n q_{\ell} x_{\ell r}^* x_{\ell s}^* w_{\ell}, \\
\kappa_{R,S\gamma} &= -\frac{1}{2\gamma} \sum_{\ell=1}^n \frac{q_{\ell R} q_{\ell S}}{q_{\ell}^2}, \\
\kappa_{\gamma,RS} &= -\frac{1}{2\gamma} \sum_{\ell=1}^n \frac{q_{\ell RS}}{q_{\ell}}, \\
\kappa_{r,ST} &= \kappa_{r,S\gamma} = \kappa_{r,\gamma\gamma} = \kappa_{R,st} = \kappa_{R,St} = \kappa_{R,s\gamma} = \kappa_{R,\gamma\gamma} = \kappa_{\gamma,rs} = \kappa_{\gamma,rS} = \kappa_{\gamma,\gamma r} = \kappa_{\gamma,\gamma R} = 0,
\end{aligned}$$

em que  $x_{\ell rs}^{**} = \partial^2 \eta_{\ell} / \partial \beta^r \partial \beta^s$ ,  $q_{\ell RST} = \partial^3 q_{\ell} / \partial \delta^R \partial \delta^S \partial \delta^T$ , para  $r, s, t = 1, \dots, p$  e  $R, S, T = 1, \dots, q$  e as quantidades  $f_{\ell}$  e  $g_{\ell}$ ,  $\ell = 1, \dots, n$ , são definidas pelas expressões (4.7) e (4.8), respectivamente.

Finalmente, os cumulantes de quarta ordem são dados por:

$$\begin{aligned}
\kappa_{rstu} &= -\frac{1}{\gamma} \sum_{\ell=1}^n q_{\ell} \left\{ \left[ \left( \frac{\partial^2 w}{\partial \eta^2} \right) + \left( \frac{\partial^2 \mu}{\partial \eta^2} \right) \left( \frac{\partial^2 \theta}{\partial \eta^2} \right) + 2 \left( \frac{\partial \mu}{\partial \eta} \right) \left( \frac{\partial^3 \theta}{\partial \eta^3} \right) \right]_{\ell} x_{\ell r}^* x_{\ell s}^* x_{\ell t}^* x_{\ell u}^* \right. \\
&\quad + w_{\ell} (x_{\ell rst}^{***} x_{\ell u}^* + x_{\ell rsu}^{***} x_{\ell t}^* + x_{\ell rut}^{***} x_{\ell s}^* + x_{\ell ust}^{***} x_{\ell r}^* + x_{\ell rs}^{**} x_{\ell tu}^{**} + x_{\ell rt}^{**} x_{\ell su}^{**} + x_{\ell ru}^{**} x_{\ell st}^{**}) \\
&\quad + (x_{\ell rs}^{**} x_{\ell t}^* x_{\ell u}^* + x_{\ell ts}^{**} x_{\ell r}^* x_{\ell u}^* + x_{\ell us}^{**} x_{\ell r}^* x_{\ell t}^* + x_{\ell ru}^{**} x_{\ell s}^* x_{\ell t}^* + x_{\ell tu}^{**} x_{\ell r}^* x_{\ell s}^* + x_{\ell rt}^{**} x_{\ell s}^* x_{\ell u}^*) \\
&\quad \left. \times \left[ 2 \left( \frac{\partial \mu}{\partial \eta} \right) \left( \frac{\partial^2 \theta}{\partial \eta^2} \right) + \left( \frac{\partial \theta}{\partial \eta} \right) \left( \frac{\partial^2 \mu}{\partial \eta^2} \right) \right]_{\ell} \right\}, \\
\kappa_{RSTU} &= -\frac{1}{2} \sum_{\ell=1}^n \frac{q_{\ell RSTU}}{q_{\ell}},
\end{aligned}$$

$$\begin{aligned}
\kappa_{RS\gamma\gamma} &= -\frac{1}{\gamma^2} \sum_{\ell=1}^n \frac{q_{\ell RS}}{q_{\ell}} \\
\kappa_{r,s,t,u} &= \frac{1}{\gamma} \sum_{\ell=1}^n q_{\ell} b_{\ell} x_{\ell_r}^* x_{\ell_s}^* x_{\ell_t}^* x_{\ell_u}^*, \\
\kappa_{R,S,T,U} &= \frac{1}{2} \sum_{\ell=1}^n \left( \frac{q_{\ell RSTU}}{q_{\ell}} - \frac{q_{\ell RST} q_{\ell U}}{q_{\ell}^2} - \frac{q_{\ell RSU} q_{\ell T}}{q_{\ell}^2} - \frac{q_{\ell RTU} q_{\ell S}}{q_{\ell}^2} - \frac{q_{\ell STU} q_{\ell R}}{q_{\ell}^2} \right. \\
&\quad + \frac{2q_{\ell R} q_{\ell S} q_{\ell TU}}{q_{\ell}^3} + \frac{2q_{\ell R} q_{\ell T} q_{\ell SU}}{q_{\ell}^3} + \frac{2q_{\ell R} q_{\ell U} q_{\ell ST}}{q_{\ell}^3} + \frac{2q_{\ell S} q_{\ell T} q_{\ell RU}}{q_{\ell}^3} \\
&\quad \left. + \frac{2q_{\ell S} q_{\ell U} q_{\ell RT}}{q_{\ell}^3} + \frac{2q_{\ell T} q_{\ell U} q_{\ell RS}}{q_{\ell}^3} - \frac{q_{\ell RS} q_{\ell TU}}{q_{\ell}^2} - \frac{q_{\ell RT} q_{\ell SU}}{q_{\ell}^2} - \frac{q_{\ell RU} q_{\ell ST}}{q_{\ell}^2} \right), \\
\kappa_{r,s,T,U} &= \frac{2}{\gamma} \sum_{\ell=1}^n q_{\ell} w_{\ell} x_{\ell_r}^* x_{\ell_s}^* \frac{q_{\ell T} q_{\ell U}}{q_{\ell}^2}, \\
\kappa_{r,s,\gamma,\gamma} &= \frac{2}{\gamma^3} \sum_{\ell=1}^n q_{\ell} w_{\ell} x_{\ell_r}^* x_{\ell_s}^*, \\
\kappa_{R,S,\gamma,\gamma} &= \frac{3}{\gamma^2} \sum_{\ell=1}^n \frac{q_{\ell RS}}{q_{\ell}}, \\
\kappa_{r,s,tu} &= \frac{1}{\gamma} \sum_{\ell=1}^n q_{\ell} \{ (h_{\ell} - b_{\ell}) x_{\ell_r}^* x_{\ell_s}^* x_{\ell_t}^* x_{\ell_u}^* + (f_{\ell} - g_{\ell}) x_{\ell_r}^* x_{\ell_s}^* x_{\ell_{tu}}^{**} \}, \\
\kappa_{R,S,TU} &= -\sum_{\ell=1}^n \frac{q_{\ell R} q_{\ell S} q_{\ell TU}}{q_{\ell}^3}, \\
\kappa_{r,s,TU} &= -\frac{1}{\gamma} \sum_{\ell=1}^n w_{\ell} x_{\ell_r}^* x_{\ell_s}^* q_{\ell TU}, \\
\kappa_{r,s,\gamma\gamma} &= -\frac{2}{\gamma^3} \sum_{\ell=1}^n q_{\ell} w_{\ell} x_{\ell_r}^* x_{\ell_s}^*, \\
\kappa_{R,S,\gamma\gamma} &= -\frac{2}{\gamma^2} \sum_{\ell=1}^n \frac{q_{\ell RS}}{q_{\ell}},
\end{aligned}$$

em que  $x_{\ell rst}^{***} = \partial^3 \eta_{\ell} / \partial \beta^r \partial \beta^s \beta^t$ ,  $q_{\ell RSTU} = \partial^4 q_{\ell} / \partial \delta^R \partial \delta^S \partial \delta^T \partial \delta^U$ , para  $r, s, t, u = 1, \dots, p$  e  $R, S, T, U = 1, \dots, q$  e a quantidade  $b_{\ell}$  é definida pela expressão (4.9).

Considerando o caso em que  $m_\ell = \exp\{z_\ell^\top \delta\}$ , obtemos  $q_\ell = \exp\{-(z_\ell - \bar{z})^\top \delta\}$ ,  $q_{\ell R} = -(z_\ell - \bar{z})_R \exp\{-(z_\ell - \bar{z})^\top \delta\} = -(z_\ell - \bar{z})_R q_\ell$ ,  $q_{\ell RS} = (z_\ell - \bar{z})_{RS} q_\ell$ ,  $q_{\ell RST} = -(z_\ell - \bar{z})_{RST} q_\ell$  e  $q_{\ell RSTU} = (z_\ell - \bar{z})_{RSTU} q_\ell$  com  $(z_\ell - \bar{z})_R = z_{\ell R} - \bar{z}_R$ ,  $(z_\ell - \bar{z})_{RS} = (z_{\ell R} - \bar{z}_R)(z_{\ell S} - \bar{z}_S)$  e assim por diante, em que  $\bar{z} = (\bar{z}_1, \dots, \bar{z}_q)^\top$  e  $\bar{z}_R = (1/n) \sum_{\ell=1}^n z_{\ell R}$ , para  $R = 1, \dots, q$ .

Desta forma, os cumulantes das derivadas do logaritmo da função de verossimilhança em relação ao parâmetro  $\delta$  são dados por

$$\begin{aligned} \kappa_{RS} &= -\frac{1}{2} \sum_{\ell=1}^n (z_\ell - \bar{z})_{RS} \\ \kappa_{RST} &= \frac{1}{2} \sum_{\ell=1}^n (z_\ell - \bar{z})_{RST}, \\ \kappa_{rsT} &= \frac{1}{\gamma} \sum_{\ell=1}^n x_{\ell r}^* w_\ell x_{\ell s}^* (z_\ell - \bar{z})_T q_\ell, \\ \kappa_{RS\gamma} &= \frac{1}{2\gamma} \sum_{\ell=1}^n (z_\ell - \bar{z})_{RS}, \\ \kappa_{R,S,T} &= \sum_{\ell=1}^n (z_\ell - \bar{z})_{RST}, \\ \kappa_{r,s,T} &= \frac{1}{\gamma} \sum_{\ell=1}^n (z_\ell - \bar{z})_R q_\ell x_{\ell r}^* x_{\ell s}^* w_\ell, \\ \kappa_{R,S,\gamma} &= \frac{1}{\gamma} \sum_{\ell=1}^n (z_\ell - \bar{z})_{RS}, \\ \kappa_{r,sT} &= -\frac{1}{\gamma} \sum_{\ell=1}^n x_{\ell r}^* w_\ell x_{\ell s}^* (z_\ell - \bar{z})_T q_\ell, \\ \kappa_{R,ST} &= -\frac{1}{2} \sum_{\ell=1}^n (z_\ell - \bar{z})_{RST}, \\ \kappa_{R,S\gamma} &= -\frac{1}{2\gamma} \sum_{\ell=1}^n (z_\ell - \bar{z})_{RS}, \end{aligned}$$



$$\begin{aligned}
\kappa_{\gamma,RS} &= -\frac{1}{2\gamma} \sum_{\ell=1}^n (z_{\ell} - \bar{z})_{RS}, \\
\kappa_{\gamma,RS} &= -\frac{1}{2\gamma} \sum_{\ell=1}^n (z_{\ell} - \bar{z})_{RS}, \\
\kappa_{RSTU} &= -\frac{1}{2} \sum_{\ell=1}^n (z_{\ell} - \bar{z})_{RSTU}, \\
\kappa_{RS\gamma\gamma} &= -\frac{1}{\gamma^2} \sum_{\ell=1}^n (z_{\ell} - \bar{z})_{RS}, \\
\kappa_{R,S,T,U} &= 3 \sum_{\ell=1}^n (z_{\ell} - \bar{z})_{RSTU}, \\
\kappa_{r,s,T,U} &= \frac{2}{\gamma} \sum_{\ell=1}^n q_{\ell} w_{\ell} x_{\ell_r}^* x_{\ell_s}^* (z_{\ell} - \bar{z})_{TU}, \\
\kappa_{R,S,\gamma,\gamma} &= \frac{3}{\gamma^2} \sum_{\ell=1}^n (z_{\ell} - \bar{z})_{RS}, \\
\kappa_{R,S,TU} &= -\sum_{\ell=1}^n (z_{\ell} - \bar{z})_{RSTU}, \\
\kappa_{r,s,TU} &= -\frac{1}{\gamma} \sum_{\ell=1}^n w_{\ell} x_{\ell_r}^* x_{\ell_s}^* (z_{\ell} - \bar{z})_{TU} q_{\ell}, \\
\kappa_{R,S,\gamma\gamma} &= -\frac{2}{\gamma^2} \sum_{\ell=1}^n (z_{\ell} - \bar{z})_{RS}.
\end{aligned}$$



## Apêndice B

### Obtenção dos coeficientes $A_1$ , $A_2$ e $A_3$ para a correção tipo-Bartlett em modelos não-lineares da família exponencial com dispersão variável

Apresentamos aqui a obtenção das quantidades  $A_1$ ,  $A_2$  e  $A_3$  referentes ao teste de hipóteses  $H_0 : \beta_1 = \beta_1^{(0)}$ ,  $\delta_1 = \delta_1^{(0)}$  versus  $H_1$  : pelo menos uma das igualdades da hipótese  $H_0$  não se verifica, em MNLFs com dispersão variável (vide Capítulo 4). Segundo Harris (1985), os coeficientes  $A_1$ ,  $A_2$  e  $A_3$  são dados pelas expressões:

$$\begin{aligned} A_1 &= 3 \sum (\kappa_{ijk} + 2\kappa_{i,jk})(\kappa_{rst} + 2\kappa_{rs,t})a_{ij}a_{st}m_{kr} \\ &- 6 \sum (\kappa_{ijk} + 2\kappa_{i,jk})\kappa_{r,s,t}a_{ij}a_{kr}m_{st} \\ &+ 6 \sum (\kappa_{i,jk} - \kappa_{i,j,k})(\kappa_{rst} + 2\kappa_{rs,t})a_{js}a_{kt}m_{ir} \\ &- 6 \sum (\kappa_{i,j,k,r} + \kappa_{i,j,kr})a_{kr}m_{ij}, \end{aligned}$$

(B.1)

$$\begin{aligned}
A_2 &= -3 \sum (\kappa_{i,j,k} \kappa_{r,s,t}) a_{kr} m_{ij} m_{st} \\
&+ 6 \sum (\kappa_{ijk} + 2\kappa_{i,jk}) \kappa_{r,s,t} a_{ij} m_{kr} m_{st} \\
&- 6 \sum \kappa_{i,j,k} \kappa_{r,s,t} a_{kt} m_{ir} m_{js} \\
&+ 3 \sum \kappa_{i,j,k,r} m_{ij} m_{kr} \\
&e
\end{aligned} \tag{B.2}$$

$$\begin{aligned}
A_3 &= 3 \sum \kappa_{i,j,k} \kappa_{r,s,t} m_{ij} m_{kr} m_{st} \\
&+ 2 \sum \kappa_{i,j,k} \kappa_{r,s,t} m_{ir} m_{js} m_{kt},
\end{aligned} \tag{B.3}$$

sendo  $a_{ij}$  e  $m_{ij}$  os elementos  $(i, j)$  das matrizes

$$A = \begin{pmatrix} 0 & 0 \\ 0 & K_{22}^{-1} \end{pmatrix} \text{ e } M = K^{-1} - A, \tag{B.4}$$

respectivamente. Os somatórios nas expressões (B.1)-(B.3) variam sobre todos os parâmetros.

Os coeficientes  $A_1$ ,  $A_2$  e  $A_3$  referentes ao teste  $H_0 : \beta_1 = \beta_1^{(0)}, \delta_1 = \delta_1^{(0)}$  em MNLFs podem ser obtidos através das expressões (B.1)-(B.3). Para facilitar os cálculos, escrevemos  $A_1 = A_{11} + A_{12} + A_{13} + A_{14}$ ,  $A_2 = A_{21} + A_{22} + A_{23} + A_{24}$  e  $A_3 = A_{31} + A_{32}$ , em que

$$A_{11} = 3 \sum (\kappa_{ijk} + 2\kappa_{i,jk}) (\kappa_{rst} + 2\kappa_{r,s,t}) a_{ij} a_{st} m_{kr},$$

$$A_{12} = -6 \sum (\kappa_{ijk} + 2\kappa_{i,jk}) \kappa_{r,s,t} a_{ij} a_{kr} m_{st},$$

$$A_{13} = 6 \sum (\kappa_{i,jk} - 2\kappa_{i,j,k})(\kappa_{rst} + 2\kappa_{r,s,t})a_{js}a_{kt}m_{ir},$$

$$A_{14} = -6 \sum (\kappa_{i,j,k,r} + \kappa_{i,j,kr})a_{kr}m_{ij},$$

$$A_{21} = -3 \sum (\kappa_{i,j,k}\kappa_{r,s,t})a_{kr}m_{ij}m_{st},$$

$$A_{22} = 6 \sum (\kappa_{ijk} + 2\kappa_{i,j,k})\kappa_{r,s,t}a_{ij}m_{kr}m_{st},$$

$$A_{23} = -6 \sum \kappa_{i,j,k}\kappa_{r,s,t}a_{kt}m_{ir}m_{js},$$

$$A_{24} = 3 \sum \kappa_{i,j,k,r}m_{ij}m_{kr},$$

$$A_{31} = 3 \sum \kappa_{i,j,k}\kappa_{r,s,t}m_{ij}m_{kr}m_{st}$$

e

$$A_{32} = 2 \sum \kappa_{i,j,k}\kappa_{r,s,t}m_{ir}m_{js}m_{kt}.$$

A partição  $\beta = (\beta_1^\top, \beta_2^\top)^\top$  e  $\delta = (\delta_1^\top, \delta_2^\top)^\top$  conduz às seguintes partições

$$K_\beta = \begin{pmatrix} K_{\beta_{11}} & K_{\beta_{12}} \\ K_{\beta_{21}} & K_{\beta_{22}} \end{pmatrix}, \quad K_\beta^{-1} = \begin{pmatrix} K_\beta^{11} & K_\beta^{12} \\ K_\beta^{21} & K_\beta^{22} \end{pmatrix},$$

$$K_\delta = \begin{pmatrix} K_{\delta_{11}} & K_{\delta_{12}} \\ K_{\delta_{21}} & K_{\delta_{22}} \end{pmatrix}, \quad K_\delta^{-1} = \begin{pmatrix} K_\delta^{11} & K_\delta^{12} \\ K_\delta^{21} & K_\delta^{22} \end{pmatrix},$$

em que  $K_\beta = E\{U(\beta)U^\top(\beta)\}$  e  $K_\delta = E\{U(\delta)U^\top(\delta)\}$  são as matrizes de informação de Fisher de  $\beta$  e  $\delta$  respectivamente,  $K_\beta^{-1}$  e  $K_\delta^{-1}$  são suas inversas.

Consideramos as matrizes  $A$  e  $M$  particionadas da seguinte forma:

$$A = \begin{pmatrix} A_\beta & 0 & 0 \\ 0 & A_\delta & 0 \\ 0 & 0 & A_\gamma \end{pmatrix} \text{ e } M = \begin{pmatrix} M_\beta & 0 & 0 \\ 0 & M_\delta & 0 \\ 0 & 0 & M_\gamma \end{pmatrix},$$

em que

$$A_\beta = \begin{pmatrix} 0 & 0 \\ 0 & K_{\beta 22}^{-1} \end{pmatrix},$$

$$A_\delta = \begin{pmatrix} 0 & 0 \\ 0 & K_{\delta 22}^{-1} \end{pmatrix}$$

e

$$A_\gamma = k_{\gamma,\gamma}^{-1}.$$

Portanto,  $M_\beta = K_\beta^{-1} - A_\beta$ ,  $M_\delta = K_\delta^{-1} - A_\delta$  e  $M_\gamma = K_\gamma^{-1} - A_\gamma = 0$ .

Assim, devido à ortogonalidade global de  $\beta$ ,  $\delta$  e  $\gamma$  e o fato de  $\kappa_{IJk} = \kappa_{I,Jk} = \kappa_{R\gamma\gamma} = \kappa_{R\gamma,\gamma} = \kappa_{r\gamma\gamma} = \kappa_{r\gamma,\gamma} = \kappa_{i,J,K} = 0$ , obtemos

$$\begin{aligned} A_{11} &= 3 \sum \{ (\kappa_{ijk} + 2\kappa_{i,jk})(\kappa_{rst} + 2\kappa_{rs,t})a_{ij}a_{kr}m_{st} + (\kappa_{ijk} + 2\kappa_{i,jk})(\kappa_{rST} + 2\kappa_{RS,t})a_{ij}a_{kr}m_{ST} \\ &+ (\kappa_{ijK} + 2\kappa_{i,jK})(\kappa_{RST} + 2\kappa_{RS,T})a_{ij}a_{KR}m_{ST} + (\kappa_{IJK} + 2\kappa_{I,Jk})(\kappa_{rST} + 2\kappa_{r,S,T})a_{IJ}a_{kr}m_{ST} \\ &+ (\kappa_{IJK} + 2\kappa_{I,JK})(\kappa_{RST} + 2\kappa_{RS,T})a_{IJ}a_{KR}m_{ST} \}. \end{aligned} \tag{B.5}$$

Substituindo os respectivos cumulantes obtidos no Apêndice A na expressão (B.5), invertendo a ordem das somas e rearranjando os termos, obtemos

$$\begin{aligned}
A_{11} = & 3 \sum_{\ell,u=1}^n \left\{ \sum \frac{q_\ell}{\gamma} x_{\ell_i}^* a_{ij} x_{j_\ell}^* f_\ell x_{\ell_s}^* m_{st} x_{t_u}^* w_u \frac{q_u}{\gamma} x_{u_k}^* a_{kr} x_{r_u}^* f_u + \sum \frac{q_\ell}{\gamma} x_{\ell_i}^* a_{ij} x_{j_\ell}^* f_\ell x_{\ell_s}^* m_{st} x_{t_u}^* w_u \frac{q_u}{\gamma} x_{u_{kr}}^{**} a_{rk} \right. \\
& + \left. \sum \frac{q_u}{\gamma} f_u x_{u_k}^* a_{kr} x_{r_u}^* x_{u_t}^* m_{ts} x_{s_\ell}^* w_\ell \frac{q_\ell}{\gamma} x_{\ell_{ij}}^{**} a_{ji} + \sum \frac{q_u}{\gamma} w_u x_{u_{kr}}^{**} a_{rk} x_{u_t}^* m_{ts} x_{s_\ell}^* w_\ell \frac{q_\ell}{\gamma} x_{\ell_{ij}}^{**} a_{ji} \right\} \\
& + 3 \sum_{\ell,u=1}^n \left\{ \sum \frac{q_\ell}{\gamma} w_\ell x_{\ell_i}^* a_{ij} x_{j_\ell}^* (z_{\ell_S} - \bar{z}_S) m_{ST} (z_{T_u} - \bar{z}_T) x_{u_k}^* a_{kr} x_{r_u}^* w_u \frac{q_u}{\gamma} \right\} \\
& + 3 \sum_{\ell,u=1}^n \left\{ \sum \frac{q_\ell}{2\gamma} w_\ell x_{\ell_i}^* a_{ij} x_{j_\ell}^* (z_{\ell_S} - \bar{z}_S) m_{ST} (z_{T_u} - \bar{z}_T) (z_{u_K} - \bar{z}_K) a_{KR} (z_{R_u} - \bar{z}_R) \right\} \\
& + 3 \sum_{\ell,u=1}^n \left\{ \sum \frac{1}{2} (z_{\ell_I} - \bar{z}_I) a_{IJ} (z_{J_\ell} - \bar{z}_J) (z_{\ell_S} - \bar{z}_S) m_{ST} (z_{T_u} - \bar{z}_T) x_{u_k}^* a_{kr} x_{r_u}^* w_u \frac{q_u}{\gamma} \right\} \\
& + \left. 3 \sum_{\ell,u=1}^n \left\{ \sum \frac{1}{4} (z_{\ell_I} - \bar{z}_I) a_{IJ} (z_{J_\ell} - \bar{z}_J) (z_{\ell_S} - \bar{z}_S) m_{ST} (z_{T_u} - \bar{z}_T) (z_{u_K} - \bar{z}_K) a_{KR} (z_{R_u} - \bar{z}_R) \right\} \right\}
\end{aligned} \tag{B.6}$$

Sejam as matrizes  $Z_\beta$ ,  $Z_{\beta_2}$ ,  $P_2$ ,  $Z_\delta$ ,  $Z_{\delta_2}$ ,  $F$ ,  $W$  e  $\Phi$  definidas de acordo com a Seção 4.3. Os termos da forma  $\sum x_{\ell_i}^* m_{ij} x_{j_\ell}^*$ ,  $\sum x_{\ell_i}^* a_{ij} x_{j_\ell}^*$ ,  $\sum (z_{\ell_I} - \bar{z}_I) m_{IJ} (z_{J_\ell} - \bar{z}_J)$  e  $\sum (z_{\ell_I} - \bar{z}_I) a_{IJ} (z_{J_\ell} - \bar{z}_J)$  são os elementos  $(\ell, u)$  das matrizes  $Z_\beta - Z_{\beta_2}$ ,  $Z_{\beta_2}$ ,  $Z_\delta - Z_{\delta_2}$  e  $Z_{\delta_2}$ , respectivamente. Assim,

$$\begin{aligned}
\sum \frac{q_\ell}{\gamma} x_{\ell_i}^* a_{ij} x_{j_\ell}^* f_\ell x_{\ell_s}^* m_{st} x_{t_u}^* w_u \frac{q_u}{\gamma} x_{u_k}^* a_{kr} x_{r_u}^* f_u &= \phi_\ell z_{\beta 2 \ell \ell} f_\ell (z_{\beta \ell u} - z_{\beta 2 \ell u}) z_{\beta 2 u} w_u \phi_u, \\
\sum \frac{q_u}{\gamma} f_u x_{u_k}^* a_{kr} x_{r_u}^* x_{u_t}^* m_{ts} x_{s_\ell}^* w_\ell \frac{q_\ell}{\gamma} x_{\ell_{ij}}^{**} a_{ji} &= \phi_\ell w_\ell p_{2\ell} (z_{\beta \ell u} - z_{\beta 2 \ell u}) z_{\beta 2 u u} f_u \phi_u, \\
\sum \frac{q_\ell}{\gamma} w_\ell x_{\ell_i}^* a_{ij} x_{j_\ell}^* (z_{\ell_S} - \bar{z}_S) m_{ST} (z_{T_u} - \bar{z}_T) x_{u_k}^* a_{kr} x_{r_u}^* w_u \frac{q_u}{\gamma} &= \phi_\ell w_\ell z_{\beta 2 \ell \ell} (z_{\delta \ell u} - z_{\delta 2 \ell u}) z_{\beta 2 u u} w_u \phi_u
\end{aligned}$$

e

$$\sum \frac{1}{4} (z_{\ell_I} - \bar{z}_I) a_{IJ} (z_{J_\ell} - \bar{z}_J) (z_{\ell_S} - \bar{z}_S) m_{ST} (z_{T_u} - \bar{z}_T) (z_{u_K} - \bar{z}_K) a_{KR} (z_{R_u} - \bar{z}_R) = \frac{1}{4} z_{\delta 2 \ell \ell} (z_{\delta \ell u} - z_{\delta 2 \ell u}) z_{\delta 2 u u},$$

em que  $\phi_\ell = q_\ell/\gamma$  e  $p_{2\ell} = \text{tr}\{X_2^*(X_2^{*\top}W\Phi X_2^*)^{-1}D_{22}^{(\ell)}\}$ , sendo

$$D^{(\ell)} = \left\{ \frac{\partial^2 \eta_\ell}{\partial \beta_i \partial \beta_j} \right\} = \begin{pmatrix} D_{11}^{(\ell)} & D_{12}^{(\ell)} \\ D_{21}^{(\ell)} & D_{22}^{(\ell)} \end{pmatrix}.$$

Assim,  $A_{11}$  se reduz a

$$\begin{aligned} A_{11} &= 3 \sum_{\ell,u=1}^n \left\{ \phi_\ell z_{\beta_{2\ell\ell}} f_\ell (z_{\beta_{\ell u}} - z_{\beta_{2\ell u}}) z_{\beta_{2uu}} w_u \phi_u + \phi_u w_u p_{2u} (z_{\beta_{u\ell}} - z_{\beta_{2u\ell}}) z_{\beta_{2\ell\ell}} f_\ell \phi_\ell \right. \\ &+ \left. \phi_\ell w_\ell p_{2\ell} (z_{\beta_{\ell u}} - z_{\beta_{2\ell u}}) z_{\beta_{2uu}} f_u \phi_u + \phi_\ell w_\ell p_{2\ell} (z_{\beta_{\ell u}} - z_{\beta_{2\ell u}}) p_{2u} w_u \phi_u \right\} \\ &+ 3 \sum_{\ell,u=1}^n \left\{ \phi_\ell w_\ell z_{\beta_{2\ell\ell}} (z_{\delta_{\ell u}} - z_{\delta_{2\ell u}}) z_{\beta_{2uu}} w_u \phi_u \right\} \\ &+ 3 \sum_{\ell,u=1}^n \left\{ \frac{1}{2} \phi_\ell w_\ell z_{\beta_{2\ell\ell}} (z_{\delta_{\ell u}} - z_{\delta_{2\ell u}}) z_{\delta_{2uu}} \right\} \\ &+ 3 \sum_{\ell,u=1}^n \left\{ \frac{1}{2} \phi_u w_u z_{\beta_{2uu}} (z_{\delta_{u\ell}} - z_{\delta_{2u\ell}}) z_{\delta_{2\ell\ell}} \right\} \\ &+ 3 \sum_{\ell,u=1}^n \left\{ \frac{1}{4} z_{\delta_{2\ell\ell}} (z_{\delta_{\ell u}} - z_{\delta_{2\ell u}}) z_{\delta_{2uu}} \right\} \end{aligned}$$

e pode ser escrita em notação matricial como

$$\begin{aligned} A_{11} &= 3\mathbf{1}^\top \Phi F Z_{\beta_{2d}} (Z_\beta - Z_{\beta_2}) Z_{\beta_{2d}} F \Phi \mathbf{1} + 6\mathbf{1}^\top \Phi W P_2 (Z_\beta - Z_{\beta_2}) Z_{\beta_{2d}} F \Phi \mathbf{1} \\ &+ 3\mathbf{1}^\top \Phi W P_2 (Z_\beta - Z_{\beta_2}) P_2 W \Phi \mathbf{1} + 3\mathbf{1}^\top \Phi W Z_{\beta_{2d}} (Z_\delta - Z_{\delta_2}) Z_{\delta_{2d}} \mathbf{1} \\ &+ 3\mathbf{1}^\top \Phi W Z_{\beta_{2d}} (Z_\delta - Z_{\delta_2}) Z_{\beta_{2d}} W \Phi \mathbf{1} + \frac{3}{4} \mathbf{1}^\top Z_{\delta_{2d}} (Z_\delta - Z_{\delta_2}) Z_{\delta_{2d}} \mathbf{1}. \end{aligned}$$



Para os demais coeficientes procedemos de maneira análoga; assim obtemos

$$\begin{aligned}
A_{12} = & -6 \sum \{ (\kappa_{ijk} + 2\kappa_{i,jk}) \kappa_{r,s,t} a_{ij} a_{kr} m_{st} + (\kappa_{ijK} + 2\kappa_{i,jK}) \kappa_{R,s,t} a_{ij} a_{KR} m_{st} \\
& + (\kappa_{ijK} + 2\kappa_{i,jK}) \kappa_{R,S,T} a_{ij} a_{KR} m_{ST} + (\kappa_{ij\gamma} + 2\kappa_{i,j\gamma}) \kappa_{\gamma,s,t} a_{ij} a_{\gamma\gamma} m_{st} \\
& + (\kappa_{ij\gamma} + 2\kappa_{i,j\gamma}) \kappa_{\gamma,S,T} a_{ij} a_{\gamma\gamma} m_{ST} + (\kappa_{IJK} + 2\kappa_{I,JK}) \kappa_{R,s,t} a_{IJ} a_{KR} m_{st} \\
& + (\kappa_{IJK} + 2\kappa_{I,JK}) \kappa_{R,S,T} a_{IJ} a_{KR} m_{ST} + (\kappa_{IJ\gamma} + 2\kappa_{I,J\gamma}) \kappa_{\gamma,s,t} a_{IJ} a_{\gamma\gamma} m_{st} \\
& + (\kappa_{IJ\gamma} + 2\kappa_{I,J\gamma}) \kappa_{\gamma,S,T} a_{IJ} a_{\gamma\gamma} m_{ST} \}, \tag{B.7}
\end{aligned}$$

daí, obtemos

$$\begin{aligned}
A_{12} = & -6 \sum_{\ell,u=1}^n \left\{ \sum \frac{-q\ell}{\gamma} f_{\ell} x_{\ell_i}^* a_{ij} x_{j_{\ell}}^* x_{\ell_k}^* a_{kr} x_{r_u}^* a_{kr} x_{r_u}^* x_{u_s}^* m_{st} x_{t_u}^* (f_u - g_u) \frac{q_u}{\gamma} \right. \\
& - \sum \frac{q\ell}{\gamma} w_{\ell} x_{\ell_i}^* a_{ji} x_{j_{\ell}}^* a_{kr} x_{r_u}^* x_{u_s}^* m_{st} x_{t_u}^* (f_u - g_u) \frac{q_u}{\gamma} \\
& - \sum \frac{q\ell}{\gamma} w_{\ell} x_{\ell_i}^* a_{ij} x_{j_{\ell}}^* (z_{\ell_K} - \bar{z}_K) a_{KR} (z_{R_u} - \bar{z}_R) x_{u_s}^* m_{st} x_{t_u}^* w_u \frac{q_u}{\gamma} \\
& - \sum \frac{q\ell}{\gamma} w_{\ell} x_{\ell_i}^* a_{ij} x_{j_{\ell}}^* (z_{\ell_K} - \bar{z}_K) a_{KR} (z_{R_u} - \bar{z}_R) (z_{u_S} - \bar{z}_S) m_{ST} (z_{T_u} - \bar{z}_T) \\
& - \sum \frac{q\ell}{\gamma} w_{\ell} x_{\ell_i}^* a_{ij} x_{j_{\ell}}^* x_{u_s}^* m_{st} x_{t_u}^* w_u \frac{q_u}{\gamma} \frac{1}{\gamma^2} \frac{2\gamma^2}{n} \\
& - \sum \frac{q\ell}{\gamma} w_{\ell} x_{\ell_i}^* a_{ij} x_{j_{\ell}}^* (z_{\ell_S} - \bar{z}_S) m_{ST} (z_{T_u} - \bar{z}_T) \frac{1}{\gamma^2} \frac{2\gamma^2}{n} \\
& - \sum \frac{1}{2} (z_{\ell_I} - \bar{z}_I) a_{IJ} (z_{J_{\ell}} - \bar{z}_J) (z_{\ell_K} - \bar{z}_K) a_{KR} (z_{R_u} - \bar{z}_R) x_{u_s}^* m_{st} x_{t_u}^* w_u \frac{q_u}{\gamma} \\
& - \sum \frac{1}{2} (z_{\ell_I} - \bar{z}_I) a_{IJ} (z_{J_{\ell}} - \bar{z}_J) (z_{\ell_K} - \bar{z}_K) a_{KR} (z_{R_u} - \bar{z}_R) (z_{u_S} - \bar{z}_S) m_{ST} (z_{T_u} - \bar{z}_T) \\
& - \sum (z_{\ell_I} - \bar{z}_I) a_{IJ} (z_{J_{\ell}} - \bar{z}_J) x_{u_s}^* m_{st} x_{t_u}^* w_u \frac{q_u}{\gamma} \frac{1}{\gamma^2} \frac{2\gamma^2}{n} \\
& \left. - \sum (z_{\ell_I} - \bar{z}_I) a_{IJ} (z_{J_{\ell}} - \bar{z}_J) (z_{R_u} - \bar{z}_R) (z_{u_S} - \bar{z}_S) m_{ST} (z_{T_u} - \bar{z}_T) \frac{1}{\gamma^2} \frac{2\gamma^2}{n} \right\},
\end{aligned}$$

que em forma matricial pode ser escrito como

$$\begin{aligned}
A_{12} &= 6\mathbf{1}^\top \Phi F Z_{\beta_2 d} Z_{\beta_2} (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1} + 6\mathbf{1}^\top \Phi W P_2 Z_{\beta_2} (Z_\beta - Z_{\beta_2})_d (F - G) \Phi \mathbf{1} \\
&+ 6\mathbf{1}^\top \Phi W Z_{\beta_2 d} Z_{\delta_2} (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} + 6\mathbf{1}^\top \Phi W Z_{\beta_2 d} Z_{\delta_2} (Z_\delta - Z_{\delta_2})_d \mathbf{1} \\
&+ \frac{12}{n} \mathbf{1}^\top \Phi W Z_{\beta_2 d} (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} + \frac{12}{n} \mathbf{1}^\top \Phi W Z_{\beta_2 d} (Z_\delta - Z_{\delta_2})_d \mathbf{1} \\
&+ 3\mathbf{1}^\top Z_{\delta_2 d} Z_{\delta_2} (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} + 3\mathbf{1}^\top Z_{\delta_2 d} Z_{\delta_2} (Z_\delta - Z_{\delta_2})_d \mathbf{1} \\
&+ \frac{6}{n} \mathbf{1}^\top Z_{\delta_2 d} (Z_\beta - Z_{\beta_2})_d W \Phi \mathbf{1} + \frac{6}{n} \mathbf{1}^\top Z_{\delta_2 d} (Z_\delta - Z_{\delta_2})_d \mathbf{1}.
\end{aligned}$$

$$\begin{aligned}
A_{13} &= 6 \sum \{ (\kappa_{i,jk} - \kappa_{i,j,k}) (\kappa_{rst} + 2\kappa_{rs,t}) a_{js} a_{kt} m_{ir} - (\kappa_{rst} + 2\kappa_{rs,t}) \kappa_{I,j,k} a_{js} a_{kt} m_{IR} \\
&+ (\kappa_{i,j\gamma} - \kappa_{i,j,\gamma}) \kappa_{rs\gamma} a_{js} a_{\gamma\gamma} m_{ir} + (\kappa_{i,Jk} - \kappa_{i,J,k}) (\kappa_{rSt} + 2\kappa_{rS,t}) a_{JS} a_{kt} m_{ir} \\
&+ (\kappa_{I,JK} - \kappa_{I,J,K}) (\kappa_{RST} + 2\kappa_{RS,T}) a_{JS} a_{KT} m_{IR} + (\kappa_{I,J\gamma} - \kappa_{I,J,\gamma}) (\kappa_{RS\gamma} + 2\kappa_{RS,\gamma}) a_{JS} a_{\gamma\gamma} m_{IR} \\
&+ (\kappa_{i,\gamma} - \kappa_{i,\gamma,k}) (\kappa_{r\gamma t} + 2\kappa_{r\gamma,t}) a_{\gamma\gamma} a_{kt} m_{ir} + (\kappa_{I,\gamma K} - \kappa_{I,\gamma,K}) (\kappa_{R\gamma t} + 2\kappa_{R\gamma,t}) a_{js} a_{kt} m_{ir}.
\end{aligned} \tag{B.8}$$

Assim, obtemos

$$\begin{aligned}
A_{13} &= 6 \sum_{\ell,u=1}^n \left\{ \sum \frac{-q_\ell}{\gamma} (2g_\ell - f_\ell) x_{\ell_i}^* m_{ir} x_{r_u}^* x_{\ell_j}^* a_{js} x_{s_u}^* x_{\ell_j}^* a_{js} x_{s_u}^* f_u \frac{q_u}{\gamma} \right. \\
&- \sum \frac{q_\ell}{\gamma} (2g_\ell - f_\ell) x_{\ell_i}^* m_{ir} x_{r_k}^* a_{kt} x_{k_\ell}^* x_{\ell_j}^* a_{js} x_{s_t}^* w_u \frac{q_u}{\gamma} \\
&- \sum \frac{q_\ell}{\gamma} w_\ell x_{\ell_i}^* m_{ir} x_{r_u}^* a_{js} x_{s_u}^* x_{u_t}^* a_{tk} x_{\ell_j k}^* f_u \frac{q_u}{\gamma} \\
&- \sum \frac{q_\ell}{\gamma} w_\ell x_{\ell_i}^* m_{ir} x_{r_u}^* x_{\ell_j k}^* a_{kt} x_{tsu}^* w_u \frac{q_u}{\gamma}
\end{aligned}$$

$$\begin{aligned}
& + \sum \frac{q\ell}{\gamma} w_\ell (z_{\ell_I} - \bar{z}_I) m_{IR} (z_{R_u} - \bar{z}_R) x_{\ell_j}^* a_{js} x_{s_u}^* x_{\ell_k}^* a_{kt} x_{t_u}^* w_u \frac{q_u}{\gamma} \\
& - \sum 2 \frac{q\ell}{\gamma} w_\ell x_{\ell_i}^* m_{ir} x_{r_u}^* x_{\ell_j}^* a_{js} x_{s_u}^* w_u \frac{q_u}{\gamma} \frac{1}{\gamma^2} \frac{2\gamma^2}{n} \\
& + \sum 2 \frac{q\ell}{\gamma} w_\ell x_{\ell_i}^* m_{ir} x_{r_u}^* x_{\ell_k}^* a_{kt} x_{t_u}^* (z_{\ell_J} - \bar{z}_J) a_{JS} (z_{S_u} - \bar{z}_S) w_u \frac{q_u}{\gamma} \\
& + \sum \frac{3}{4} (z_{\ell_I} - \bar{z}_I) m_{IR} (z_{R_u} - \bar{z}_R) (z_{\ell_J} - \bar{z}_J) a_{JS} (z_{S_u} - \bar{z}_S) (z_{\ell_K} - \bar{z}_K) a_{KT} (z_{T_u} - \bar{z}_T) \\
& + \sum \frac{3}{2} (z_{\ell_I} - \bar{z}_I) m_{IR} (z_{R_u} - \bar{z}_R) (z_{\ell_J} - \bar{z}_J) a_{JS} (z_{S_u} - \bar{z}_S) \frac{1}{2\gamma^2} \frac{2\gamma^2}{n} \\
& + \sum 2 \frac{q\ell}{\gamma} w_\ell x_{\ell_i}^* m_{ir} x_{r_u}^* x_{\ell_k}^* a_{kt} x_{t_u}^* w_u \frac{q_u}{\gamma} \frac{1}{\gamma^2} \frac{2\gamma^2}{n} \\
& + \sum \frac{3}{2} (z_{\ell_I} - \bar{z}_I) m_{IR} (z_{R_u} - \bar{z}_R) (z_{\ell_K} - \bar{z}_K) a_{KT} (z_{T_u} - \bar{z}_T) \frac{1}{2\gamma^2} \frac{2\gamma^2}{n}.
\end{aligned}$$

Definimos  $C_2 = \{c_{\ell u}\}$  e  $J_2 = \{j_{\ell u}\}$ , matrizes  $n \times n$ , com

$$c_{\ell u} = x_{2u}^* (X_2^{*\top} W \Phi X_2^*)^{-1} D_{22}^{(\ell)} (X_2^{*\top} W \Phi X_2^*)^{-1} x_{2u}^{*\top}$$

e

$$j_{\ell u} = \text{tr}\{D_{22}^{(\ell)} (X_2^{*\top} W \Phi X_2^*)^{-1} D_{22}^{(u)} (X_2^{*\top} W \Phi X_2^*)^{-1}\},$$

em que  $x_{2u}^*$  denota a  $u$ -ésima linha de  $X_2^*$ .

Assim, escrevemos  $A_{13}$  matricialmente como

$$\begin{aligned}
A_{13} & = -6\mathbf{1}^\top \Phi (2G - F) [Z_{\beta_2}^{(2)} \odot (Z_\beta - Z_{\beta_2})] F \Phi \mathbf{1} - 6\mathbf{1}^\top \Phi W [(Z_\beta - Z_{\beta_2}) \odot J_2] W \Phi \mathbf{1} \\
& - 6\mathbf{1}^\top \Phi W [(Z_\beta - Z_{\beta_2}) \odot C_2^\top] F \Phi \mathbf{1} + \frac{9}{2} \mathbf{1}^\top [Z_{\delta_2}^{(2)} \odot (Z_\delta - Z_{\delta_2})] \mathbf{1} + \frac{18}{n} \mathbf{1}^\top [Z_{\delta_2} \odot (Z_\delta - Z_{\delta_2})] \mathbf{1} \\
& - 6\mathbf{1}^\top \Phi (2G - F) [(Z_\beta - Z_{\beta_2}) \odot C_2] W \Phi \mathbf{1}
\end{aligned}$$

$$+ \mathbf{61}^\top \Phi W [(Z_\delta - Z_{\delta_2}) \odot Z_{\beta_2}^{(2)}] W \Phi \mathbf{1} + \mathbf{121}^\top \Phi W [(Z_\beta - Z_{\beta_2}) \odot Z_{\beta_2} Z_{\delta_2}] W \Phi \mathbf{1}.$$

Os passos para a determinação de fórmulas matriciais para os demais termos são análogos e, portanto, são aqui omitidos.

## Apêndice C

### Conjuntos de dados

#### C.1 Elasticidade constante de substituição da função de produção

Tabela C.1: *Output* agregado em uma certa indústria ( $y_t$ ), trabalho ( $L_t$ ) e capital ( $K_t$ ), numa amostra de 30 observações.

$y_t$	$L_t$	$K_t$	$y_t$	$L_t$	$K_t$
0,256918	0,228	0,802	0,186747	0,664	0,129
0,183599	0,258	0,249	0,020671	0,631	0,017
1,212883	0,821	0,771	0,100159	0,059	0,906
0,522568	0,767	0,511	0,252334	0,811	0,223
0,847894	0,495	0,758	0,103312	0,758	0,145
0,763379	0,487	0,425	0,078945	0,050	0,161
0,623130	0,678	0,452	0,005799	0,823	0,006
1,031485	0,748	0,817	0,723250	0,483	0,836
0,569498	0,727	0,845	0,776468	0,682	0,521
0,882497	0,695	0,958	0,216536	0,116	0,930
0,108827	0,458	0,084	0,541182	0,440	0,495
0,026437	0,981	0,021	0,316320	0,456	0,185
0,003750	0,002	0,295	0,123811	0,342	0,092
0,461626	0,429	0,277	0,38654	0,358	0,485
0,268474	0,231	0,546	0,279431	0,162	0,934

Fonte: Griffiths *et al.* (1993, p. 722)

## C.2 Modelo de crescimento do pasto

Tabela C.2: Taxa de crescimento do pasto ( $y_i$ ) e o tempo decorrido desde o último corte do pasto ( $x_i$ ) para uma amostra de 9 observações.

$y_i$	$x_i$
8,93	9
10,80	14
18,59	21
39,35	28
22,33	42
56,11	57
61,73	63
64,62	70
67,08	79

Fonte: Ratkowsky (1983)

## Referências Bibliográficas

- Bates, D. M. & Watts, D. G. (1988). *Nonlinear regression analysis and its applications*. Wiley, New York. 2
- Botter, D. A. & Cordeiro, G. M. (1998). Improved estimates for generalized linear models with dispersion covariates. *Journal of Statistical Computation and Simulation.*, **62**, 91–104. 2
- Bowman, K. O. & Shenton, L. R. (1998). Asymptotic skewness and the distribution of maximum likelihood estimators. *Communications Statistics Theory and Methods*, **27**, 2743–2760. 3, 19
- Cordeiro, G. M. (1983). Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistics Society, B*, **45**, 404–413. 1
- Cordeiro, G. M. (1987). On the corrections to the likelihood ratio statistics. *Biometrika*, **74**, 265–274. 1
- Cordeiro, G. M. (2004). Second-order covariance matrix of maximum likelihood estimates in generalized linear models. *Statistics and Probability Letters*, **66**, 153–160. 2, 3, 5, 8
- 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. 1, 31, 36
- Cordeiro, G. M. & McCullagh, P. (1991). Bias correction in generalized linear models. *Journal of the Royal Statistics Society, B*, **53**, 629–643. 2, 10
- Cordeiro, G. M. & Paula, G. A. (1989). Improved likelihood ratio statistics for exponential family nonlinear models. *Biometrika*, **76**, 93–100. 1, 2, 16
- Cordeiro, G. M. & Santana, R. G. (2008). Covariance matrix formula for exponential family nonlinear models. *Communications Statistics Theory and Methods*, **37**, 2724–2734. 2
- Cordeiro, G. M. & Udo, M. C. T. (2008). Bias correction in generalized nonlinear models with dispersion covariates. *Communications Statistics Theory and Methods*, **37**, 2219–2225. 2
- Cordeiro, G. M., Ferrari, S. L. P. & Paula, G. A. (1993). Improved score tests for generalized linear models. *Journal of the Royal Statistical Society, B*, **55**, 661–674. 1

- Cordeiro, G. M., Ferrari, S. L. P. & Cysneiros, A. H. M. A. (1998). A formula to improve score test statistics. *Journal of Statistical Computation and Simulation*, **62**, 123–136. 1
- Cordeiro, G. M., Botter, D. A., Barroso, L. P. & Ferrari, S. L. P. (2003). Three corrected score tests for generalized linear models with dispersion covariates. *Statistica Neerlandica*, **57**, 391–409. 1
- Cordeiro, G. M., Barroso, L. P. & Botter, D. A. (2006). Covariance matrix formula for generalized linear models with unknown dispersion. *Communications Statistics Theory and Methods*, **35**, 113–120. 2
- Cordeiro, H. H. & Cordeiro, G. M. (2001). Skewness for parameters in generalized linear models. *Communications Statistics Theory and Methods*, **30**, 1317–1334. 2, 4, 16, 21, 22
- Cox, D. R. & Hinkley, D. V. (1974). *Theoretical statistics*. John Wiley, New York. 17, 33
- Cox, D. R. & Reid, N. (1987). Parameter orthogonality and approximate conditional inference (with discussion). *Journal of the Royal Statistics Society, B*, **49**, 1–39. 4, 17
- Cribari-Neto, F. & Ferrari, S. L. P. (1995). Second order asymptotics for score tests in generalized linear models. *Biometrika*, **82**, 426–432. 1
- Cysneiros, A. H. M. A. & Ferrari, S. L. P. (2006). An improved likelihood ratio test for varying dispersion in exponential family nonlinear models. *Statistics and Probability Letters*, **76**, 255–265. 2
- Ferrari, S. L. P. & Cordeiro, G. M. (1996). Corrected score tests for exponential family nonlinear models. *Statistics and Probability Letters*, **26**, 7–12. 2
- Ferrari, S. L. P. & Cysneiros, A. H. M. A. (2008). Skovgaard’s adjustment to likelihood ratio tests in exponential family nonlinear models. *Statistics and Probability Letters*, **78**, 3047–3055. 2
- Ferrari, S. L. P., Botter, D. A. & Cribari-Neto, F. (1996). Second- and third-order bias reduction for one-parameter family models. *Statistics and Probability Letters*, **30**, 339–345. 5
- Ferrari, S. L. P., Uribe-Opazo, M. A. & Cribari-Neto, F. (1997). Second order asymptotics for score tests in exponential family nonlinear models. *Journal of Statistical Computation and Simulation*, **59**, 179–194. 2
- Ferrari, S. L. P., Cysneiros, A. H. M. A. & Cribari-Neto, F. (2004). An improved test for heteroskedasticity using adjusted modified profile likelihood inference. *Journal of Statistical Planning and Inference*, **124**, 423–437. 55
- Gallant, A. R. (1987). *Nonlinear statistical models*. Wiley, New York. 2



- Griffiths, W. E., Hill, R. C. & Judge, G. G. (1993). *Learning and practicing econometrics*. John Wiley, New York. 28, 73
- Harris, P. (1985). An asymptotic expansion for the null distribution of the efficient score statistic. *Biometrika*, **72**, 653–659. 36, 63
- Jørgensen, B. (1987). Exponential dispersion models (with discussion). *Journal of the Royal Statistics Society, B*, **49**, 127–162. 16
- Kakizawa, Y. (1996). Higher order monotone bartlett-type adjustment for some multivariate test statistics. *Biometrika*, **83**, 923–927. 1
- Lehmann, E. L. & Casella, E. (1998). *Theory of point estimation*. Springer-Verlag, New York. 17
- McCullagh, P. & Nelder, J. A. (1989). *Generalized linear models*. Chapman and Hall, London. 2
- Nelder, J. A. & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistics Society, A*, **135**, 370–384. 1, 16
- Pace, L. & Salvan, A. (1997). *Principles of statistical inference*. World Scientific, Singapore. 7
- Paula, G. A. (1992). Bias correction for exponential family nonlinear models. *Journal of Statistical Computation and Simulation*, **40**, 43–54. 2, 21
- Rao, C. R. (1973). *Linear statistical inference and its applications*. Wiley, New York. 9, 21
- Ratkowsky, D. A. (1983). *Nonlinear regression in modeling: a unified practical approach*. Marcel Dekker, New York. 2, 16, 29, 74
- Ratkowsky, D. A. (1990). *Handbook of nonlinear regression models*. Marcel Dekker, New York. 2
- Seber, G. A. F. & Wild, C. J. (1989). *Nonlinear regression*. John Wiley, New York. 16
- Skovgaard, I. M. (2001). Likelihood asymptotics. *Scandinavian Journal of Statistics*, **28**, 3–32. 2
- Smyth, G. K. (1989). Generalized linear models with varying dispersion. *Journal of the Royal Statistics Society, B*, **51**, 47–60. 31
- Wei, D. A. (1998). *Exponential family nonlinear models*. Springer-Verlag, Singapore. 2