

**CORREÇÃO DE VIÉS E DE BARTLETT EM MODELOS EM SÉRIES DE
POTÊNCIA NÃO-LINEARES GENERALIZADOS**

Priscila Gonçalves da Silva

Orientadora: Prof^a Dr^a Audrey Helen Mariz de Aquino Cysneiros

Área de Concentração: Estatística Matemática

**Dissertação submetida como requerimento parcial para obtenção do grau de
Mestre em Estatística pela Universidade Federal de Pernambuco**

Recife, fevereiro de 2010

Silva, Priscila Gonçalves da
Correção de viés e de Bartlett em modelos em séries de
potência não-lineares generalizados / Priscila Gonçalves da
Silva. - Recife: O Autor, 2010.
iii, 89 folhas : il., fig., tab.

Dissertação (mestrado) – Universidade Federal de
Pernambuco. CCEN. Estatística, 2010.

Inclui bibliografia e apêndice.

1. Estatística Matemática. I. Título.

519.9

CDD (22. ed.)

MEI2010 – 034

Universidade Federal de Pernambuco
Pós-Graduação em Estatística

26 de fevereiro de 2010
(data)

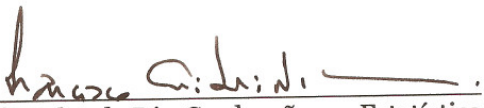
Nós recomendamos que a dissertação de mestrado de autoria de

Priscila Gonçalves da Silva

intitulada

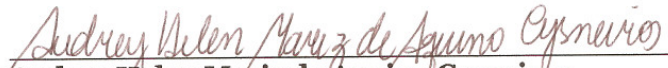
“Correção de viés em modelos em séries de potência não-lineares generalizados”

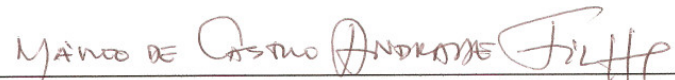
seja aceita como cumprimento parcial dos requerimentos para o grau de Mestre em Estatística.

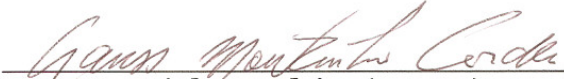

Coordenador da Pós-Graduação em Estatística

Banca Examinadora:

 Prof. Francisco Cribari Neto
Coordenador
Pós-Graduação em Estatística UFPE


Audrey Helen Mariz de Aquino Cysneiros orientadora


Mário de Castro Andrade Filho (USP/São Carlos)


Gauss Moutinho Cordeiro (UFRPE)

Este documento será anexado à versão final da dissertação.

-

Agradecimentos

- Primeiramente ao meu Deus, por ter me proporcionado mais esta conquista.
- À professora Audrey Helen Mariz de Aquino Cysneiros, por sua sublime orientação com que direcionou esse trabalho.
- Aos meus pais, Josabete e Natanael, pela educação que a mim foi dada, pelo apoio e compreensão.
- A toda minha família pelo apoio e incentivo, em especial aos meus irmãos, Jenilson e Patricia, e meus sobrinhos, Thiago e João.
- Aos meus colegas de Mestrado por compartilharmos momentos de dificuldades e superação.
- A todos os professores e funcionários do Departamento de Estatística da UFPE por seus trabalhos realizados.
- Aos professores Gauss Moutinho Cordeiro e Mário de Castro Andrade Filho pelas sugestões.
- À CAPES, ao CNPq e à FACEPE pelo apoio financeiro oferecido.

Resumo

Esta dissertação tem dois objetivos. O primeiro é a obtenção da correção de viés de segunda ordem dos estimadores de máxima verossimilhança na classe dos Modelos em Série de Potência Não-Lineares Generalizados, considerando o parâmetro de dispersão conhecido, via Cox & Snell (1968) e bootstrap (Efron, 1979). O segundo objetivo é a obtenção da correção de Bartlett à estatística da razão de verossimilhanças nesta classe de modelos. Desenvolvemos estudos de simulação para avaliar e comparar numericamente o comportamento dos estimadores de máxima verossimilhança, bem como o de suas versões corrigidas, em amostras finitas. Adicionalmente, avaliamos numericamente o desempenho dos testes da razão de verossimilhanças e suas versões corrigidas em relação ao tamanho e poder em amostras finitas. Por fim, realizamos uma aplicação empírica.

Palavras-chave: Modelos em Série de Potência; Correção de Viés; Correção de Bartlett.

Abstract

This dissertation has two purposes. The first one is to obtain the second-order bias correction of the maximum likelihood estimators in the class of the in Power Series Generalized Nonlinear Models, considering the dispersion parameter known, via Cox & Snell (1968) and bootstrap (Efron, 1979). The second objective is to obtain the Bartlett correction to the likelihood ratio statistic in this class of models. Numerical evaluation is performed involving the different estimators. Additionally, we have numerically evaluated the finite sample performance of likelihood ratio tests and its Bartlett-corrected versions on the size and power. Finally, we present one empirical application.

Keywords: Power Series Models; Bias Correction; Bartlett Correction.

Índice

	Página
1 Introdução	1
2 Modelos em séries de potência não-lineares generalizados (MSPNLGs)	4
2.1 Introdução	4
2.2 Definição	5
2.3 Aspectos inferenciais	9
2.3.1 Estimação dos parâmetros de regressão	9
2.3.2 Testes da razão de verossimilhanças em MSPNLG	11
3 Correção de viés em MSPNLG	12
3.1 Introdução	12
3.2 Correção de Cox & Snell	14
3.2.1 Correção de viés dos EsMV dos MSPNLGs	15
3.3 Correção via bootstrap	18
3.4 Resultados numéricos	20
3.4.1 Modelos lineares	21
3.4.2 Modelos não-lineares	29
3.5 Aplicação	39
3.6 Comentários	42

4	Correção de Bartlett em MSPNLG	44
4.1	Introdução	44
4.2	Correção de Bartlett	46
4.2.1	Correção de Bartlett em MSPNLG	48
4.3	Resultados numéricos	55
4.3.1	Modelos lineares	56
4.3.2	Modelos não-lineares	61
4.4	Aplicação	66
4.5	Comentários	68
5	Conclusões	69
	Apêndice	71
A	Cálculo dos Momentos	71
A.1	Derivadas do logaritmo da função de verossimilhança	73
A.2	Cálculo de cumulantes	75
A.2.1	Derivadas dos cumulantes	75
A.3	Cálculo de $\sum \lambda_{rstuvw}$	76
B	Conjuntos de dados	82
	Referências bibliográficas	84

Capítulo 1

Introdução

Com o intuito de unificar vários modelos discretos importantes em uma única estrutura conceitual, Cordeiro *et al.* (2009) propuseram uma nova classe de modelos em séries de potências não-lineares generalizados (MSPNLG). Esta classe de modelos é definida pela família de distribuições em séries de potências modificada para representar a variável resposta em termos da média e uma função de ligação não-linear para a média da mesma. Desta forma, esta classe de modelos abrange modelos tradicionais tais como os modelos log-não-lineares, binomial não-lineares e binomial negativa não-lineares. Nesta dissertação destacamos dois aspectos inferenciais no MSPNLG: o primeiro corresponde à obtenção da expressão do viés de segunda ordem dos estimadores de máxima verossimilhança (EsMV) dos parâmetros do modelo e o segundo visa à obtenção de ajustes para a estatística da razão de verossimilhanças (LR).

A estimação dos parâmetros no modelo MSPNLG é feita pelo método da máxima verossimilhança, que por sua vez fornece, em geral, estimadores viesados. Em alguns casos, o viés é considerado insignificante quando comparado ao erro-padrão dos EsMV, visto que ele é de ordem n^{-1} , enquanto o desvio padrão da estimativa é de ordem $n^{-1/2}$. Porém, no caso de modelos não-lineares quando o tamanho da amostra é pequeno ou a informação de Fisher é reduzida, o viés passa a ter uma magnitude comparável ao erro padrão do EMV (Cordeiro, 1999). Deste modo, é de suma importância o cálculo dos vieses de segunda ordem dos EsMV a fim de obtermos estimadores mais precisos.

Já a estatística da razão de verossimilhanças, em problemas regulares e sob a hipótese

nula, tem uma distribuição qui-quadrado aproximadamente, em grandes amostras, e o erro desta aproximação é de ordem n^{-1} . Desta forma, torna-se importante obter ajustes para a estatística LR que reduzam esse erro de aproximação. A idéia é modificar essa estatística por um fator de correção, visando produzir uma nova estatística com o primeiro momento igual ao da distribuição qui-quadrado de referência.

Diante disso, o objetivo deste trabalho é fornecer uma expressão do viés de segunda ordem de Cox & Snell (1968) dos estimadores de máxima verossimilhança dos parâmetros do MSPNLG e outra expressão da correção de Bartlett para a estatística da razão de verossimilhanças.

No Capítulo 2, discorremos sobre a família de distribuições em séries de potência, expondo algumas características e propriedades, bem como a definição do MSPNLG e seus aspectos inferenciais. No Capítulo 3, fornecemos uma expressão, em forma fechada, para o viés dos EsMV dos parâmetros do MSPNLG, considerando o parâmetro de dispersão conhecido. Adicionalmente, discorremos sobre a obtenção da correção de viés via a metodologia bootstrap (Efron, 1979). Resultados numéricos sobre o desempenho dos EsMV, bem como das suas versões corrigidas, em amostras de tamanho finito, aplicações a dados reais e algumas considerações finais também são apresentados neste capítulo. O Capítulo 4 trata de ajustes para a estatística da razão de verossimilhanças, com base na correção de Bartlett. Comportamentos em amostras finitas dos testes baseados na estatística LR e nas suas versões corrigidas são apresentados em relação ao tamanho e poder. Complementando, uma aplicação com dados reais e alguns comentários são encontrados no Capítulo 4. Por fim, no Capítulo 5, são expostas as conclusões deste estudo. Vale salientar aqui que esta dissertação foi escrita de tal forma que os Capítulos 2, 3 e 4 são independentes, significando que alguns resultados básicos e notação são apresentados mais de uma vez.

Finalmente, deve ser enfatizado que não há resultados na literatura relacionados com os MSPNLGs. Nosso trabalho veio preencher esta lacuna, tornando-se o pioneiro em explorar essa classe de modelos. Deve ainda ser destacado que os Capítulos 3 e 4 são as principais contribuições teóricas desta dissertação de mestrado.

Nesta dissertação, os resultados numéricos foram obtidos utilizando a versão 4.10 da linguagem matricial de programação `0x` para sistema operacional Windows. Esta linguagem foi criada por Jurgen Doornik, em 1994, na Universidade de Oxford (Inglaterra). Ela é

muito flexível com sintaxe similar às sintaxes das linguagens de programação C e C++. Mais detalhes sobre esta linguagem de programação podem ser encontrados em Doornik (2001) e em Cribari-Neto e Zarkos (2003). As apresentações gráficas foram produzidas com o ambiente de programação R, tendo sido utilizada a versão 2.8.0 para a plataforma Windows. O R é um ambiente integrado que possui grandes facilidades para manipulação de dados, geração de gráficos e modelagem estatística em geral (vide Cribari-Neto e Zarkos, 1999; Ihaka e Gentleman 1996; Venables e Ripley, 2002).

Capítulo 2

Modelos em séries de potência não-lineares generalizados (MSPNLGs)

2.1 Introdução

Dados na forma de contagem são frequentemente analisados utilizando modelos de regressão Poisson e binomial negativa (Cameron e Trivedi, 1998). No entanto, em muitas situações podem ocorrer os fenômenos conhecidos como superdispersão e subdispersão, que acontecem quando a variância da variável resposta é maior ou menor do que a sua média, respectivamente. Nesses casos, a suposição de distribuição de Poisson para a resposta é inadequada, sendo necessário o uso de modelos alternativos. Neste contexto, a nova classe de modelos em séries de potências não-lineares generalizados (MSPNLG) foi proposta por Cordeiro *et al.* (2009) com o objetivo de acomodar as diferentes relações de dispersão.

A classe MSPNLG é definida por um conjunto de variáveis aleatórias independentes pertencentes à família de distribuições em séries de potências, adotando o componente sistemático dos modelos não-lineares da família exponencial (Cordeiro e Paula, 1989). Este componente consiste de uma função de ligação não-linear entre a média da variável resposta e a estrutura não-linear do modelo. Já o componente aleatório do MSPNLG é definido por uma subclasse de distribuições em série de potências, originalmente proposta por Gupta (1974) e, posteriormente, expressa em termos de sua média por Consul (1990).

2.2 Definição

Sejam $Y_1, \dots, Y_n$ variáveis aleatórias discretas independentes e tais que Y_i segue uma família de distribuições com parâmetros de média $\mu_i > 0$ e parâmetro de dispersão $\phi > 0$, com função de probabilidade na forma

$$\pi(y; \mu_i, \phi) = \frac{a(y, \phi)g(\mu_i, \phi)^y}{f(\mu_i, \phi)}, \quad y \in A_\epsilon, \quad (2.1)$$

em que o suporte de Y_i é um subconjunto A_ϵ dos inteiros $\{\epsilon, \epsilon + 1, \dots\}$ e que não depende de parâmetros desconhecidos, $\epsilon \geq 0$, $a(y, \phi)$ é positiva e as funções analíticas $g(\mu_i, \phi)$ e $f(\mu_i, \phi)$ dos parâmetros μ_i e ϕ são positivas, finitas e duas vezes diferenciáveis. Satisfeita a suposição de que o parâmetro ϕ , que assumimos ser conhecido, é maior do que 0, temos que a variável aleatória Y tem uma distribuição de probabilidade completamente determinada por sua função de variância. A função $f(\mu, \phi)$ é tal que

$$f(\mu, \phi) = \sum_{y \in A_\epsilon} a(y, \phi)g(\mu, \phi)^y.$$

Para a família de distribuições dada em (2.1), valem as seguintes relações:

$$E(Y) = \mu = \frac{f'g}{fg'} \quad \text{e} \quad Var(Y) = V(\mu, \phi) = \frac{g}{g'}, \quad (2.2)$$

em que $f = f(\mu_i, \phi)$, $g = g(\mu_i, \phi)$ e o símbolo “ r ” indica a diferenciação em relação a μ . Observe que a função de variância depende apenas da função $g(\mu, \phi)$ e pode ser expressa como um fator multiplicativo da média dado por $V(\mu, \phi) = [(\log f)']^{-1}\mu$. A média de Y_i , está relacionada com o componente sistemático através de uma função de ligação da forma

$$h(\mu_i) = \eta_i = \eta(x_i; \beta), \quad i = 1, \dots, n, \quad (2.3)$$

em que $h(\cdot)$ é uma função de ligação conhecida e duplamente diferenciável, $\beta = (\beta_1, \dots, \beta_p)^\top$ é um vetor de p ($p < n$) parâmetros desconhecidos a serem estimados, $x_i = (x_{i1}, \dots, x_{ik})^\top$ representa os valores de k variáveis explicativas e $\eta(\cdot; \cdot)$ é uma função possivelmente não-linear no segundo argumento, contínua e diferenciável com respeito aos componentes de β tal que a matriz de derivadas $\tilde{X} = \tilde{X}(\beta) = \partial\eta/\partial\beta^\top$, com $\eta = (\eta_1, \dots, \eta_n)^\top$, tem posto p para todo β . A matriz $\tilde{X}$ tem elementos que são, em geral, funções do vetor de parâmetros β desconhecidos.

As distribuições Poisson, Binomial, Binomial Negativa, Poisson Generalizada, Binomial Negativa Generalizada, Borel, Consul, Borel-Tanner, Geeta- m e Haight são algumas das distribuições pertencentes à família (2.1). Características destas distribuições são apresentadas na Tabela 2.1.

Consideremos um MSPNLG definido por (2.1) e (2.3). O logaritmo da função de verossimilhança dos parâmetros do modelo, dado o vetor de observações $(y_1, \dots, y_n)^\top$, é dado por

$$l(\beta; y) = \sum_{i=1}^n \log\{a(y_i, \phi)\} + \sum_{i=1}^n [y_i \log\{g(\mu_i, \phi)\} - \log\{f(\mu_i, \phi)\}]. \quad (2.4)$$

Com o parâmetro ϕ conhecido, (2.1) é uma família exponencial natural discreta uniparamétrica e a função de variância $V(\mu, \phi)$ determina a função desvio do MSPNLG em consideração, que pode ser escrito como $D^{(\phi)} = 2 \sum_{i=1}^n D_i(y_i, \hat{\mu}_i^{(\phi)})$, em que

$$D_i(y_i, \mu_i) = \left[y_i \log \left\{ \frac{g(y_i, \phi)}{g(\mu_i, \phi)} \right\} + \log \left\{ \frac{f(\mu_i, \phi)}{f(y_i, \phi)} \right\} \right]$$

e $\hat{\mu}^{(\phi)}$ é a estimativa de máxima verossimilhança de μ , considerando conhecidos o parâmetro ν da distribuição Binomial Negativa Generalizada e o parâmetro ϕ . O desvio $D^{(\phi)}$, que depende do parâmetro ϕ conhecido ou consistentemente estimado, tem uma distribuição χ_{n-p}^2 aproximada, embora em geral esta aproximação possa não ser válida, pois a dimensão do modelo saturado depende de n e os argumentos assintóticos usuais não são aplicáveis. Apresentaremos a seguir a função desvio dos modelos referentes às distribuições apresentadas na Tabela 2.1, que nestes casos, tem distribuição χ_{n-p}^2 aproximada pelo menos quando todos os μ_i 's são grandes ou n é grande.

1. Poisson generalizada:

$$D(y, \mu) = 2 \left[y \log \left\{ \frac{y(1 + \phi\mu)}{\mu(1 + \phi y)} \right\} - \frac{y - \mu}{1 + \mu\phi} \right].$$

Aqui $V(\mu, \phi) = \mu(1 + \phi\mu)^2$. Esse modelo reduz-se ao modelo de Poisson quando $\phi = 0$.

2. Binomial negativa generalizada (BNG):

$$D(y, \mu) = 2y \log \left\{ \frac{y}{\mu} \left(\frac{\nu + \phi\mu}{\nu + \phi y} \right)^\phi \left(\frac{\nu + \phi y - y}{\nu + \phi\mu - \mu} \right)^{\phi-1} \right\} + 2\nu \log \left\{ \frac{(\nu + \phi\mu)(\nu + \phi y - y)}{(\nu + \phi y)(\nu + \phi\mu - \mu)} \right\},$$

em que $\nu > 0$ é suposto conhecido mas não necessariamente inteiro e $V(\mu, \phi) = \mu(1 + \frac{\phi\mu}{\nu})\{1 + \frac{(\phi-1)\mu}{\nu}\}$. O modelo BNG se reduz ao modelo Binomial e Binomial Negativa, quando $\phi = 0$ e $\phi = 1$, respectivamente.

3. Borel-Tanner:

$$D(y, \mu) = 2 \left[m \left(1 - \frac{y}{\mu} \right) + (y - m) \log \left\{ \left(\frac{\mu}{y} \right) \left(\frac{y - m}{\mu - m} \right) \right\} \right].$$

Nesse caso, $\phi = 1$ e $V(\mu, \phi) = (\mu - m)\mu^2/m^2$. Quando $m = 1$, temos o modelo da distribuição Borel.

4. Delta binomial:

$$D(y, \mu) = 2y\phi \log \left\{ \frac{\mu}{y} \left(\frac{\phi y - y + m}{\phi\mu - \mu + m} \right) \right\} + 2(y - m) \log \left\{ \frac{(y - m)}{(\mu - m)} \left(\frac{\phi\mu - \mu + m}{\phi y - y + m} \right) \right\}.$$

Aqui $V(\mu, \phi) = \mu(\mu - m)\{(\phi - 1)\mu + m\}/(\phi m^2)$. O caso especial $m = 1$ representa a distribuição Consul.

5. Geeta- m :

$$D(y, \mu) = 2y \log \left\{ \frac{(y - m)}{(\mu - m)} \left(\frac{y}{\mu} \right)^{\phi-1} \left(\frac{\phi\mu - m}{\phi y - m} \right)^{\phi} \right\} + 2m \log \left\{ \frac{(\mu - m)(\phi y - m)}{(y - m)(\phi\mu - m)} \right\}.$$

Aqui $V(\mu, \phi) = \mu(\mu/m - 1)(\phi\mu/m - 1)/(\phi - 1)$. A distribuição Geeta- m decorre da soma de m variáveis i.i.d. com distribuição Geeta, já esta se reduz à distribuição Haight quando $\phi = 2$.

Tabela 2.1: Funções f , g , a e o suporte de algumas distribuições da família (2.1).

Distribuição	$f(\mu, \phi)$	$g(\mu, \phi)$	$a(y, \phi)$	Suporte (A_ϵ)
1. Poisson	e^μ	μ	$\frac{1}{y!}$	$\{0, 1, 2, \dots\}$
2. Binomial	$\left(1 + \frac{\mu}{m-\mu}\right)^m$	$\frac{\mu}{m-\mu}$	$\binom{m}{y}$	$\{0, 1, 2, \dots, m\}$
3. Binomial negativa	$\left(1 - \frac{\mu}{\mu+\phi}\right)^{-\phi}$	$\frac{\mu}{\mu+\phi}$	$\frac{\Gamma(\phi+y)}{y! \Gamma(\phi)}$	$\{0, 1, 2, \dots\}$
4. Poisson generalizada	$e^{\mu(1+\mu\phi)^{-1}}$	$\frac{\mu e^{-\mu\phi(1+\mu\phi)^{-1}}}{1+\mu\phi}$	$\frac{(1+\phi y)^{y-1}}{y!}$	$\{0, 1, 2, \dots\}$
5. Borel	$1 - \frac{1}{\mu}$	$\left(1 - \frac{1}{\mu}\right) e^{-1+1/\mu}$	$\frac{y^{y-2}}{(y-1)!}$	$\{1, 2, \dots\}$
6. Consul	$\frac{\mu-1}{\mu(\phi-1)+1}$	$\phi^{-\phi} (1 - \mu^{-1})(\phi - 1 + \mu^{-1})^{\phi-1}$	$\frac{\Gamma(\phi y+1)}{y! \Gamma(\phi y-y+2)}$	$\{1, 2, \dots\}$
7. Binomial negativa generalizada	$\left(\frac{\phi-1+\nu/\mu}{\phi+\nu/\mu}\right)^{-\nu}$	$\frac{1}{\phi+\nu/\mu} \left(\frac{\phi-1+\nu/\mu}{\phi+\nu/\mu}\right)^{\phi-1}$	$\frac{\nu \Gamma(\phi y+\nu+1)}{(\phi y+\nu) y! \Gamma(\phi y-y+\nu+1)}$	$\{0, 1, 2, \dots\}$
8. Borel-Tanner	$\left(1 - \frac{m}{\mu}\right)^m$	$\left(1 - \frac{m}{\mu}\right) e^{-1+m/\mu}$	$\frac{m y^{y-m-1}}{(y-m)!}$	$\{m, m+1, \dots\}$
9. Delta binomial	$\left\{ \frac{\mu-m}{\mu(\phi-1)+m} \right\}^m$	$\frac{1}{\phi^\phi} \left(1 - \frac{m}{\mu}\right) \left(\phi - 1 + \frac{m}{\mu}\right)^{\phi-1}$	$\frac{m \Gamma(\phi y+1)}{y(y-m)! \Gamma(\phi y-y+m+1)}$	$\{m, m+1, \dots\}$
10. Geeta	$\frac{\mu-1}{\phi\mu-1}$	$\frac{\mu-1}{\phi\mu-1} \left\{ \frac{(\phi-1)\mu}{\phi\mu-1} \right\}^{\phi-1}$	$\frac{\Gamma(\phi y-1)}{y! \Gamma(\phi y-y)}$	$\{1, 2, \dots\}$
11. Geeta- m	$\left(\frac{\mu-m}{\phi\mu-m}\right)^m$	$\frac{\mu-m}{\phi\mu-m} \left\{ \frac{(\phi-1)\mu}{\phi\mu-m} \right\}^{\phi-1}$	$\frac{m \Gamma(\phi y-m)}{y(y-m)! \Gamma(\phi y-y)}$	$\{m, m+1, \dots\}$
12. Haigh	$\frac{\mu-1}{2\mu-1}$	$\frac{\mu(\mu-1)}{(2\mu-1)^2}$	$\frac{(2y-2)!}{y!(y-1)!}$	$\{1, 2, \dots\}$

2.3 Aspectos inferenciais

2.3.1 Estimação dos parâmetros de regressão

A função escore para o parâmetro β , condicionando em ϕ , é dada por

$$U_\beta = \frac{\partial l(\beta; y)}{\partial \beta} = \tilde{X}^\top (Ty - Q),$$

em que $T = \text{diag}\{t_1, \dots, t_n\}$ é uma matriz diagonal de dimensão $n \times n$ cujo i -ésimo elemento é $t_i = \frac{g'_i}{g_i h'_i}$ e $Q = (q_1, \dots, q_n)^\top$ é um vetor $n \times 1$ cujo i -ésimo elemento é $q_i = \frac{f'_i}{f_i h'_i}$. A matriz de informação de β , dado ϕ , é

$$K_\beta = E \left\{ - \frac{\partial^2 l(\beta; y)}{\partial \beta \partial \beta^\top} \right\} = \tilde{X}^\top W \tilde{X}, \quad (2.5)$$

em que W é uma matriz diagonal $n \times n$ de pesos dados por

$$w_i = \left(q'_i - \frac{f'_i g_i}{f_i g'_i} t'_i \right) \frac{1}{h'_i}.$$

Sejam $\bar{f} = \frac{f'}{f}$ e $\bar{\bar{f}} = \frac{f''}{f}$ (mantendo a mesma notação para a função g). Temos que

$$q'_i = \frac{f_i h'_i f''_i - (f'_i h'_i + f_i h''_i) f'_i}{(f_i h'_i)^2} = \frac{\bar{\bar{f}}_i}{h'_i} - \frac{\bar{f}_i^2}{h'_i} - \frac{\bar{f}_i h''_i}{(h'_i)^2}.$$

Do mesmo modo,

$$t'_i = \frac{\bar{\bar{g}}_i}{h'_i} - \frac{\bar{g}_i^2}{h'_i} - \frac{\bar{g}_i h''_i}{(h'_i)^2}.$$

Com isso, podemos mostrar que

$$w_i = \{ \bar{g}_i (\bar{\bar{f}}_i - \bar{f}_i^2) + \bar{f}_i (\bar{\bar{g}}_i^2 - \bar{g}_i^2) \} (\bar{g}_i)^{-1} h_i'^{-2}. \quad (2.6)$$

De (2.2) vem $\bar{f}_i = \mu_i \bar{g}_i$, ou seja, $f'_i = f_i \mu_i \bar{g}_i$. Dessa forma, temos que

$$f''_i = f'_i \mu_i \bar{g}_i + f_i \bar{g}_i + f_i \mu_i \left\{ \frac{g_i g''_i - (g'_i)^2}{g_i^2} \right\} = f'_i \mu_i \bar{g}_i + f_i \bar{g}_i + f_i \mu_i (\bar{\bar{g}}_i - \bar{g}_i^2).$$

Consequentemente,

$$\bar{\bar{f}}_i = \bar{f}_i \mu_i \bar{g}_i + \bar{g}_i + \mu_i (\bar{\bar{g}}_i - \bar{g}_i^2) = (\mu_i \bar{g}_i)^2 + \bar{g}_i + \mu_i (\bar{\bar{g}}_i - \bar{g}_i^2). \quad (2.7)$$

Substituindo $\bar{f}_i = \mu_i \bar{g}_i$ e (2.7) em (2.6), teremos que os elementos da matriz de pesos W que dependem da distribuição de Y_i reduzem-se à $w_i = \bar{g}_i h_i'^{-2} = V_i^{-1} h_i'^{-2}$, em que $V_i = V(\mu_i, \phi)$ e, portanto, $W = (LV L)^{-1}$ em que $V = \text{diag}\{V_1, \dots, V_n\}$ e $L = \text{diag}\{h'_1, \dots, h'_n\}$.

A inferência sobre os parâmetros β e ϕ , baseada no método de máxima verossimilhança, pode ser realizada maximizando (2.4) numericamente. Alternativamente, podemos supor ϕ fixo e utilizar o processo iterativo de Newton-Raphson a fim de obter a estimativa de β . Usando a notação em que (ϕ) explicita a dependência da estimativa de β neste parâmetro, o processo iterativo escoring de Fisher é definido como

$$\beta^{(\phi)(k+1)} = \beta^{(\phi)(k)} + K^{-1}(\beta^{(\phi)(k)}) U(\beta^{(\phi)(k)}), \quad k = 0, 1, \dots$$

Ressaltando que t_i e q_i podem ser reescritos como $t_i = (V_i h'_i)^{-1}$ e $q_i = \mu_i (V_i h'_i)^{-1}$, fazendo com que a matriz T seja expressa como $(VL)^{-1}$ e o vetor Q como $(VL)^{-1} \mu$, em que $\mu = (\mu_1, \dots, \mu_n)^\top$ é um vetor $n \times 1$. Esse processo iterativo pode ser reescrito como um processo de mínimos quadrados ponderados, como se segue:

$$\begin{aligned} \beta^{(\phi)(k+1)} &= \beta^{(\phi)(k)} + (\tilde{X}^{(\phi)(k)\top} W^{(\phi)(k)} \tilde{X}^{(\phi)(k)})^{-1} \tilde{X}^{(\phi)(k)\top} (T^{(\phi)(k)} y - Q^{(\phi)(k)}) \\ &= (\tilde{X}^{(\phi)(k)\top} W^{(\phi)(k)} \tilde{X}^{(\phi)(k)})^{-1} \tilde{X}^{(\phi)(k)\top} W^{(\phi)(k)} \delta^{(\phi)(k)}, \quad k = 0, 1, \dots, \end{aligned} \quad (2.8)$$

em que

$$\begin{aligned} \delta^{(\phi)(k)} &= \tilde{X}^{(\phi)(k)} \beta^{(\phi)(k)} + (W^{(\phi)(k)})^{-1} (T^{(\phi)(k)} y - Q^{(\phi)(k)}) \\ &= \tilde{X}^{(\phi)(k)} \beta^{(\phi)(k)} + (L^{(\phi)(k)} V^{(\phi)(k)} L^{(\phi)(k)}) \left((V^{(\phi)(k)} L^{(\phi)(k)})^{-1} y - (V^{(\phi)(k)} L^{(\phi)(k)})^{-1} \mu^{(\phi)(k+1)} \right) \\ &= \tilde{X}^{(\phi)(k)} \beta^{(\phi)(k)} + L^{(\phi)(k)} (y - \mu^{(\phi)(k+1)}). \end{aligned}$$

Em (2.8), $\delta^{(\phi)(k)}$ desempenha o papel de uma variável dependente modificada, enquanto $W^{(\phi)(k)}$ é uma matriz de pesos que muda a cada passo do processo iterativo. O valor inicial $\beta^{(\phi)(0)}$ pode ser obtido, por exemplo, ajustando um modelo log-não linear. O estimador restrito $\hat{\beta}^{(\phi)}$ tem uma distribuição assintoticamente normal com média β e matriz de covariâncias $(K_\beta)^{-1}$ consistentemente estimada por $(\hat{\tilde{X}}^{(\phi)\top} \hat{W}^{(\phi)} \hat{\tilde{X}}^{(\phi)})^{-1}$.

A estimação do parâmetro ϕ , quando o mesmo é desconhecido, pode ser feita diretamente pelo método da máxima verossimilhança. No entanto, esse método torna-se bastante

complexo para algumas distribuições pertencentes à família (2.1), visto que a função $a(y, \phi)$ usualmente envolve razão de funções gama com argumentos que dependem de ambos y e ϕ e possíveis constantes. Portanto, a equação não-linear para ϕ contém soma de funções digama. Esta dificuldade pode ser evitada estimando o parâmetro ϕ por métodos indiretos. A seguir, descrevemos um desses métodos.

De posse de $\hat{\beta}^{(\phi)}$, a estimativa de ϕ pode ser obtida inserindo $\hat{\mu}_i^{(\phi)} = h^{-1}(\eta(x_i; \hat{\beta}^{(\phi)}))$ em (2.4) e maximizando o logaritmo da função de verossimilhança perfilada $l_p(\phi) = l(\hat{\beta}^{(\phi)}, \phi)$ de ϕ dada por

$$l_p(\phi) = \sum_{i=1}^n \log\{a(y_i, \phi)\} + \sum_{i=1}^n [y_i \log\{g(\hat{\mu}_i^{(\phi)}, \phi)\} - \log\{f(\hat{\mu}_i^{(\phi)}, \phi)\}].$$

2.3.2 Testes da razão de verossimilhanças em MSPNLG

Seja $Y = (Y_1, \dots, Y_n)^\top$ uma amostra aleatória de tamanho n e cujo logaritmo da função de verossimilhança $l(\beta; y)$, dado por (2.4), depende do parâmetro desconhecido $\beta = (\beta_1, \dots, \beta_p)^\top$. Assumimos que $l(\beta; y)$ seja regular com respeito aos componentes de β até quarta ordem.

Considerando que o vetor de parâmetros β pode ser decomposto como $\beta = (\beta_1^\top, \beta_2^\top)^\top$, sendo $\beta_1 = (\beta_1, \dots, \beta_q)^\top$ o vetor de parâmetros de interesse e $\beta_2 = (\beta_{q+1}, \dots, \beta_p)^\top$ o vetor de parâmetros de perturbação. Em muitas situações há interesse em testar hipóteses sobre uma parte do vetor de parâmetros β , digamos $H_0 : \beta_1 = \beta_1^{(0)}$ versus $H_1 : \beta_1 \neq \beta_1^{(0)}$, em que $\beta_1^{(0)}$ é um vetor especificado de dimensão q ($q \leq p$). Podemos definir a estatística da razão de verossimilhanças, assumindo um valor fixo para ϕ , como

$$LR = 2\{l(\hat{\beta}^{(\phi)}; y) - l(\tilde{\beta}^{(\phi)}; y)\}, \quad (2.9)$$

em que $\hat{\beta}^{(\phi)}$ é o estimador de máxima verossimilhança de β sob a hipótese alternativa H_1 e $\tilde{\beta}^{(\phi)} = (\beta_1^{(0)\top}, \tilde{\beta}_2^{(\phi)\top})^\top$ é o estimador correspondente de β sob a hipótese nula H_0 . A estatística da razão de verossimilhanças tem, sob a hipótese nula, distribuição assintótica χ_q^2 . Rejeitamos a hipótese nula H_0 , ao nível de significância α , se $LR > \chi_{(\alpha; q)}^2$, em que $\chi_{(\alpha; q)}^2$ é o percentil $(1 - \alpha)$ da distribuição χ_q^2 .

Capítulo 3

Correção de viés em MSPNLG

3.1 Introdução

O viés estatístico é uma medida de qualidade de um estimador e é calculado como a diferença entre o verdadeiro valor do parâmetro e o valor esperado do estimador em apreço. Em geral, o método da máxima verossimilhança fornece estimadores viesados quando o tamanho de amostra n é pequeno ou quando a informação de Fisher é reduzida. Em alguns casos, o viés pode até ser considerado insignificante quando comparado ao erro-padrão dos EsMV, visto que ele é de ordem n^{-1} , enquanto o desvio padrão da estimativa é de ordem $n^{-1/2}$. Porém, encontrar estimadores corrigidos pelo viés de ordem n^{-2} pode melhorar a qualidade das estimativas, principalmente, em amostras pequenas.

Na literatura, várias propostas sobre a correção de viés vêm sendo estudadas. Bartlett (1953) apresentou uma expressão simples para o viés de ordem n^{-1} do EMV no caso uniparamétrico. Haldane (1953) e Haldane e Smith (1956) desenvolveram expressões de ordem n^{-1} para os primeiros quatros cumulantes em amostras aleatórias de um ou dois parâmetros desconhecidos. Cox e Snell (1968) obtiveram uma expressão geral para o viés de ordem n^{-1} dos EsMV nos casos uniparamétrico e multiparamétrico, supondo observações independentes mas não necessariamente identicamente distribuídas. O viés de ordem n^{-2} em modelos não-lineares em que a matriz de covariâncias é conhecida foi obtido por Box (1971). Cook, Tsai e Wei (1986) analisaram os vieses das estimativas dos resíduos e dos estimadores de máxima

verossimilhança em modelos de regressão não-linear. Estimadores corrigidos em modelos de regressão log-gama generalizada foram apresentados por Young e Bakir (1987).

Cordeiro e McCullagh (1991) derivaram uma fórmula geral para o viés de ordem n^{-2} dos EsMV em modelos lineares generalizados. A expressão do viés de ordem n^{-2} dos EsMV em uma ampla classe de modelos multivariados normais não-lineares foi apresentada por Cordeiro e Vasconcellos (1997). Cordeiro e Vasconcellos (1999) forneceram EsMV corrigidos para o modelo de regressão von Mises. Expressões em forma matricial para o viés de ordem n^{-2} dos EsMV em modelos de regressão multivariado não-linear com erros t de Student e em modelos de regressão não-lineares simétricos foram desenvolvidos por Vasconcellos e Cordeiro (2000) e Cordeiro, Ferrari, Uribe-Opazo e Vasconcellos (2000), respectivamente. Cribari-Neto e Vasconcellos (2002) analisaram o comportamento, em amostras finitas, de três procedimentos alternativos para corrigir o viés de ordem n^{-1} dos EsMV dos parâmetros da distribuição Beta. Vasconcellos e Silva (2005) obtiveram estimadores corrigidos em um modelo de regressão t de Student não-linear quando o número de graus de liberdade é desconhecido. Ospina, Cribari-Neto e Vasconcellos (2006) derivaram expressões em forma fechada para os vieses de ordem n^{-1} dos EsMV em modelo de regressão Beta.

Lemonte *et al.* (2007) desenvolveram estimadores não-viesados para os parâmetros que indexam a distribuição Birnbaum-Saunders. Cordeiro e Barroso (2007) obtiveram uma fórmula matricial para o viés de ordem n^{-2} dos EsMV em modelos lineares generalizados e mostraram, através de simulações de Monte Carlo, que a nova estimativa pode fornecer melhorias substanciais em termos de viés e erro médio quadrático em relação às estimativas EsMV usuais e às estimativas corrigidas propostas por Cordeiro e McCullagh (1991).

Cordeiro *et al.* (2008) desenvolveram correção de viés do EMV em modelos não-lineares superdispersados. Cordeiro e Udo (2008) derivaram fórmulas gerais para o viés de ordem n^{-2} das estimativas de máxima verossimilhança dos parâmetros em modelos não-lineares generalizados com dispersão nas covariáveis. Cordeiro e Demétrio (2008) obtiveram correção de viés no modelo de quasi-verossimilhança estendido. Recentemente, Cordeiro *et al.* (2009) derivaram uma fórmula matricial para o viés de ordem n^{-2} dos estimadores de máxima verossimilhança dos parâmetros de média e variância em modelos não-lineares heterocedásticos. Cysneiros *et al.* (2009) derivaram uma fórmula geral para o viés de ordem n^{-2} dos estimadores de máxima verossimilhança nos modelos de regressão não-lineares simétricos

heterocedásticos.

3.2 Correção de Cox & Snell

Nesta seção, nossa atenção está dirigida à obtenção da correção do viés de estimadores de máxima verossimilhança. Para isto é necessário obter derivadas do logaritmo da função de verossimilhança com relação ao vetor de parâmetros $\theta = (\theta_1, \dots, \theta_p)^\top$ desconhecido e alguns momentos destas derivadas. Assumimos, portanto, no que segue, que tais derivadas e momentos existem. Por outro lado, introduziremos a seguinte notação: $U_r = \partial l / \partial \theta_r$, $U_{rs} = \partial^2 l / \partial \theta_r \partial \theta_s$, $U_{rst} = \partial^3 l / \partial \theta_r \partial \theta_s \partial \theta_t$ e assim por diante, em que l é o logaritmo da função de verossimilhança e b, r, s, t , são indexadores do espaço paramétrico. Consequentemente, os cumulantes conjuntos das derivadas do logaritmo da função de verossimilhança são definidos como $\kappa_{rs} = E(U_{rs})$, $\kappa_{r,s} = E(U_r U_s)$, $\kappa_{rst} = E(U_{rst})$, $\kappa_{rs,t} = E(U_{rs} U_t)$, etc. Denotamos as derivadas dos momentos em relação aos componentes do vetor θ por $\kappa_{rs}^{(t)} = \partial \kappa_{rs} / \partial \theta_t$. A metodologia para encontrar o viés dos EsMV segue o trabalho de Cox & Snell (1968), no qual eles mostraram que para observações independentes, mas não necessariamente identicamente distribuídas, o viés de ordem n^{-1} do EMV $\hat{\theta}_b$ de θ_b é expresso da seguinte forma:

$$B(\hat{\theta}_b) = E(\hat{\theta}_b - \theta_b) = \sum_{r,s,t} \kappa^{br} \kappa^{st} \left(\kappa_{rs,t} + \frac{1}{2} \kappa_{rst} \right), \text{ para } b = 1, 2, \dots, p, \quad (3.1)$$

em que $-\kappa^{rs} = \kappa^{r,s}$ representa o elemento (r, s) da inversa da matriz de informação de Fisher K_θ de θ . Em decorrência da identidade de Bartlett, a saber $\kappa_{rs,t} + \kappa_{rst} - \kappa_{rs}^{(t)} = 0$, a expressão (3.1) pode ser reescrita substituindo o termo $\kappa_{rs,t} + \frac{1}{2} \kappa_{rst}$ pelo termo $\kappa_{rs}^{(t)} - \frac{1}{2} \kappa_{rst}$, ou seja,

$$B(\hat{\theta}_b) = \sum_{r,s,t} \kappa^{br} \kappa^{st} \left\{ \kappa_{rs}^{(t)} - \frac{1}{2} \kappa_{rst} \right\}, \text{ para } b = 1, 2, \dots, p. \quad (3.2)$$

A partir da expressão (3.2), definimos um estimador de máxima verossimilhança corrigido $\tilde{\theta}_b$ para o parâmetro θ_b da seguinte forma:

$$\tilde{\theta}_b = \hat{\theta}_b - \hat{B}(\hat{\theta}_b), \text{ para } b = 1, 2, \dots, p,$$

sendo $\hat{B}(\hat{\theta}_b)$ o estimador de máxima verossimilhança do viés (3.2), ou seja, os parâmetros desconhecidos são substituídos por suas respectivas estimativas de máxima verossimilhança.

Este novo estimador tem viés de ordem n^{-2} , pois $E(\tilde{\theta}_b) = \theta_b + O(n^{-2})$. Consequentemente, esperamos que o estimador corrigido $\tilde{\theta}_b$ tenha melhores propriedades em amostras finitas do que $\hat{\theta}_b$, cujo viés é de ordem n^{-1} .

3.2.1 Correção de viés dos EsMV dos MSPNLGs

O objetivo desta subseção é obter uma expressão geral, em forma fechada, para o viés de ordem n^{-1} dos EsMV dos parâmetros dos MSPNLGs. No Apêndice A são apresentados os cálculos dos cumulantes e as derivadas dos cumulantes necessários para a obtenção de $B(\hat{\beta}_b)$. Por simplicidade, aqui apresentamos somente o processo pelo qual a expressão (3.2) foi conduzida à forma matricial.

Consideremos n variáveis aleatórias discretas independentes $Y_1, \dots, Y_n$ cada qual com função de probabilidade da forma

$$\pi(y; \mu_i, \phi) = \frac{a(y, \phi)g(\mu_i, \phi)^y}{f(\mu_i, \phi)}, \quad y \in A_\epsilon, \quad i = 1, \dots, n \quad (3.3)$$

em que o suporte de Y_i é um subconjunto A_ϵ dos inteiros $\{\epsilon, \epsilon + 1, \dots\}$, $\epsilon \geq 0$, e que não depende de parâmetros desconhecidos, $a(y, \phi)$ é uma função positiva, as funções analíticas $f_i = f(\mu_i, \phi)$ e $g_i = g(\mu_i, \phi)$ são positivas, finitas e duas vezes diferenciáveis, $\phi > 0$ e $\mu_i > 0$ são chamados de parâmetros de dispersão e de média, respectivamente. Para a família de distribuições dada em (3.3) as seguintes relações são válidas:

$$E(Y) = \mu = \frac{f^{(1)}g}{fg^{(1)}} \quad \text{e} \quad Var(Y) = V(\mu, \phi) = \frac{g}{g^{(1)}},$$

em que o índice sobrescrito (1) indica a primeira diferenciação em relação a μ . Os modelos em séries de potência não-lineares generalizados são definidos por (3.3) e pelo componente sistemático

$$h(\mu_i) = \eta_i = \eta(x_i; \beta), \quad i = 1, \dots, n,$$

em que $h(\cdot)$ é uma função de ligação conhecida e duplamente diferenciável, $\beta = (\beta_1, \dots, \beta_p)^\top$ é um vetor de p ($p < n$) parâmetros desconhecidos a serem estimados, $x_i = (x_{i1}, \dots, x_{ik})^\top$ representa os valores de k variáveis explicativas e $\eta(\cdot; \cdot)$ é uma função possivelmente não-linear no segundo argumento, contínua e diferenciável com respeito aos componentes de β

tal que a matriz de derivadas $\tilde{X} = \tilde{X}(\beta) = \partial\eta/\partial\beta^\top$, com $\eta = (\eta_1, \dots, \eta_n)^\top$, tem posto p para todo β . A matriz $\tilde{X}$ tem elementos que são, em geral, funções do vetor de parâmetros β desconhecidos.

O logaritmo da função de verossimilhança do vetor dos parâmetros β , dado o vetor de observações $(y_1, \dots, y_n)$, dos MSPNLGs pode ser expresso na forma

$$l(\beta; y) = \sum_{i=1}^n \log\{a(y_i, \phi)\} + \sum_{i=1}^n [y_i \log\{g(\mu_i, \phi)\} - \log\{f(\mu_i, \phi)\}].$$

Assumimos que a função $l(\beta; y)$ é regular com respeito às derivadas em relação aos componentes de β até a quarta ordem. Para a obtenção dessas derivadas, utilizamos a notação $\tilde{x}_{ir} = \partial\eta_i/\partial\beta_r$, $\tilde{x}_{irs} = \partial^2\eta_i/\partial\beta_r\partial\beta_s$, $\tilde{x}_{irst} = \partial^3\eta_i/\partial\beta_r\partial\beta_s\partial\beta_t$, etc. Assumindo que ϕ é conhecido, temos que as três primeiras derivadas de $l(\beta; y)$ são

$$\begin{aligned} U_r &= \sum_{i=1}^n d_{0i} \tilde{x}_{ir}, \\ U_{rs} &= \sum_{i=1}^n \left\{ d_{1i} \tilde{x}_{is} \tilde{x}_{ir} + d_{0i} \tilde{x}_{irs} \right\} \text{ e} \\ U_{rst} &= \sum_{i=1}^n \left\{ \left[d_{2i} - \frac{d_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + d_{1i} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs}) + d_{0i} \tilde{x}_{irst} \right\}, \end{aligned}$$

em que

$$d_{0i} = y_i t_i - q_i \text{ e } d_{ji} = \frac{y_i t_i^{(j)} - q_i^{(j)}}{(h_i^{(1)})^j},$$

com $t_i = \frac{g_i^{(1)}}{g_i h_i^{(1)}}$, $q_i = \frac{f_i^{(1)}}{f_i h_i^{(1)}}$ e o índice sobrescrito (j) indicando a j -ésima derivada em relação μ com $j = 1, 2$ e $i = 1, \dots, n$. Tomando as esperanças das duas últimas expressões, encontramos:

$$\begin{aligned} \kappa_{rs} &= \sum_{i=1}^n w_{1i} \tilde{x}_{is} \tilde{x}_{ir} \text{ e} \\ \kappa_{rst} &= \sum_{i=1}^n \left\{ \left[w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + w_{1i} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs}) \right\}, \end{aligned}$$

sendo a derivada do primeiro cumulante dada por

$$\kappa_{rs}^{(t)} = \sum_{i=1}^n \tilde{w}_{1i} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + \sum_i w_{1i} \tilde{x}_{ist} \tilde{x}_{ir} + \sum_{i=1}^n w_{1i} \tilde{x}_{is} \tilde{x}_{irt},$$

em que w_{ji} e $\tilde{w}_{ji}$ são escalares definidos, respectivamente, por

$$w_{ji} = \left(\frac{f_i^{(1)} g_i t_i^{(j)}}{f_i g_i^{(1)} t_i^{(j)}} - q_i^{(j)} \right) \frac{1}{h_i^{(1)}} \quad \text{e}$$

$$\tilde{w}_{ji} = \varphi_{ji} - \frac{(j-1) q_i V_i t_i^{(j)} h_i^{(2)} - q_i^{(j+1)}}{(h_i^{(1)})^{j+1}} + j \frac{q_i^{(j)} h_i^{(2)}}{(h_i^{(1)})^{j+2}},$$

com

$$\varphi_{ji} = \frac{q_i^{(1)} V_i t_i^{(j)} + q_i V_i^{(1)} t_i^{(j)} + q_i V_i t_i^{(j+1)}}{(h_i^{(1)})^j},$$

$j = 1, 2$ e $i = 1, \dots, n$. Vale ressaltar que as quantidades acima envolvem derivadas que dependem das formas específicas das funções f , g , h e V nas diversas distribuições pertencentes à família de série de potência. Temos, então, que a quantidade $\kappa_{rs}^{(t)} - \frac{1}{2} \kappa_{rst}$ dada na expressão (3.2) tem a seguinte forma:

$$\begin{aligned} \kappa_{rs}^{(t)} - \frac{1}{2} \kappa_{rst} &= \sum_{i=1}^n \left\{ \tilde{w}_{1i} - \frac{1}{2} \left[w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \right\} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + \frac{1}{2} \sum_{i=1}^n w_{1i} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} - \tilde{x}_{it} \tilde{x}_{irs}) \\ &= \sum_{i=1}^n c_i \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + \frac{1}{2} \sum_{i=1}^n w_{1i} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} - \tilde{x}_{it} \tilde{x}_{irs}), \end{aligned}$$

sendo $c_i = \tilde{w}_{1i} - \frac{1}{2} \{w_{2i} - w_{1i} h_i^{(2)} (h_i^{(1)})^{-2}\}$, $i = 1, \dots, n$. Então, a partir da expressão (3.2) obtemos a fórmula para calcular o viés de segunda ordem da b -ésima componente de $\hat{\beta}$, a qual é dada por

$$B(\hat{\beta}_b) = \sum_{r,s,t} \kappa^{br} \kappa^{st} \sum_i c_i \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + \frac{1}{2} \sum_{r,s,t} \kappa^{br} \kappa^{st} \sum_i w_{1i} \tilde{x}_{ist} \tilde{x}_{ir}, \quad b = 1, \dots, p,$$

em que r, s e t variam em $E_S = \{1, \dots, p\}$, o índice i varre todas as observações e $-\kappa^{rs} = \kappa^{r,s}$ representa o elemento (r, s) da inversa da matriz de informação de Fisher K_β de β , dada por

$$K_\beta = E \left\{ - \frac{\partial^2 l(\beta)}{\partial \beta \partial \beta^\top} \right\} = \tilde{X}^\top W \tilde{X},$$

em que W é uma matriz diagonal $n \times n$ de pesos dados por

$$w_i = \left(q_i^{(1)} - \frac{f_i^{(1)} g_i}{f_i g_i^{(1)}} t_i^{(1)} \right) \frac{1}{h_i^{(1)}}, \quad i = 1, \dots, n.$$

Portanto, o viés de ordem n^{-2} pode ser escrito, em notação matricial, como

$$B(\hat{\beta}) = (\tilde{X}^\top W \tilde{X})^{-1} \tilde{X}^\top \delta, \quad (3.4)$$

em que $\delta = (Z_d C + \frac{1}{2} D W_1)$, com Z_d e D sendo matrizes diagonais de ordem $n \times n$, em que a diagonal da primeira é igual à diagonal da matriz $Z = \tilde{X}(\tilde{X}^\top W \tilde{X})^{-1} \tilde{X}^\top$ e o i -ésimo elemento da segunda é igual ao traço de $(\tilde{X}^\top W \tilde{X})^{-1} \tilde{\tilde{X}}_i$, em que $\tilde{\tilde{X}}_i$ é uma matriz $p \times p$ cujo elemento (r, s) é $\tilde{x}_{irs}$, C e W_1 são vetores coluna de ordem n com respectivos elementos c_i e w_{1i} , $i = 1, \dots, n$, dados acima. Observamos, portanto, que $B(\hat{\beta})$ pode ser obtido através de uma regressão de mínimos quadrados ponderados.

Na construção da matriz W e dos vetores C e W_1 , presentes na equação (3.4), necessitamos da função de ligação e das funções t e q com suas respectivas primeiras e segundas derivadas, das funções f , g e de variância com suas primeiras derivadas, respectivamente. Para obtermos as matrizes Z , Z_d e D precisamos da matriz modelo $\tilde{X}$ e das matrizes quadradas $\tilde{\tilde{X}}_i$, $i = 1, \dots, n$. Uma vez computadas as matrizes acima, o cálculo de $B(\hat{\beta})$ é imediato. É óbvio que a expressão (3.4) depende muito do particular modelo adotado.

Se $\eta(\cdot; \cdot)$ for linear temos que $\tilde{\tilde{X}}_i = 0$ e, conseqüentemente, $\delta = Z_d C$. Este resultado mostra que para os modelos pertencentes tanto à classe dos modelos em séries de potência lineares generalizados quanto à classe dos modelos lineares generalizados, a fórmula matricial dada em (3.4) coincide com a fórmula de Cordeiro e McCullagh (1991, p. 635, equação 4.2).

3.3 Correção via bootstrap

Uma forma alternativa de se obter a correção de viés é através da técnica bootstrap. Este é um método computacionalmente intensivo, introduzido por Bradley Efron em 1979 (Efron, 1979) e utilizado para obter soluções aproximadas de problemas estatísticos cujas soluções analíticas são complicadas ou desconhecidas.

Considere uma amostra aleatória $y = (y_1, \dots, y_n)^\top$ de uma variável Y , cuja distribuição está completamente determinada por sua função de distribuição $F = F_Y(y)$, e indexada pelo

parâmetro desconhecido $\theta = \tau(F)$. Seja $\hat{\theta} = s(y)$ uma função da amostra aleatória y e um estimador para θ . Uma vez que é impossível extrair repetidas amostras da população descrita pela função de distribuição desconhecida F , a idéia é obter a partir da amostra original, um grande número de subamostras $y^* = (y_1^*, \dots, y_n^*)^\top$ com base numa estimativa $\hat{F}$ de F , visando com isso obter um estimador de $\theta = \tau(F)$, denotado por $\hat{\theta} = \tau(\hat{F})$. A reamostragem pode ser feita de forma paramétrica ou não-paramétrica.

Na versão não-paramétrica, a reamostragem é feita retirando-se uma amostra a partir de uma estimativa não-paramétrica $\hat{F}_n$ de F , que é a função de distribuição empírica da amostra original, a qual atribui probabilidade $1/n$ a cada y_i , $i = 1, \dots, n$, isto é,

$$\hat{F}_n(y) = \frac{\#\{y_i \leq y\}}{n},$$

que representa a proporção amostral de valores observados menores ou iguais a y . Na versão paramétrica, quando se conhece previamente o modelo paramétrico F_ξ ao qual F pertence, a amostra y^* é formada realizando-se a amostragem com base na estimativa paramétrica $\hat{F}_\xi$, em que os parâmetros desconhecidos são substituído por suas respectivas estimativas paramétricas.

A implementação do método bootstrap permite a estimação do erro padrão, viés, variâncias, intervalos de confiança e outras quantidades de interesse da inferência estatística. Nesta subseção, no entanto, nosso foco será na obtenção da estimativa de viés por bootstrap. Com essa finalidade, denote por $B_F(\hat{\theta}, \theta)$ o viés do estimador $\hat{\theta}$, ou seja,

$$B_F(\hat{\theta}, \theta) = E_F[s(y)] - \tau(F), \quad (3.5)$$

em que o subscrito F indica que a esperança matemática é calculada com base na função de distribuição F . Substituindo F por suas respectivas estimativas bootstrap em (3.5), obtemos os estimadores bootstrap para o viés nas versões não-paramétrica e paramétrica, respectivamente, dados por

$$B_{\hat{F}_n}(\hat{\theta}, \theta) = E_{\hat{F}_n}[s(y)] - \tau(\hat{F}_n) \text{ e } B_{\hat{F}_\xi}(\hat{\theta}, \theta) = E_{\hat{F}_\xi}[s(y)] - \tau(\hat{F}_\xi).$$

A partir da geração de N subamostras $y^{*1}, \dots, y^{*N}$ podemos ter uma aproximação para $E_{\hat{F}_\xi}[s(y)]$ e $E_{\hat{F}_n}[s(y)]$, por meio da média aritmética das respectivas réplicas bootstrap $\hat{\theta}^{*1}, \dots, \hat{\theta}^{*N}$, onde $\hat{\theta}^{*i} = s(y^{*i})$, $i = 1, \dots, N$, ou seja,

$$\hat{\theta}^{*(\cdot)} = \frac{1}{N} \sum_{i=1}^N \hat{\theta}^{*i}.$$

Com isso, as estimativas do viés via bootstrap não-paramétrico e paramétrico são, respectivamente, dadas por

$$\hat{B}_{\hat{F}_n}(\hat{\theta}, \theta) = \hat{\theta}^{*(\cdot)} - s(y) \text{ e } \hat{B}_{\hat{F}_\xi}(\hat{\theta}, \theta) = \hat{\theta}^{*(\cdot)} - s(y),$$

as quais diferenciam-se na forma de obter as subamostras. Tendo em mãos as estimativas do viés, podemos definir estimadores corrigidos até segunda ordem de θ da forma

$$\begin{aligned} \check{\theta}_1 &= s(y) - \hat{B}_{\hat{F}_n}(\hat{\theta}, \theta) = 2s(y) - \hat{\theta}^{*(\cdot)} \text{ e} \\ \check{\theta}_2 &= s(y) - \hat{B}_{\hat{F}_\xi}(\hat{\theta}, \theta) = 2s(y) - \hat{\theta}^{*(\cdot)}. \end{aligned}$$

3.4 Resultados numéricos

Nesta seção, o objetivo é comparar, por meio do método de simulação de Monte Carlo, os desempenhos dos estimadores de máxima verossimilhança dos parâmetros que indexam os modelos em séries de potências lineares e não-lineares generalizados e das suas versões corrigidas via Cox & Snell e por bootstrap paramétrico, em amostras de tamanho finito e sob diferentes cenários, tanto no que se refere aos vieses quanto à eficiência. Para isso foram calculados medidas de qualidade para a estimação pontual como: viés, viés relativo e erro quadrático médio (EQM). O viés relativo é definido como $100 \times (\text{viés} / \text{valor verdadeiro do parâmetro})\%$. Os estimadores avaliados foram: estimador de máxima verossimilhança ($\hat{\beta}$), estimador de máxima verossimilhança corrigido via correção de Cox & Snell ($\tilde{\beta}$) e corrigido via bootstrap ($\check{\beta}$).

Formulamos um experimento de simulação de Monte Carlo baseado em 10000 réplicas considerando os seguintes tamanhos amostrais: $n = 25, 35, 45, 100$. Para cada réplica de Monte Carlo, consideramos $B = 600$ réplicas bootstrap. Foram selecionadas três distribuições da classe MSPNLG a saber, Binomial Negativa Generalizada (BNG), Poisson

Generalizada (GPO) e Consul. Os resultados numéricos basearam-se nos seguintes preditores:

$$\eta_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} \quad \text{e} \quad (3.6)$$

$$\eta_i = \beta_0 + \beta_1 x_{1i} + \exp(\beta_2 x_{2i}), \quad (3.7)$$

em que $i = 1, \dots, n$ e as covariáveis foram tomadas como amostras aleatórias da distribuição uniforme $U(0, 1)$. No caso da distribuição BNG, fixamos os parâmetros $\phi = 1, 5$ e $\nu = 5$, já para a GPO e a Consul fixamos $\phi = 0, 2$ e $\phi = 1, 0$, respectivamente. Os valores verdadeiros considerados para o vetor de parâmetros $\beta = (\beta_0, \beta_1, \beta_2)^\top$ variaram entre 0, 25, 0, 5, 0, 75 e 1. O processo de simulação foi realizado utilizando a linguagem matricial de programação 0x (Doornik, 2001).

3.4.1 Modelos lineares

Nesta seção, daremos ênfase aos modelos em séries de potências lineares generalizados, cujo preditor linear é dado em (3.6).

Inicialmente, temos como objetivo analisar a influência do tamanho da amostra no desempenho dos estimadores. Os resultados apresentados nas Tabelas 3.1, 3.2 e 3.3 correspondem às estimativas do viés do vetor de parâmetros β considerando amostras de diferentes tamanhos e a variável resposta proveniente das distribuições BNG, Consul e GPO, respectivamente, tendo sido geradas amostras assumindo que $\beta_0 = \beta_1 = \beta_2 = 0, 25$. Notamos que, em todos os modelos, as estimativas dos vieses das versões corrigidas do estimador de máxima verossimilhança, $\tilde{\beta}$ e $\check{\beta}$, são, em módulo, menores do que as correspondentes estimativas $\hat{\beta}$, independentes do tamanho amostral, com apenas uma exceção no modelo Consul para $n = 35$, no qual $\check{\beta}_1$ apresentou uma estimativa de viés, em módulo, igual a 0,01369, enquanto que no mesmo contexto $\hat{\beta}_1$ forneceu uma estimativa igual a 0,00970. Entre os estimadores $\tilde{\beta}$ e $\check{\beta}$, o estimador $\tilde{\beta}$ foi o mais eficaz no sentido de que, na maioria das vezes, forneceu a menor estimativa de viés em valor absoluto. Para o modelo BNG e $n = 25$, por exemplo, as estimativas dos vieses, em módulo, para o vetor de parâmetros β provenientes do estimador $\tilde{\beta}$ foram 0,00384, 0,00676 e 0,00212, ao passo que as provenientes do estimador $\check{\beta}$ foram 0,00611, 0,01382 e 0,00939. As estimativas do viés relativo e do erro quadrático

médio fornecidas pelos novos estimadores, $\tilde{\beta}$ e $\check{\beta}$, refletem o ganho de precisão conseguido pelas correções feitas no estimador de máxima verossimilhança. Este fato é mais notório nas estimativas referentes ao parâmetro β_0 do que naquelas referentes à β_1 e β_2 . Em relação ao tamanho da amostra, como era de se esperar, vemos que todos estimadores apresentam uma melhora nas estimativas à medida que o tamanho de amostra cresce.

Analizamos também o comportamento dos estimadores para diferentes valores de β . Neste caso, fixamos o tamanho da amostra em $n = 35$, os valores de $\beta_0 = \beta_1 = 0,25$ e variamos o valor de β_2 . As Tabelas 3.4, 3.5 e 3.6 apresentam esses resultados para os modelos BNG, Consul e GPO, respectivamente. Observamos que as estimativas, em módulo, apresentadas pelo estimador $\tilde{\beta}$ são bem melhores do que os demais estimadores em todos os modelos. No modelo Consul, por exemplo, quando $\beta_2 = 0,5$, as estimativas dos vieses de $\tilde{\beta}$ para o vetor de parâmetros β são, em módulo, iguais à 0,00160, 0,00662 e 0,00273, enquanto que para $\hat{\beta}$ temos 0,07738, 0,01256 e 0,01659 e para $\check{\beta}$, 0,01424, 0,01168 e 0,00592. É notório também que à medida que aumentamos o valor de β_2 , para este mesmo parâmetro há uma redução, em módulo, nas estimativas do viés relativo produzidas por todos estimadores. No entanto, essa redução nas estimativas do viés relativo não é o bastante para dispensar o uso das correções. Um fato importante que observamos é que as estimativas do viés produzidas por $\check{\beta}$ para o parâmetro β_1 são, em módulo, na maioria das vezes, superiores às estimativas correspondente de $\hat{\beta}$.

Observamos também que, em todos os modelos, o estimador $\hat{\beta}$ subestima β_0 e superestima β_1 e β_2 . As correções feitas no estimador de máxima verossimilhança corrige essa tendência de diferentes maneiras. Enquanto que a correção de Cox & Snell leva o estimador à fornecer vieses relativos, na maioria das vezes, positivos para todos os parâmetros, a correção por bootstrap faz com que o estimador forneça vieses relativos sempre positivos para β_0 e vieses relativos quase sempre negativos para β_1 e β_2 .

Tabela 3.1: Resultados da estimação pontual do vetor β no modelo linear Binomial Negativa Generalizada indexado pelos parâmetros $\beta_0 = \beta_1 = \beta_2 = 0, 25, \phi = 1, 5$ e $\nu = 5$ para $n = 25, 35, 45$ e 100 .

n	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
25	$\hat{\beta}$	-0,08218	-32,8720	0,00675	0,02986	11,9430	0,00089	0,01050	4,1986	0,00011
	$\tilde{\beta}$	0,00384	1,5348	0,00001	0,00676	2,7041	0,00005	-0,00212	-0,8494	0,00000
	$\check{\beta}$	0,00611	2,4454	0,00004	-0,01382	-5,5268	0,00019	0,00939	3,7548	0,00009
35	$\hat{\beta}$	-0,05286	-21,1450	0,00279	0,01016	4,0647	0,00010	0,01173	4,6919	0,00014
	$\tilde{\beta}$	0,00151	0,6036	0,00000	0,00516	2,0652	0,00003	0,00305	1,2203	0,00001
	$\check{\beta}$	0,01045	4,1786	0,00011	-0,00976	-3,9047	0,00010	-0,00364	-1,4552	0,00001
45	$\hat{\beta}$	-0,04031	-16,1240	0,00163	0,00905	3,6184	0,00008	0,00356	1,4247	0,00001
	$\tilde{\beta}$	0,00123	0,4935	0,00000	0,00362	1,4476	0,00001	-0,00099	-0,3962	0,00000
	$\check{\beta}$	0,00452	1,8092	0,00002	-0,00410	-1,6413	0,00002	-0,00050	-0,1984	0,00000
100	$\hat{\beta}$	-0,02029	-8,1178	0,00041	0,00395	1,5799	0,00002	0,00663	2,6513	0,00004
	$\tilde{\beta}$	-0,00064	-0,2559	0,00000	0,00150	0,5988	0,00000	0,00212	0,8470	0,00000
	$\check{\beta}$	0,00116	0,4655	0,00000	-0,00009	-0,0358	0,00000	0,00166	0,6636	0,00000

Tabela 3.2: Resultados da estimação pontual do vetor β no modelo linear Consul indexado pelos parâmetros $\beta_0 = \beta_1 = \beta_2 = 0, 25$ e $\phi = 1, 0$ para $n = 25, 35, 45$ e 100 .

n	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
25	$\hat{\beta}$	-0,12453	-49,8120	0,01551	0,03815	15,2590	0,00146	0,01729	6,9174	0,00030
	$\tilde{\beta}$	0,00487	1,9484	0,00002	0,00537	2,1480	0,00003	-0,00026	-0,1047	0,00000
	$\check{\beta}$	0,01549	6,1973	0,00024	-0,01824	-7,2953	0,00033	0,00636	2,5448	0,00004
35	$\hat{\beta}$	-0,07787	-31,1480	0,00606	0,00970	3,8816	0,00009	0,01484	5,9340	0,00022
	$\tilde{\beta}$	0,00352	1,4091	0,00001	0,00376	1,5041	0,00001	0,00302	1,2074	0,00001
	$\check{\beta}$	0,01613	6,4515	0,00026	-0,01369	-5,4761	0,00019	-0,00579	-2,3145	0,00003
45	$\hat{\beta}$	-0,05931	-23,7220	0,00352	0,01134	4,5372	0,00013	0,00351	1,4026	0,00001
	$\tilde{\beta}$	0,00280	1,1206	0,00001	0,00435	1,7406	0,00002	-0,00233	-0,9338	0,00001
	$\check{\beta}$	0,00641	2,5659	0,00004	-0,00595	-2,3791	0,00004	0,00030	0,1187	0,00000
100	$\hat{\beta}$	-0,02988	-11,9520	0,00089	0,00436	1,7457	0,00002	0,00860	3,4393	0,00007
	$\tilde{\beta}$	-0,00047	-0,1887	0,00000	0,00117	0,4693	0,00000	0,00216	0,8651	0,00000
	$\check{\beta}$	0,00228	0,9129	0,00001	-0,00123	-0,4935	0,00000	0,00091	0,3658	0,00000

Tabela 3.3: Resultados da estimação pontual do vetor β no modelo linear Poisson Generalizada indexado pelos parâmetros $\beta_0 = \beta_1 = \beta_2 = 0, 25$ e $\phi = 0, 2$ para $n = 25, 35, 45$ e 100 .

n	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
25	$\hat{\beta}$	-0,08327	-33,3070	0,00693	0,03017	12,0680	0,00091	0,01078	4,3104	0,00012
	$\tilde{\beta}$	0,00394	1,5750	0,00002	0,00695	2,7814	0,00005	-0,00185	-0,7399	0,00000
	$\check{\beta}$	0,00628	2,5125	0,00004	-0,01407	-5,6277	0,00020	0,00978	3,9119	0,00010
35	$\hat{\beta}$	-0,05372	-21,4880	0,00289	0,00993	3,9714	0,00010	0,01229	4,9140	0,00015
	$\tilde{\beta}$	0,00137	0,5484	0,00000	0,00503	2,0135	0,00003	0,00362	1,4472	0,00001
	$\check{\beta}$	0,01072	4,2862	0,00011	-0,00960	-3,8409	0,00009	-0,00425	-1,7008	0,00002
45	$\hat{\beta}$	-0,04084	-16,3350	0,00167	0,00912	3,6490	0,00008	0,00371	1,4835	0,00001
	$\tilde{\beta}$	0,00125	0,4999	0,00000	0,00375	1,4985	0,00001	-0,00080	-0,3198	0,00000
	$\check{\beta}$	0,00414	1,6559	0,00002	-0,00384	-1,5347	0,00001	-0,00006	-0,0249	0,00000
100	$\hat{\beta}$	-0,02063	-8,2499	0,00043	0,00383	1,5326	0,00001	0,00694	2,7752	0,00005
	$\tilde{\beta}$	-0,00071	-0,2829	0,00000	0,00140	0,5582	0,00000	0,00240	0,9593	0,00001
	$\check{\beta}$	0,00145	0,5805	0,00000	-0,00034	-0,1362	0,00000	0,00132	0,5295	0,00000

Tabela 3.4: Resultados da estimação pontual do vetor β no modelo linear Binomial Negativa Generalizada indexado pelos parâmetros $\beta_0 = \beta_1 = 0, 25, \beta_2 = 0, 25, 0, 5, 0, 75, 1, \phi = 1, 5$ e $\nu = 5$ para $n = 35$.

	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
β_2	$\hat{\beta}$	-0,05286	-21,1450	0,00279	0,01016	4,0647	0,00010	0,01173	4,6919	0,00014
0,25	$\tilde{\beta}$	0,00151	0,6036	0,00000	0,00516	2,0652	0,00003	0,00305	1,2203	0,00001
	$\check{\beta}$	0,01549	6,1973	0,00024	-0,01824	-7,2953	0,00033	0,00636	2,5448	0,00004
0,5	$\hat{\beta}$	-0,05088	-20,3530	0,00259	0,01023	4,0924	0,00010	0,01231	2,4610	0,00015
	$\tilde{\beta}$	0,00129	0,5162	0,00000	0,00539	2,1543	0,00003	0,00197	0,3934	0,00000
	$\check{\beta}$	0,01424	5,6941	0,00020	-0,01168	-4,6705	0,00014	-0,00592	-1,1833	0,00004
0,75	$\hat{\beta}$	-0,04947	-19,7890	0,00245	0,00968	3,8725	0,00009	0,01472	1,9622	0,00022
	$\tilde{\beta}$	0,00067	0,2697	0,00000	0,00506	2,0231	0,00003	0,00334	0,4460	0,00001
	$\check{\beta}$	0,01301	5,2048	0,00017	-0,01256	-5,0257	0,00016	-0,00448	-0,5976	0,00002
1	$\hat{\beta}$	-0,04738	-18,9530	0,00225	0,00912	3,6462	0,00008	0,01414	1,4141	0,00020
	$\tilde{\beta}$	0,00093	0,3707	0,00000	0,00475	1,9011	0,00002	0,00229	0,2290	0,00001
	$\check{\beta}$	0,01327	5,3072	0,00018	-0,01322	-5,2895	0,00017	-0,00527	-0,5268	0,00003

Tabela 3.5: Resultados da estimação pontual do vetor β no modelo linear Consul indexado pelos parâmetros $\beta_0 = \beta_1 = 0, 25, \beta_2 = 0, 25, 0, 5, 0, 75, 1$ e $\phi = 1, 0$ para $n = 35$.

	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
β_2	$\hat{\beta}$	-0,07787	-31,1480	0,00606	0,00970	3,8816	0,00009	0,01484	5,9340	0,00022
0,25	$\tilde{\beta}$	0,00352	1,4091	0,00001	0,00376	1,5041	0,00001	0,00302	1,2074	0,00001
	$\check{\beta}$	0,01613	6,4515	0,00026	-0,01369	-5,4761	0,00019	-0,00579	-2,3145	0,00003
0,5	$\hat{\beta}$	-0,07738	-30,9540	0,00599	0,01256	5,0231	0,00016	0,01659	3,3172	0,00028
	$\tilde{\beta}$	0,00160	0,6406	0,00000	0,00662	2,6498	0,00004	0,00273	0,5459	0,00001
	$\check{\beta}$	0,01424	5,6941	0,00020	-0,01168	-4,6705	0,00014	-0,00592	-1,1833	0,00004
0,75	$\hat{\beta}$	-0,07505	-30,0210	0,00563	0,01195	4,7787	0,00014	0,01787	2,3833	0,00032
	$\tilde{\beta}$	0,00166	0,6655	0,00000	0,00614	2,4550	0,00004	0,00256	0,3412	0,00001
	$\check{\beta}$	0,01301	5,2048	0,00017	-0,01256	-5,0257	0,00016	-0,00448	-0,5976	0,00002
1	$\hat{\beta}$	-0,07289	-29,1540	0,00531	0,01097	4,3865	0,00012	0,01904	1,9037	0,00036
	$\tilde{\beta}$	0,00174	0,6969	0,00000	0,00530	2,1191	0,00003	0,00272	0,2720	0,00001
	$\check{\beta}$	0,01327	5,3072	0,00018	-0,01322	-5,2895	0,00017	-0,00527	-0,5268	0,00003

Tabela 3.6: Resultados da estimação pontual do vetor β no modelo linear Poisson Generalizada indexado pelos parâmetros $\beta_0 = \beta_1 = 0, 25, \beta_2 = 0, 25, 0, 5, 0, 75, 1$ e $\phi = 0, 2$ para $n = 35$.

β_2	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
0,25	$\hat{\beta}$	-0,05372	-21,4880	0,00289	0,00993	3,9714	0,00010	0,01229	4,9140	0,00015
	$\tilde{\beta}$	0,00137	0,5484	0,00000	0,00503	2,0135	0,00003	0,00362	1,4472	0,00001
	$\check{\beta}$	0,01072	4,2862	0,00011	-0,00960	-3,8409	0,00009	-0,00425	-1,7008	0,00002
0,5	$\hat{\beta}$	-0,05142	-20,5690	0,00264	0,01043	4,1701	0,00011	0,01216	2,4317	0,00015
	$\tilde{\beta}$	0,00151	0,6032	0,00000	0,00571	2,2835	0,00003	0,00193	0,3865	0,00000
	$\check{\beta}$	0,01073	4,2915	0,00012	-0,01115	-4,4600	0,00012	-0,00394	-0,7881	0,00002
0,75	$\hat{\beta}$	-0,04999	-19,9960	0,00250	0,00885	3,5389	0,00008	0,01445	1,9272	0,00021
	$\tilde{\beta}$	0,00098	0,3915	0,00000	0,00439	1,7579	0,00002	0,00330	0,4405	0,00001
	$\check{\beta}$	0,00890	3,5615	0,00008	-0,00903	-3,6119	0,00008	-0,00295	-0,3938	0,00001
1	$\hat{\beta}$	-0,04818	-19,2740	0,00232	0,00860	3,4399	0,00007	0,01452	1,4517	0,00021
	$\tilde{\beta}$	0,00101	0,4059	0,00000	0,00444	1,7764	0,00002	0,00303	0,3030	0,00001
	$\check{\beta}$	0,00781	3,1221	0,00006	-0,00848	-3,3926	0,00007	-0,00200	-0,1996	0,00000

3.4.2 Modelos não-lineares

Nesta seção, enfatizaremos os modelos em séries de potências não-lineares generalizados, cujo preditor não-linear é dado em (3.7).

Os resultados apresentados nas Tabelas 3.7, 3.8 e 3.9 correspondem às estimativas do vetor de parâmetro β nos modelos BNG, Consul e GPO, respectivamente, nos quais foram consideradas amostras de diferentes tamanhos. No caso dos modelos BNG e GPO, a variável resposta foi gerada assumindo que $\beta_0 = \beta_1 = \beta_2 = 0,25$, já no caso do modelo Consul assumimos que $\beta_0 = \beta_1 = 0,25$ e $\beta_2 = 1$. Os resultados mostram que as estimativas dos vieses das versões corrigidas do estimador de máxima verossimilhança, $\tilde{\beta}$ e $\check{\beta}$ são, em módulo, menores do que as estimativas de $\hat{\beta}$, independente do tamanho amostral, com exceção de $\check{\beta}_0$, em $n = 25$, o qual apresentou, nos modelos BNG e GPO, estimativas de viés iguais à 0,02302 e 0,02582, respectivamente, superiores às estimativas correspondentes fornecidas por $\hat{\beta}_0$, as quais foram, em módulo, iguais à 0,01346 e 0,01477. O ganho de precisão conseguido pelos estimadores $\tilde{\beta}$ e $\check{\beta}$ é refletido nas estimativas do viés relativo e do erro quadrático médio proveniente desses estimadores. Nos modelos BNG e GPO, entre os estimadores $\tilde{\beta}$ e $\check{\beta}$ não há indício que um tenha um melhor comportamento que o outro. Por outro lado, no modelo Consul, o estimador $\tilde{\beta}$ mostrou-se mais eficiente do que o $\check{\beta}$ ao apresentar, na maioria dos casos, menores estimativas, em módulo, do viés, do viés relativo e do erro quadrático médio. Observamos também que todos os estimadores tornam-se mais eficientes à medida que o tamanho da amostra cresce, conforme era esperado.

Para avaliar o desempenho dos estimadores quando o vetor de parâmetros β assume diferentes valores, fixamos o tamanho de amostra em $n = 35$, os valores de $\beta_0 = \beta_1 = 0,25$ e variamos β_2 . As Tabelas 3.10, 3.11 e 3.12 apresentam esses resultados para os modelos BNG, Consul e GPO, respectivamente. Dessa vez, os resultados mostram que, em todos os modelos, o estimador $\tilde{\beta}$ é bem mais preciso do que os demais estimadores. No modelo GPO, quando $\beta_2 = 0,5$, por exemplo, as estimativas do viés de $\tilde{\beta}$ são, em módulo, iguais a 0,00029, 0,00487 e 0,02325, enquanto para $\hat{\beta}$ temos 0,02128, 0,00967 e 0,08210 e para $\check{\beta}$, 0,00581, 0,00745 e 0,05261. Assim como nos modelos lineares, quando o valor do parâmetro β_2 cresce, observamos que, para este mesmo parâmetro, há uma redução nos valores absolutos das estimativas do viés relativo fornecidas pelos estimadores em estudo. Vemos também que o estimador $\hat{\beta}$ tem uma tendência de subestimar β_0 e β_2 e superestimar β_1 , enquanto que os

estimadores corrigidos apresentam tendências diferentes. O estimador $\tilde{\beta}$ fornece, na maioria das vezes, vieses relativos positivos para todos os parâmetros e $\check{\beta}$, por sua vez, fornece vieses relativos quase sempre positivos para β_0 e β_2 , e quase sempre negativos para β_1 .

Outro estudo de simulação utilizando modelos não-lineares foi realizado com o objetivo de analisar o desempenho dos estimadores em diferentes distribuições. Para isso, utilizamos os dados analisados por Previdelli (2005), os quais foram obtidos durante um teste de aprendizagem e memória espacial aplicado em ratos portadores de lesão cerebral isquêmica, ou seja, falta de sangue no cérebro. No experimento, descrito aqui de forma bem sucinta, foram utilizados 51 ratos, sendo que 25 deles foram submetidos à isquemia cerebral global e transitória (lesionados) e os outros 26 animais foram designados como grupo falso isquêmico (não-lesionado). Foi utilizado no experimento um labirinto radial de oito braços aversivo, o mesmo é considerado como um modelo de aprendizagem que pretende imitar situações em que o animal possa encontrar no ambiente natural. Na Figura 3.1 temos uma representação esquemática do labirinto radial de oito braços aversivo. Esse tipo de experimento parte do pressuposto que alguns comportamentos aprendidos pelo animal são úteis em sua sobrevivência no meio selvagem, como por exemplo, a procura por água e comida. No labirinto utilizado no experimento, os braços se originam num ponto central e a comunicação dos braços com a arena central tem trânsito livre. Nas extremidades dos braços, uma abertura permite o acesso do animal a uma pequena caixa escura localizada logo abaixo de cada orifício, a qual pode ser inserida e removida como uma gaveta abaixo, servindo como refúgio para o rato em relação às áreas iluminadas do labirinto. Dentre os oito braços, somente um contem o refúgio verdadeiro, sendo que nos demais braços os esconderijos são de fundo falso. As funções cognitivas de todos os ratos foram testadas através do teste do labirinto, no qual era avaliada a capacidade do rato em encontrar o esconderijo. Cerca de vinte dias após a indução da isquemia cerebral, os ratos foram colocados diariamente no labirinto. O experimento durou 15 dias, e a cada dia de teste foram dadas três tentativas ao animal para encontrar o esconderijo. A variável resposta corresponde ao número de erros cometidos pelos ratos e as covariáveis foram

$$x_0 = \begin{cases} 1, & \text{se o } i\text{-rato for lesionado} \\ 0, & \text{caso contrário,} \end{cases}$$

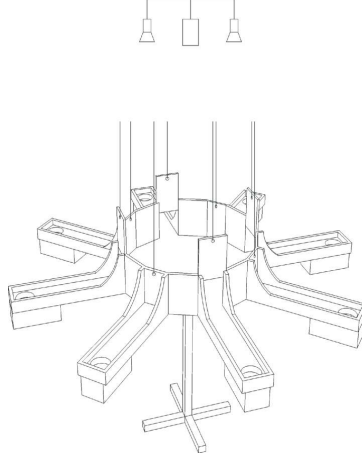


Figura 3.1: Representação esquemática do labirinto radial de oito braços aversivo.

$$x_1 = \begin{cases} 1, & \text{se o } i\text{-rato for não-lesionado} \\ 0, & \text{caso contrário} \end{cases}$$

e x_2 que corresponde ao tempo representado em cinco blocos de três dias cada, conforme a Tabela B.2. O modelo utilizado por Previdelli (2005) foi o modelo de regressão não-linear generalizado superdispersado Poisson, tendo como função de ligação a função identidade e no qual

$$\eta_i = \beta_0 x_{0i} + x_{1i} \exp(\beta_1 x_{2i}) + x_{0i} \exp(\beta_2 x_{2i}), \quad i = 1, \dots, 255. \quad (3.8)$$

Em nosso estudo, utilizamos o preditor não-linear acima, com $\beta_0 = 0,25$ e $\beta_1 = \beta_2 = 0,1$, para gerar a variável resposta proveniente das distribuições BNG, GPO, Consul e Poisson. Novamente, consideramos 10000 réplicas de Monte Carlo e $B = 600$ réplicas bootstrap. Os resultados, apresentados na Tabela 3.13, mostraram que em todos os modelos, de um modo geral, há um ganho considerável com o uso da correção de Cox & Snell no estimador de máxima verossimilhança. Nos modelos BNG e PO, o estimador $\tilde{\beta}$ foi o que teve o melhor desempenho, uma vez que apresentou as menores estimativas de viés em valores absolutos. No modelo BNG, particularmente, o estimador $\tilde{\beta}$ forneceu as maiores estimativas de viés, tendo assim o pior desempenho entre os estimadores. Nos modelos GPO e Consul, ambos os estimadores corrigidos tiveram um bom desempenho, fornecendo estimativas de viés inferiores às do estimador $\hat{\beta}$.

Tabela 3.7: Resultados da estimação pontual do vetor β no modelo não-linear Binomial Negativa Generalizada indexado pelos parâmetros $\beta_0 = \beta_1 = \beta_2 = 0, 25$, $\phi = 1, 5$ e $\nu = 5$ para $n = 25, 35, 45$ e 100 .

n	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
25	$\hat{\beta}$	-0,01346	-5,3830	0,00018	0,01991	7,9626	0,00040	-0,36879	-147,5200	0,13601
	$\tilde{\beta}$	0,01081	4,3219	0,00012	0,01274	5,0974	0,00016	0,29076	116,3000	0,08454
	$\check{\beta}$	0,02302	9,2065	0,00053	0,00259	1,0360	0,00001	-0,12964	-51,8560	0,01681
35	$\hat{\beta}$	-0,01428	-5,7134	0,00020	0,00891	3,5657	0,00008	-0,18050	-72,2020	0,03258
	$\tilde{\beta}$	0,00413	1,6527	0,00002	0,00483	1,9303	0,00002	0,05699	22,7950	0,00325
	$\check{\beta}$	0,01267	5,0671	0,00016	-0,00714	-2,8548	0,00005	0,01945	7,7794	0,00038
45	$\hat{\beta}$	-0,01095	-4,3806	0,00012	0,00733	2,9328	0,00005	-0,12069	-48,2760	0,01457
	$\tilde{\beta}$	0,00157	0,6277	0,00000	0,00377	1,5079	0,00001	0,01979	7,9149	0,00039
	$\check{\beta}$	0,00418	1,6701	0,00002	-0,00211	-0,8436	0,00000	0,02625	10,5000	0,00069
100	$\hat{\beta}$	-0,00638	-2,5503	0,00004	0,00329	1,3179	0,00001	-0,03543	-14,1730	0,00126
	$\tilde{\beta}$	-0,00027	-0,1093	0,00000	0,00088	0,3531	0,00000	0,00258	1,0336	0,00001
	$\check{\beta}$	0,00004	0,0179	0,00000	-0,00032	-0,1279	0,00000	0,02822	11,2860	0,00080

Tabela 3.8: Resultados da estimação pontual do vetor β no modelo não-linear Consul indexado pelos parâmetros $\beta_0 = \beta_1 = 0, 25, \beta_2 = 1$ e $\phi = 1$ para $n = 25, 35, 45$ e 100 .

n	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
25	$\hat{\beta}$	-0,04912	-19,6480	0,00241	0,04250	17,0000	0,00181	-0,09605	-9,6052	0,00923
	$\tilde{\beta}$	0,00162	0,6481	0,00000	0,00678	2,7120	0,00005	0,03357	3,3573	0,00113
	$\check{\beta}$	0,00240	0,9583	0,00001	-0,00810	-3,2395	0,00007	0,03982	3,9815	0,00159
35	$\hat{\beta}$	-0,03749	-14,9980	0,00141	0,01488	5,9517	0,00022	-0,03500	-3,5004	0,00123
	$\tilde{\beta}$	0,00026	0,1024	0,00000	0,00496	1,9853	0,00002	0,00401	0,4006	0,00002
	$\check{\beta}$	0,00708	2,8332	0,00005	-0,01029	-4,1159	0,00011	0,01781	1,7808	0,00032
45	$\hat{\beta}$	-0,03018	-12,0710	0,00091	0,01474	5,8950	0,00022	-0,02971	-2,9714	0,00088
	$\tilde{\beta}$	-0,00099	-0,3971	0,00000	0,00521	2,0840	0,00003	0,00228	0,2285	0,00001
	$\check{\beta}$	0,00254	1,0157	0,00001	-0,00446	-1,7848	0,00002	0,00980	0,9801	0,00010
100	$\hat{\beta}$	-0,01278	-5,1129	0,00016	0,00388	1,5520	0,00002	-0,01204	-1,2035	0,00014
	$\tilde{\beta}$	-0,00011	-0,0456	0,00000	0,00088	0,3512	0,00000	0,00082	0,0822	0,00000
	$\check{\beta}$	0,00187	0,7470	0,00000	-0,00058	-0,2337	0,00000	0,00088	0,0876	0,00000

Tabela 3.9: Resultados da estimação pontual do vetor β no modelo não-linear Poisson Generalizada indexado pelos parâmetros $\beta_0 = \beta_1 = \beta_2 = 0, 25$ e $\phi = 0, 2$ para $n = 25, 35, 45$ e 100 .

n	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
25	$\hat{\beta}$	-0,01477	-5,9072	0,00022	0,01921	7,6838	0,00037	-0,38494	-153,9800	0,14818
	$\tilde{\beta}$	0,01101	4,4019	0,00012	0,01354	5,4166	0,00018	0,34852	139,4100	0,12146
	$\check{\beta}$	0,02582	10,3280	0,00067	0,00305	1,2188	0,00001	-0,15162	-60,6460	0,02299
35	$\hat{\beta}$	-0,01491	-5,9645	0,00022	0,00894	3,5778	0,00008	-0,18724	-74,8940	0,03506
	$\tilde{\beta}$	0,00439	1,7553	0,00002	0,00513	2,0533	0,00003	0,06964	27,8570	0,00485
	$\check{\beta}$	0,01488	5,9534	0,00022	-0,00756	-3,0221	0,00006	0,00205	0,8208	0,00000
45	$\hat{\beta}$	-0,01020	-4,0782	0,00010	0,00731	2,9228	0,00005	-0,25702	-102,8100	0,06606
	$\tilde{\beta}$	0,00291	1,1633	0,00001	0,00389	1,5575	0,00002	-0,08515	-34,0600	0,00725
	$\check{\beta}$	0,00531	2,1250	0,00003	-0,00220	-0,8800	0,00000	0,02368	9,4729	0,00056
100	$\hat{\beta}$	-0,00664	-2,6571	0,00004	0,00337	1,3489	0,00001	-0,03711	-14,8450	0,00138
	$\tilde{\beta}$	-0,00032	-0,1270	0,00000	0,00093	0,3708	0,00000	0,00377	1,5099	0,00001
	$\check{\beta}$	0,00007	0,0287	0,00000	-0,00001	-0,0034	0,00000	0,02493	9,9719	0,00062

Tabela 3.10: Resultados da estimação pontual do vetor β no modelo não-linear Binomial Negativa Generalizada indexado pelos parâmetros $\beta_0 = \beta_1 = 0,25$ e $\beta_2 = 0,25, 0,5, 0,75, 1, \phi = 1, 5$ e $\nu = 5$ para $n = 35$.

	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
β_2	$\hat{\beta}$	-0,01428	-5,7134	0,00020	0,00891	3,5657	0,00008	-0,18050	-72,2020	0,03258
	$\tilde{\beta}$	0,00413	1,6527	0,00002	0,00483	1,9303	0,00002	0,05699	22,7950	0,00325
	$\check{\beta}$	0,01267	5,0671	0,00016	-0,00714	-2,8548	0,00005	0,01945	7,7794	0,00038
0,5	$\hat{\beta}$	-0,02102	-8,4071	0,00044	0,01004	4,0162	0,00010	-0,07023	-14,0450	0,00493
	$\tilde{\beta}$	-0,00033	-0,1310	0,00000	0,00506	2,0246	0,00003	0,01622	3,2440	0,00026
	$\check{\beta}$	0,00474	1,8960	0,00002	-0,00772	-3,0890	0,00006	0,05825	11,6500	0,00339
0,75	$\hat{\beta}$	-0,02324	-9,2954	0,00054	0,00970	3,8808	0,00009	-0,03610	-4,8128	0,00130
	$\tilde{\beta}$	-0,00097	-0,3862	0,00000	0,00459	1,8369	0,00002	0,00700	0,9340	0,00005
	$\check{\beta}$	0,00481	1,9242	0,00002	-0,00807	-3,2281	0,00007	0,02531	3,3747	0,00064
1	$\hat{\beta}$	-0,02421	-9,6828	0,00059	0,00945	3,7819	0,00009	-0,02691	-2,6913	0,00072
	$\tilde{\beta}$	-0,00156	-0,6236	0,00000	0,00440	1,7601	0,00002	0,00495	0,4948	0,00002
	$\check{\beta}$	0,00522	2,0878	0,00003	-0,00772	-3,0881	0,00006	0,00556	0,5556	0,00003

Tabela 3.11: Resultados da estimação pontual do vetor β no modelo não-linear Consul indexado pelos parâmetros $\beta_0 = \beta_1 = 0, 25$ e $\beta_2 = 0, 25, 0, 5, 0, 75, 1, \phi = 1$ para $n = 35$.

	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
β_2	$\hat{\beta}$	-0,02424	-9,6954	0,00059	0,01209	4,8341	0,00015	-0,49742	-198,9700	0,24743
	$\tilde{\beta}$	0,00847	3,3889	0,00007	0,00651	2,6025	0,00004	1,58820	635,2700	2,52230
	$\check{\beta}$	0,03081	12,3250	0,00095	-0,00953	-3,8137	0,00009	0,69055	276,2200	0,47686
0,5	$\hat{\beta}$	-0,03138	-12,5510	0,00098	0,01367	5,4674	0,00019	-0,28810	-57,6190	0,08300
	$\tilde{\beta}$	0,00251	1,0046	0,00001	0,00526	2,1042	0,00003	-0,02344	-4,6882	0,00055
	$\check{\beta}$	0,01308	5,2311	0,00017	-0,00957	-3,8289	0,00009	0,49339	98,6780	0,24343
0,75	$\hat{\beta}$	-0,03604	-14,4140	0,00130	0,01455	5,8185	0,00021	-0,06748	-8,9966	0,00455
	$\tilde{\beta}$	0,00030	0,1192	0,00000	0,00494	1,9777	0,00002	0,02725	3,6327	0,00074
	$\check{\beta}$	0,00719	2,8743	0,00005	-0,00961	-3,8437	0,00009	0,10189	13,5860	0,01038
1	$\hat{\beta}$	-0,03749	-14,9980	0,00141	0,01488	5,9517	0,00022	-0,03500	-3,5004	0,00123
	$\tilde{\beta}$	0,00026	0,1024	0,00000	0,00496	1,9853	0,00002	0,00401	0,4006	0,00002
	$\check{\beta}$	0,00708	2,8332	0,00005	-0,01029	-4,1159	0,00011	0,01781	1,7808	0,00032

Tabela 3.12: Resultados da estimação pontual do vetor β no modelo não-linear Poisson Generalizada indexado pelos parâmetros $\beta_0 = \beta_1 = 0, 25$ e $\beta_2 = 0, 25, 0, 5, 0, 75, 1, \phi = 0, 2$ para $n = 35$.

β_2	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
0,25	$\hat{\beta}$	-0,01491	-5,9645	0,00022	0,00894	3,5778	0,00008	-0,18724	-74,8940	0,03506
	$\tilde{\beta}$	0,00439	1,7553	0,00002	0,00513	2,0533	0,00003	0,06964	27,8570	0,00485
	$\check{\beta}$	0,01488	5,9534	0,00022	-0,00756	-3,0221	0,00006	0,00205	0,8208	0,00000
0,5	$\hat{\beta}$	-0,02128	-8,5119	0,00045	0,00967	3,8681	0,00009	-0,08210	-16,4190	0,00674
	$\tilde{\beta}$	0,00029	0,1153	0,00000	0,00487	1,9487	0,00002	0,02325	4,6502	0,00054
	$\check{\beta}$	0,00581	2,3258	0,00003	-0,00745	-2,9800	0,00006	0,05261	10,5220	0,00277
0,75	$\hat{\beta}$	-0,02421	-9,6819	0,00059	0,00916	3,6634	0,00008	-0,04057	-5,4091	0,00165
	$\tilde{\beta}$	-0,00082	-0,3265	0,00000	0,00424	1,6979	0,00002	0,00863	1,1511	0,00007
	$\check{\beta}$	0,00479	1,9146	0,00002	-0,00799	-3,1969	0,00006	0,03449	4,5987	0,00119
1	$\hat{\beta}$	-0,02564	-10,2550	0,00066	0,00877	3,5066	0,00008	-0,03433	-3,4326	0,00118
	$\tilde{\beta}$	-0,00176	-0,7053	0,00000	0,00396	1,5857	0,00002	0,01903	1,9033	0,00036
	$\check{\beta}$	0,00555	2,2211	0,00003	-0,00818	-3,2702	0,00007	0,02874	2,8744	0,00083

Tabela 3.13: Resultados da estimação pontual do vetor β no modelo (3.8) para diferentes distribuições.

Distribuições	Estimadores	β_0			β_1			β_2		
		Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM	Estimativa do Viés	Viés Relativo	EQM
Binomial Negativa Generalizada	$\hat{\beta}$	0,00182	0,7271	0,00000	-0,00272	-2,7221	0,00001	-0,00780	-7,8031	0,00006
	$\tilde{\beta}$	0,00057	0,2288	0,00000	-0,00005	-0,0509	0,00000	0,00205	2,0530	0,00000
	$\check{\beta}$	-0,11864	-47,4580	0,01408	-0,04535	-45,3450	0,00206	1,66770	1667,7000	2,78120
Consul	$\hat{\beta}$	0,00413	1,6539	0,00002	-0,00523	-5,2267	0,00003	-0,02519	-25,1880	0,00063
	$\tilde{\beta}$	-0,00316	-1,2650	0,00001	0,00005	0,0510	0,00000	0,01193	11,9290	0,00014
	$\check{\beta}$	0,00377	1,5098	0,00001	0,00060	0,6001	0,00000	0,00342	3,4228	0,00001
Poisson Generalizada	$\hat{\beta}$	0,00333	1,3310	0,00001	-0,00402	-4,0242	0,00002	-0,01425	-14,2470	0,00020
	$\tilde{\beta}$	-0,00125	-0,4983	0,00000	0,00001	0,0061	0,00000	0,00796	7,9647	0,00006
	$\check{\beta}$	0,00063	0,1258	0,00000	0,00018	0,1817	0,00000	0,00399	3,9934	0,00002
Poisson	$\hat{\beta}$	-0,00061	-0,2427	0,00000	-0,00033	-0,3261	0,00000	-0,00075	-0,7471	0,00000
	$\tilde{\beta}$	-0,00011	-0,0440	0,00000	0,00004	0,0413	0,00000	-0,00003	-0,0294	0,00000
	$\check{\beta}$	-0,00102	-0,4093	0,00000	0,00005	0,0509	0,00000	0,00013	0,1290	0,00000

3.5 Aplicação

Nesta seção, apresentaremos uma ilustração numérica da correção de viés via Cox & Snell em um conjunto de dados reais. Os dados, apresentados na Tabela B.1, correspondem ao número de espécies de peixe em um lago (variável resposta) e o logaritmo da área do lago, em km^2 , (x). Esses dados foram analisados inicialmente por Barbour e Brown (1974) e, posteriormente, por Rigby et al. (2008) e por Cordeiro *et al.* (2009). Estes últimos discutem a flexibilidade dos MSPNLGs em ajustar esses dados, adotando como preditores lineares

$$\eta_i = \beta_0 + \beta_1 \log(x_i) \quad (3.9)$$

e

$$\eta_i = \beta_0 + \beta_1 \log(x_i) + \beta_2 \{\log(x_i)\}^2, \quad (3.10)$$

$i = 1, \dots, 70$, com $\eta_i = \log(\mu_i - m)$, em que m denota o valor mínimo do suporte da distribuição associada ao modelo. As Tabelas 3.14 e 3.15 apresentam as estimativas do vetor de parâmetros β dos modelos analisados por Cordeiro *et al.* (2009). Observamos que as estimativas de $\hat{\beta}$ e $\tilde{\beta}$ não diferem muito para o preditor (3.9). Para a distribuição Delta Binomial, por exemplo, as estimativas de $\hat{\beta}$ e $\tilde{\beta}$ para β_0 foram em torno de 2,12 e para β_1 foram em torno de 0,18. Já para o preditor (3.10), em alguns modelos, $\hat{\beta}$ e $\tilde{\beta}$ apresentaram estimativas razoavelmente diferentes, como por exemplo o modelo GPO, no qual as estimativas de $\hat{\beta}$ foram 2,84570, -0,03851 e 0,01688 e as de $\tilde{\beta}$ foram 3,02570, -0,09795 e 0,02243. Pelo Critério de Informação de Akaike (AIC), dentre os modelos analisados, o modelo Delta Binomial com o preditor linear (3.10) foi o modelo mais adequado para o ajuste do número de espécies de peixe, uma vez que forneceu o menor AIC, a saber 612,1. Este resultado coincide com o resultado obtido por Cordeiro *et al.* (2009). A Figura 3.2 apresenta os valores ajustados obtidos ao estimar este modelo a partir das estimativas de $\hat{\beta}$ e de $\tilde{\beta}$. Observamos que não há diferença entre os valores ajustados obtidos a partir de ambas estimativas quando área do lago é pequena. No entanto, à medida que aumenta a área do lago, os valores ajustados obtidos a partir de $\tilde{\beta}$ se aproxima mais do número de espécie de peixes existente no lago do que os valores ajustados obtidos a partir de $\hat{\beta}$.

Tabela 3.14: Estimação dos parâmetros β nos modelos com preditor linear dado em (3.9).

Distribuições	Parâmetros	$\hat{\beta}$		$\tilde{\beta}$	
		Estimativa	Erro-Padrão	Estimativa	Erro-Padrão
Binomial	β_0	2,39010	0,27833	2,40890	0,27820
Negativa	β_1	0,17292	0,03693	0,17232	0,03692
GPO	β_0	2,53280	0,66835	2,64440	0,70120
$(\phi = 5)$	β_1	0,14890	0,10717	0,15171	0,11326
BNG	β_0	2,38070	0,18626	2,38900	0,18617
$(\phi = 1, \nu = 2, 43)$	β_1	0,17404	0,02459	0,17376	0,02458
Delta Binomial	β_0	2,11480	0,26548	2,13170	0,26643
$(\phi = 5, m = 5)$	β_1	0,18194	0,04142	0,18217	0,04164
Geeta- m	β_0	2,15640	0,61525	2,25360	0,65666
$(\phi = 1, 1, m = 5)$	β_1	0,17544	0,10974	0,18257	0,11834

Tabela 3.15: Estimação dos parâmetros β nos modelos com preditor linear dado em (3.10).

Distribuições	Parâmetros	$\hat{\beta}$		$\tilde{\beta}$	
		Estimativa	Erro-Padrão	Estimativa	Erro-Padrão
Poisson	β_0	2,68330	0,09214	2,68660	0,09204
	β_1	0,03034	0,02513	0,02960	0,02510
	β_2	0,01164	0,00164	0,01168	0,00164
Binomial	β_0	2,83420	0,41083	2,90700	0,41023
Negativa	β_1	-0,03361	0,13757	-0,05374	0,13743
	β_2	0,01651	0,01047	0,01802	0,01046
GPO	β_0	2,84570	0,62688	3,02570	0,66123
$(\phi = 0, 3)$	β_1	-0,03851	0,24313	-0,09795	0,25657
	β_2	0,01688	0,02109	0,02243	0,02242
BNG	β_0	2,83610	0,27435	2,86850	0,27398
$(\phi = 1, \nu = 2, 43)$	β_1	-0,03477	0,09160	-0,04371	0,09152
	β_2	0,01662	0,00696	0,01729	0,00696
Delta Binomial	β_0	2,47760	0,36148	2,53710	0,36482
$(\phi = 3, m = 5)$	β_1	-0,02307	0,13866	-0,04238	0,14025
	β_2	0,01818	0,01195	0,01994	0,01212

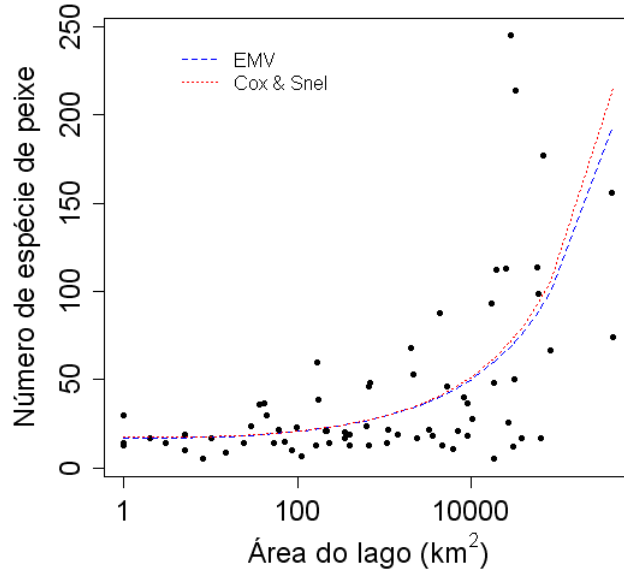


Figura 3.2: Número de espécie de peixe versus área do lago, juntamente com valores ajustados do modelo Delta Binomial($\phi = 3, m = 5$), cujo preditor linear é dado em (3.10).

3.6 Comentários

Neste capítulo, obtivemos uma expressão em forma matricial do viés de segunda ordem via Cox & Snell (1968) para os estimadores de máxima verossimilhança dos parâmetros dos modelos em séries de potência lineares e não-lineares generalizados, considerando fixo o parâmetro de dispersão. A partir desta expressão, definimos um estimador de máxima verossimilhança corrigido, o qual apresenta um viés de ordem n^{-2} inferior ao viés apresentado pelo estimador de máxima verossimilhança, que por sua vez é de ordem n^{-1} . Adicionalmente, discutiremos sobre a correção de viés através do método bootstrap. Resultados de simulação foram obtidos tanto para os modelos lineares quanto para os modelos não-lineares, envolvendo o estimador de máxima verossimilhança ($\hat{\beta}$), o estimador de máxima verossimilhança corrigido via correção de Cox & Snell ($\tilde{\beta}$) e corrigido via bootstrap ($\check{\beta}$).

Os resultados mostraram que o estimador $\hat{\beta}$, entre os estimadores em estudo, teve o pior desempenho por apresentar as maiores estimativas de viés em valor absoluto. Já o estimador

corrigido $\tilde{\beta}$ foi o mais eficaz, uma vez que apresentou estimativas de viés, em módulo, sempre menor do que o estimador $\hat{\beta}$ e, na maioria das vezes, menor do que o estimador corrigido $\check{\beta}$. Este, por sua vez, em algumas situações, apresentou estimativas de viés, em módulo, maiores do que o estimador $\hat{\beta}$.

À medida que o tamanho de amostra cresce, como era de se esperar, todos os estimadores apresentaram uma redução nos valores absolutos das estimativas do viés, mesmo assim as correções mostraram-se necessárias, uma vez que as diferenças entre as estimativas do estimador de máxima verossimilhança usual e as estimativas dos novos estimadores foram bem distintas, mesmo quando consideramos um número grande de observações. Quando aumentamos o valor de um dos parâmetros, no caso do nosso estudo aumentamos o valor de β_2 , para este mesmo parâmetro houve uma redução, em módulo, nas estimativas do viés relativo produzidas por todos estimadores. Vimos também que a correção de Cox & Snell feita no estimador de máxima verossimilhança produziu estimativas, na maioria das vezes, superiores aos valores verdadeiros dos parâmetros e a correção por bootstrap mudou, quase sempre, os sinais das estimativas do viés do estimador de máxima verossimilhança usual.

Em suma, de acordo com os resultados obtidos nas Seções 3.4.1 e 3.4.2, recomendamos o uso da correção de viés dos estimadores de máxima verossimilhança dos parâmetros dos modelos em séries de potência lineares e não-lineares generalizados. Para isso, sugerimos o uso da correção de Cox & Snell, no caso dos modelos lineares, e ambas correções de Cox & Snell e via bootstrap, no caso dos modelos não-lineares.

Capítulo 4

Correção de Bartlett em MSPNLG

4.1 Introdução

Estatísticas de testes, em que as suas distribuições são baseadas em aproximações para grandes amostras, são bastantes utilizadas quando há uma grande dificuldade em se determinar as suas distribuições exatas, como é o caso das estatísticas da razão de verossimilhanças (LR), Wald e escore. Estas estatísticas possuem a mesma distribuição de referência χ^2 e, portanto, são assintoticamente equivalentes. Testes baseados nessas estatísticas são denominados assintóticos de primeira ordem, isto é, são baseados em valores críticos obtidos de uma distribuição nula limite conhecida. No entanto, em pequenas amostras ou mesmo em amostras de tamanho moderado, a aproximação da distribuição dessas estatísticas pela distribuição χ^2 pode não ser satisfatória, podendo conduzir a taxas de rejeição sob a hipótese nula bastante distorcidas, tornando, portanto, uma preocupação recorrente verificar a qualidade dessa aproximação.

Com esse intuito, Bartlett (1937) propôs um fator de correção para o teste da razão de verossimilhanças originando, assim, uma estatística da razão de verossimilhanças modificada LR^* , cuja média está mais próxima do valor esperado da distribuição χ^2 de referência. Isso porque, sob a hipótese nula, o valor esperado $E(LR)$ corresponde a $q\{1 + b + O(n^{-2})\}$, em que q é o número de restrições impostas por H_0 , n é o tamanho da amostra e b uma constante de ordem n^{-1} , que pode ser estimada consistentemente sob H_0 , enquanto que o

valor esperado de LR^* corresponde a $q + O(n^{-2})$. Além disso, para testes de homogeneidade de variâncias, Bartlett mostrou que os três primeiros momentos de LR^* concordam com os momentos correspondentes da distribuição χ^2 até ordem n^{-1} . Consequentemente, temos que a distribuição de LR^* melhor se aproxima da distribuição χ^2 do que a distribuição de LR .

Lawley (1956) desenvolveu um método geral de obtenção para o fator de correção que envolve momentos das quatro primeiras derivadas do logaritmo da função de verossimilhança e mostrou que a estatística LR^* tem todos os momentos concordando com os respectivos da distribuição χ^2 de referência, ignorando os termos de ordem n^{-2} . Posteriormente, Hayakawa (1977) obteve uma expansão assintótica de ordem n^{-1} para da distribuição nula de LR e mostrou que, se a hipótese nula for simples, a estatística LR^* tem distribuição χ^2 até a ordem n^{-1} . Porém, para hipóteses compostas, só seria possível obter o fator de correção se um determinado coeficiente da expansão fosse nulo, o que parecia conflitar com os resultados de Lawley (1956). Chesher e Smith (1995) resolveram esse impasse quando notaram um erro na fórmula do coeficiente em questão e mostraram que este, depois de corrigido, é sempre igual a 0.

Dentre os diversos artigos produzidos na literatura que apresentam correções de Bartlett para a estatística da razão de verossimilhanças em modelos variados e em situações específicas, destacam-se os seguintes trabalhos: Cordeiro (1983, 1987) para os modelos lineares generalizados (MLGs) quando o fator de escala é conhecido e desconhecido, respectivamente; Cordeiro e Paula (1989) para os modelos não-lineares da família exponencial com parâmetro de dispersão conhecido; Cribari-Neto e Ferrari (1995) para os modelos lineares normais heteroscedásticos; Cribari-Neto e Zarkos (1995) para os modelos de regressão multivariada; Cordeiro *et al.* (1995) para a família exponencial uniparamétrica; Ferrari e Arellano-Valle (1996) para os modelos de regressão com erros t de Student; Ferrari e Uribe-Opazo (2001) para os modelos lineares simétricos; Montenegro e Cordeiro (2002) para os modelos não-lineares de locação e escala supondo que o parâmetro de escala é conhecido; Cordeiro (2004) para os modelos não-lineares simétricos, generalizando os resultados de Ferrari e Uribe-Opazo (2001). O fator de correção de Bartlett para a estatística da razão de verossimilhanças perfilada foi obtido por Ferrari *et al.* (2004) para o modelo de regressão normal linear heteroscedástico e por Cysneiros e Ferrari (2006) para os modelos de regressão não-lineares da família exponencial.

Frydenberg e Jensen (1989) mostraram, por meio de simulação, que os resultados teóricos que garantem que a distribuição da estatística corrigida tenha uma boa aproximação com a distribuição χ^2 são válidos apenas para os modelos contínuos. No entanto, Cordeiro (1982) fez vários estudos de simulação envolvendo distribuições multinomial e de Poisson que mostraram que os testes modificados, baseados em LR^* , apresentam taxas de rejeição da hipótese nula bem mais próximas dos respectivos níveis nominais do que o teste não modificado. As correções nos modelos discretos também se mostraram eficazes nos estudos de simulação realizados por Cysneiros (1997), que obteve correções de Bartlett e tipo-Bartlett nos modelos lineares generalizados contínuos e discretos.

4.2 Correção de Bartlett

Considere um modelo multiparamétrico com vetor de parâmetros desconhecidos $\theta = (\theta_1^\top, \theta_2^\top)^\top$, sendo θ_1 e θ_2 vetores de dimensões q e $p - q$, respectivamente, e $l = l(\theta)$ representando o logaritmo da função de verossimilhança. Para testar $H_0 : \theta_1 = \theta_1^{(0)}$ versus $H_1 : \theta_1 \neq \theta_1^{(0)}$, sendo $\theta_1^{(0)}$ um vetor de constantes conhecidas, a estatística da razão de verossimilhanças é definida como

$$LR = 2\{l(\hat{\theta}) - l(\tilde{\theta})\}, \quad (4.1)$$

em que $\hat{\theta} = (\hat{\theta}_1^\top, \hat{\theta}_2^\top)^\top$ e $\tilde{\theta} = (\theta_1^{(0)\top}, \tilde{\theta}_2^\top)^\top$ são os estimadores de máxima verossimilhança de $\theta = (\theta_1^\top, \theta_2^\top)^\top$, segundo H_1 e H_0 , respectivamente. Adotando r, s, t, u, v, w como indexadores do espaço paramétrico, as derivadas do logaritmo da função de verossimilhança podem ser denotadas da seguinte maneira: $U_r = \partial l / \partial \theta_r$, $U_{rs} = \partial^2 l / \partial \theta_r \partial \theta_s$, $U_{rst} = \partial^3 l / \partial \theta_r \partial \theta_s \partial \theta_t$ e assim por diante. Consequentemente, os cumulantes conjuntos dessas derivadas são definidos como $\kappa_{rs} = E(U_{rs})$, $\kappa_{r,s} = E(U_r U_s)$, $\kappa_{rst} = E(U_{rst})$, $\kappa_{rs,t} = E(U_{rs} U_t)$, etc. Denotamos as derivadas dos momentos em relação aos componentes do vetor θ por $\kappa_{rs}^{(t)} = \partial \kappa_{rs} / \partial \theta_t$ e $\kappa_{rs}^{(tu)} = \partial^2 \kappa_{rs} / \partial \theta_t \partial \theta_u$.

Sob condições gerais de regularidade, Lawley (1956) obteve uma expansão de $l(\hat{\theta})$ em série de Taylor sob a hipótese nula até termos de ordem n^{-1} envolvendo derivadas até de

quarta ordem do logaritmo da função de verossimilhança. Assim, ele mostrou que

$$2E\{l(\hat{\theta}_1, \hat{\theta}_2) - l(\theta_1, \theta_2)\} = p + \epsilon_p + O(n^{-2}), \quad (4.2)$$

sendo o termo ϵ_p de ordem n^{-1} expresso da seguinte forma:

$$\epsilon_p = \sum (\lambda_{rstu} - \lambda_{rstuvw}), \quad (4.3)$$

em que $\sum$ denota o somatório sobre todas as componentes do vetor θ ,

$$\lambda_{rstu} = \kappa^{rs} \kappa^{tu} \left(\kappa_{rstu}/4 - \kappa_{rst}^{(u)} + \kappa_{rt}^{(su)} \right) \quad \text{e} \quad (4.4)$$

$$\begin{aligned} \lambda_{rstuvw} = & \kappa^{rs} \kappa^{tu} \kappa^{vw} \left\{ \kappa_{rtv} \left(\kappa_{suw}/6 - \kappa_{sw}^{(u)} \right) + \kappa_{rtu} \left(\kappa_{svw}/4 - \kappa_{sw}^{(v)} \right) \right. \\ & \left. + \kappa_{rt}^{(v)} \kappa_{sw}^{(u)} + \kappa_{rt}^{(u)} \kappa_{sw}^{(v)} \right\}, \end{aligned} \quad (4.5)$$

com $-\kappa^{rs} = \kappa^{r,s}$ representando o elemento (r, s) da inversa da matriz de informação de Fisher K_θ de θ . Além disso, Lawley (1956) também demonstrou que

$$2E\{l(\theta_1^{(0)}, \tilde{\theta}_2) - l(\theta_1, \theta_2)\} = p - q + \epsilon_{p-q} + O(n^{-2}), \quad (4.6)$$

sendo o termo ϵ_{p-q} de ordem n^{-1} obtido analogamente ao termo ϵ_p dado em (4.3) com o somatório $\sum$ estendendo-se apenas sobre os componentes do vetor θ_2 , ou seja, sobre os $p - q$ parâmetros de perturbação. A estatística da razão de verossimilhanças definida em (4.1) pode ser reescrita como

$$LR = 2 \left[\{l(\hat{\theta}_1, \hat{\theta}_2) - l(\theta_1, \theta_2)\} - \{l(\theta_1^{(0)}, \tilde{\theta}_2) - l(\theta_1, \theta_2)\} \right].$$

A partir de (4.2) e (4.6) segue que, sob a hipótese nula, o valor esperado de LR é dado por

$$\begin{aligned} E(LR) &= 2E \left[\{l(\hat{\theta}_1, \hat{\theta}_2) - l(\theta_1, \theta_2)\} - \{l(\theta_1^{(0)}, \tilde{\theta}_2) - l(\theta_1, \theta_2)\} \right] \\ &= q + \epsilon_p - \epsilon_{p-q} + O(n^{-2}) \\ &= q \left(1 + \frac{\epsilon_p - \epsilon_{p-q}}{q} \right) + O(n^{-2}). \end{aligned}$$

Desse modo, a aproximação da distribuição da estatística da razão de verossimilhanças pela distribuição χ_q^2 pode ser melhorada substituindo LR pela estatística modificada LR^* dada por

$$LR^* = \frac{LR}{1+d},$$

ou, equivalentemente,

$$LR_1^* = LR(1-d),$$

em que os fatores de correção de Bartlett, $1/(1+d)$ e $(1-d)$, são determinados através de

$$d = \frac{\epsilon_p - \epsilon_{p-q}}{q}. \quad (4.7)$$

As estatísticas modificadas LR^* e LR_1^* possuem distribuição χ_q^2 até ordem n^{-1} sob H_0 e sob certas condições de regularidade, segundo Hayakawa (1977) (vide correção de Chesher e Smith, 1995). Um teste da razão de verossimilhanças aperfeiçoado compara as estatísticas LR^* e LR_1^* com a distribuição χ_q^2 de referência. Deve-se destacar que os fatores de correção não dependem do valor da estatística da razão de verossimilhanças, mas podem depender de parâmetros desconhecidos e, neste caso, estes devem ser substituídos por suas respectivas estimativas de máxima verossimilhança sob H_0 , o que não afeta a ordem de aproximação resultante. Vale a pena ressaltar que no caso do teste da hipótese nula simples $H_0 : \theta = \theta^{(0)}$, a quantidade d dada em (4.7) que determina os fatores de correção de Bartlett se reduz a $d = \epsilon_p/p$, em que ϵ_p é calculado pela expressão dada em (4.3).

4.2.1 Correção de Bartlett em MSPNLG

Os fatores de correção de Bartlett dependem da quantidade ϵ_p , dada em (4.3), que é uma função aparentemente complicada dos cumulantes conjuntos κ 's de derivadas do logaritmo da função de verossimilhança. O objetivo desta seção é apresentar ϵ_p em forma matricial e de fácil computação para a classe dos MSPNLGs.

Com essa finalidade, considere $Y_1, \dots, Y_n$ variáveis aleatórias discretas independentes, cada qual com função de probabilidade na forma

$$\pi(y; \mu_i, \phi) = \frac{a(y, \phi)g(\mu_i, \phi)^y}{f(\mu_i, \phi)}, \quad y \in A_\epsilon, \quad (4.8)$$

em que o suporte de Y_i é um subconjunto A_ϵ dos inteiros $\{\epsilon, \epsilon + 1, \dots\}$, $\epsilon \geq 0$, e que não depende de parâmetros desconhecidos, $a(y, \phi)$ é uma função positiva, as funções analíticas $f_i = f(\mu_i, \phi)$ e $g_i = g(\mu_i, \phi)$ são positivas, finitas e duas vezes diferenciáveis, $\phi > 0$ e $\mu_i > 0$ são chamados de parâmetros de dispersão e de média, respectivamente.

Para a família de distribuições dada em (4.8), as seguintes relações são válidas:

$$E(Y) = \mu = \frac{f^{(1)}g}{fg^{(1)}} \quad \text{e} \quad Var(Y) = V(\mu, \phi) = \frac{g}{g^{(1)}},$$

em que o índice sobrescrito (1) indica a primeira diferenciação em relação a μ . Os modelos em séries de potência não-lineares generalizados são definidos por (4.8) e pelo componente sistemático

$$h(\mu_i) = \eta_i = \eta(x_i; \beta), \quad i = 1, \dots, n$$

em que $h(\cdot)$ é uma função de ligação conhecida e duplamente diferenciável, $\beta = (\beta_1, \dots, \beta_p)^\top$ é um vetor de p ($p < n$) parâmetros desconhecidos a serem estimados, $x_i = (x_{i1}, \dots, x_{ik})^\top$ representa os valores de k variáveis explicativas e $\eta(\cdot; \cdot)$ é uma função possivelmente não-linear no segundo argumento, contínua e diferenciável com respeito aos componentes de β tal que a matriz de derivadas $\tilde{X} = \tilde{X}(\beta) = \partial\eta/\partial\beta^\top$, com $\eta = (\eta_1, \dots, \eta_n)^\top$, tem posto p para todo β . A matriz $\tilde{X}$ tem elementos que são, em geral, funções do vetor de parâmetros β desconhecidos.

O logaritmo da função de verossimilhança do vetor dos parâmetros β , dado o vetor de observações $(y_1, \dots, y_n)$, dos MSPNLGs pode ser expresso na forma

$$l(\beta) = \sum_{i=1}^n \log\{a(y_i, \phi)\} + \sum_{i=1}^n [y_i \log\{g(\mu_i, \phi)\} - \log\{f(\mu_i, \phi)\}].$$

A função escore para o parâmetro β , condicionando em ϕ , é dada por

$$U_\beta = \frac{\partial l(\beta)}{\partial \beta} = \tilde{X}^\top (Ty - Q),$$

em que $T = \text{diag}\{t_1, \dots, t_n\}$ é uma matriz diagonal de dimensão $n \times n$ cujo i -ésimo elemento é $t_i = \frac{g'_i}{g_i h'_i}$ e $Q = (q_1, \dots, q_n)^\top$ é um vetor $n \times 1$ cujo i -ésimo elemento é $q_i = \frac{f'_i}{f_i h'_i}$. A matriz de informação de Fisher de β dado ϕ é

$$K_\beta = E \left\{ -\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta^\top} \right\} = \tilde{X}^\top W \tilde{X}, \quad (4.9)$$

em que W é uma matriz diagonal $n \times n$ de pesos dados por

$$w_i = \left(q_i^{(1)} - \frac{f_i^{(1)} g_i}{f_i g_i^{(1)}} t_i^{(1)} \right) \frac{1}{h_i^{(1)}}, \quad i = 1, \dots, n.$$

Considere que o vetor de parâmetros β pode ser decomposto como $\beta = (\beta_1^\top, \beta_2^\top)^\top$, sendo $\beta_1 = (\beta_1, \dots, \beta_q)^\top$ o vetor de parâmetros de interesse e $\beta_2 = (\beta_{q+1}, \dots, \beta_p)^\top$ o vetor de parâmetros de perturbação. Essa decomposição induz a correspondente partição $\tilde{X} = (\tilde{X}_1, \tilde{X}_2)$, sendo $\tilde{X}$ a matriz de derivadas com $\tilde{X}_1 = \partial\eta/\partial\beta_1^\top$ e $\tilde{X}_2 = \partial\eta/\partial\beta_2^\top$. O nosso interesse é testar a hipótese nula $H_0 : \beta_1 = \beta_1^{(0)}$ versus a alternativa $H_1 : \beta_1 \neq \beta_1^{(0)}$, em que $\beta_1^{(0)}$ é um vetor especificado de dimensão q ($q \leq p$). A estatística da razão de verossimilhanças para o teste de H_0 , assumindo um valor fixo para ϕ , pode ser escrita da forma

$$LR = 2\{l(\hat{\beta}) - l(\tilde{\beta})\},$$

sendo $\hat{\beta}$ o estimador de máxima verossimilhança irrestrito de β e $\tilde{\beta} = (\beta_1^{(0)\top}, \tilde{\beta}_2^{(\phi)\top})^\top$ o estimador correspondente de β sob a hipótese nula.

Utilizamos a notação $\tilde{x}_{ir} = \partial\eta_i/\partial\beta_r$, $\tilde{x}_{irs} = \partial^2\eta_i/\partial\beta_r\partial\beta_s$, $\tilde{x}_{irst} = \partial^3\eta_i/\partial\beta_r\partial\beta_s\partial\beta_t$ e denotamos por $\sum_i$ o somatório sobre os dados. Definimos também os escalares w_{ji} , $\tilde{w}_{ji}$ e $\tilde{w}_{1i}^*$, respectivamente, por

$$\begin{aligned} w_{ji} &= \left(\frac{f_i^{(1)} g_i}{f_i g_i^{(1)}} t_i^{(j)} - q_i^{(j)} \right) \frac{1}{h_i^{(1)}}, \\ \tilde{w}_{ji} &= \varphi_{ji} - \frac{(j-1)q_i V_i t_i^{(j)} h_i^{(2)} - q_i^{(j+1)}}{(h_i^{(1)})^{j+1}} + j \frac{q_i^{(j)} h_i^{(2)}}{(h_i^{(1)})^{j+2}} \quad \text{e} \\ \tilde{w}_{1i}^* &= 2\varphi_{2i} - \frac{q_i V_i t_i^{(3)}}{(h_i^{(1)})^2} + \frac{t_i^{(1)}(q_i^{(2)} V_i + 2q_i^{(1)} V_i^{(1)} + q_i^{(1)} V_i^{(2)})}{(h_i^{(1)})^2} - \frac{h_i^{(2)} \varphi_{1i}}{(h_i^{(1)})^2} - \frac{q_i^{(3)}}{(h_i^{(1)})^3} + 3 \frac{q_i^{(2)} h_i^{(2)}}{(h_i^{(1)})^4} \\ &\quad + q_i^{(1)} \left(\frac{h_i^{(3)}}{(h_i^{(1)})^4} - 3 \frac{(h_i^{(2)})^2}{(h_i^{(1)})^5} \right), \end{aligned}$$

com

$$\varphi_{ji} = \frac{q_i^{(1)} V_i t_i^{(j)} + q_i V_i^{(1)} t_i^{(j)} + q_i V_i t_i^{(j+1)}}{(h_i^{(1)})^j},$$

para $j = 1, 2, 3$ e $i = 1, \dots, n$. Aqui o índice sobrescrito (j) indica a j -ésima derivada em relação a μ . Vale ressaltar que as quantidades acima envolvem derivadas que dependem das

formas específicas das funções f , g , h e V nas diversas distribuições pertencentes à família de série de potência. Temos, então, os seguintes cumulantes para os MSPNLGs (vide Apêndice A):

$$\begin{aligned}
\kappa_{rs} &= \sum_{i=1}^n w_{1i} \tilde{x}_{ir} \tilde{x}_{is}, \\
\kappa_{rst} &= \sum_{i=1}^n \left\{ \left[w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} + w_{1i} [\tilde{x}_{ir} \tilde{x}_{ist} + \tilde{x}_{irt} \tilde{x}_{is} + \tilde{x}_{irs} \tilde{x}_{it}] \right\} \quad \text{e} \\
\kappa_{rstu} &= \sum_{i=1}^n \left\{ \left[w_{3i} - 3 \frac{w_{2i} h_i^{(2)}}{(h_i^{(1)})^2} - \frac{w_{1i} h_i^{(3)}}{(h_i^{(1)})^3} + 3 \frac{w_{1i} (h_i^{(2)})^2}{(h_i^{(1)})^4} \right] \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} \right. \\
&\quad + \left[w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] [\tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu} + \tilde{x}_{ir} \tilde{x}_{isu} \tilde{x}_{it} + \tilde{x}_{iru} \tilde{x}_{is} \tilde{x}_{it} + (\tilde{x}_{ir} \tilde{x}_{ist} + \tilde{x}_{irt} \tilde{x}_{is} \\
&\quad + \tilde{x}_{irs} \tilde{x}_{it}) \tilde{x}_{iu}] + w_{1i} (\tilde{x}_{ir} \tilde{x}_{istu} + \tilde{x}_{iru} \tilde{x}_{ist} + \tilde{x}_{irt} \tilde{x}_{isu} + \tilde{x}_{irtu} \tilde{x}_{is} + \tilde{x}_{irs} \tilde{x}_{itu} + \tilde{x}_{irsu} \tilde{x}_{it} \\
&\quad \left. + \tilde{x}_{irst} \tilde{x}_{iu}) \right\}.
\end{aligned}$$

As derivadas dos cumulantes dados acima são

$$\begin{aligned}
\kappa_{rs}^{(t)} &= \sum_{i=1}^n \{ \tilde{w}_{1i} \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} + w_{1i} \tilde{x}_{irt} \tilde{x}_{is} + w_{1i} \tilde{x}_{ist} \tilde{x}_{ir} \}, \\
\kappa_{rs}^{(tu)} &= \sum_{i=1}^n \left\{ \tilde{w}_{1i}^* \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} + \tilde{w}_{1i} \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu} + \tilde{w}_{1i} \tilde{x}_{iru} \tilde{x}_{is} \tilde{x}_{it} + \tilde{w}_{1i} \tilde{x}_{ir} \tilde{x}_{isu} \tilde{x}_{it} + \tilde{w}_{1i} \tilde{x}_{irt} \tilde{x}_{is} \tilde{x}_{iu} \right. \\
&\quad \left. + w_{1i} \tilde{x}_{irtu} \tilde{x}_{is} + w_{1i} \tilde{x}_{irt} \tilde{x}_{isu} + \tilde{w}_{1i} \tilde{x}_{ir} \tilde{x}_{ist} \tilde{x}_{iu} + w_{1i} \tilde{x}_{iru} \tilde{x}_{ist} + w_{1i} \tilde{x}_{ir} \tilde{x}_{istu} \right\} \quad \text{e} \\
\kappa_{rst}^{(u)} &= \sum_{i=1}^n \left\{ \tilde{w}_{2i} - \frac{\tilde{w}_{1i} h_i^{(2)}}{(h_i^{(1)})^2} - \frac{w_{1i} h_i^{(3)}}{(h_i^{(1)})^3} + 2 \frac{w_{1i} (h_i^{(2)})^2}{(h_i^{(1)})^4} \right\} \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} \\
&\quad + \sum_{i=1}^n \left\{ w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right\} \{ \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu} + \tilde{x}_{ir} \tilde{x}_{isu} \tilde{x}_{it} + \tilde{x}_{iru} \tilde{x}_{is} \tilde{x}_{it} \} \\
&\quad + \sum_{i=1}^n \tilde{w}_{1i} \tilde{x}_{iu} \{ \tilde{x}_{ir} \tilde{x}_{ist} + \tilde{x}_{irt} \tilde{x}_{is} + \tilde{x}_{irs} \tilde{x}_{it} \} \\
&\quad + \sum_{i=1}^n w_{1i} \{ \tilde{x}_{ir} \tilde{x}_{istu} + \tilde{x}_{iru} \tilde{x}_{ist} + \tilde{x}_{irt} \tilde{x}_{isu} + \tilde{x}_{irtu} \tilde{x}_{is} + \tilde{x}_{irs} \tilde{x}_{itu} + \tilde{x}_{irsu} \tilde{x}_{it} \}.
\end{aligned}$$

Sejam $\tilde{x}_i^\top$ a i -ésima linha da matriz $\tilde{X}$, $\tilde{\tilde{X}}_i$ uma matriz $p \times p$ cujo elemento (r, s) é $\tilde{x}_{irs}$, $i = 1, \dots, n$, e K_β^{-1} a inversa da matriz de informação de Fisher K_β dada pela equação (4.9). Definimos

$$Z = \tilde{X}(\tilde{X}^\top W \tilde{X})^{-1} \tilde{X}^\top,$$

uma matriz de dimensão $n \times n$ positiva semi-definida de posto p com elementos $z_{ij} = \tilde{x}_i^\top K_\beta^{-1} \tilde{x}_j$, as matrizes quadradas B e C de dimensão $n \times n$, com elementos dados por $b_{ij} = \text{tr}(K_\beta^{-1} \tilde{\tilde{X}}_i K_\beta^{-1} \tilde{\tilde{X}}_j)$ e $c_{ij} = \tilde{x}_i^\top K_\beta^{-1} \tilde{\tilde{X}}_j K_\beta^{-1} \tilde{x}_i$, respectivamente, e a matriz diagonal $D = \text{diag}\{d_{11}, \dots, d_{1n}\}$ com $d_{1i} = \text{tr}(K_\beta^{-1} \tilde{\tilde{X}}_i)$. Utilizamos a notação Z_d , B_d e C_d para representar matrizes diagonais formadas pelos correspondentes elementos das diagonais das matrizes Z , B e C , respectivamente. Denotamos $Z^{(3)} = Z^{(2)} \odot Z$, $Z^{(2)} = Z \odot Z$, em que $\odot$ denota o produto de Hadamard (Rao, 1973, p. 30), ou seja, o elemento (i, j) de $Z^{(3)}$ é z_{ij}^3 . Adicionalmente, definimos as matrizes diagonais Q_1 , Q_2 , Q_3 e Q_4 de dimensão $n \times n$, cujos elementos são dados, respectivamente, por

$$q_{1i} = w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2}, \quad (4.10)$$

$$q_{2i} = \frac{1}{6} \left(w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right) - \tilde{w}_{1i}, \quad (4.11)$$

$$q_{3i} = \frac{1}{4} \left(w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right) - \tilde{w}_{1i} \quad \text{e} \quad (4.12)$$

$$q_{4i} = \frac{1}{4} w_{3i} - \frac{3}{4} \frac{w_{2i} h_i^{(2)}}{(h_i^{(1)})^2} + \frac{3}{4} \frac{w_{1i} h_i^{(3)}}{(h_i^{(1)})^3} - \frac{5}{4} \frac{w_{1i} (h_i^{(2)})^2}{(h_i^{(1)})^4} + \frac{\tilde{w}_{1i} h_i^{(2)}}{(h_i^{(1)})^2} - \tilde{w}_{2i} + \tilde{w}_{1i}^*, \quad (4.13)$$

$i = 1, \dots, n$. Uma expressão simples para o termo ϵ_p em notação matricial pode ser obtida substituindo os κ 's na expressão (4.3) e efetuando as somas sobre os parâmetros seguidas das somas sobre as amostras. Ao procedermos assim, aparecerão termos da forma $-\sum \tilde{x}_{ir} \kappa^{rs} \tilde{x}_{js}$, $\sum \kappa^{rs} \tilde{x}_{jsu} \kappa^{ut} \tilde{x}_{itr}$, $\sum \tilde{x}_{it} \kappa^{tu} \tilde{x}_{jus} \kappa^{sr} \tilde{x}_{ir}$ e $-\sum \kappa^{rs} \tilde{x}_{irs}$, em que $-\kappa^{rs} = \kappa^{r,s}$ representa o elemento (r, s) da matriz K_β^{-1} , $r, s = 1, \dots, p$. Esses termos correspondem aos elementos das matrizes Z , B , C e D , respectivamente.

Detalhamos agora a obtenção da parcela $\sum \lambda_{rstu}$ de ϵ_p . Substituindo κ_{rstu} , $\kappa_{rst}^{(u)}$ e $\kappa_{rt}^{(su)}$

na expressão de $\sum \lambda_{rstu}$ dada em (4.4) temos que

$$\begin{aligned} \sum \lambda_{rstu} = & \sum \kappa^{rs} \kappa^{tu} \sum_{i=1}^n \left\{ q_{4i} \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} + \left(\tilde{w}_{1i} - \frac{3}{4} q_{1i} \right) (\tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu} + \tilde{x}_{ir} \tilde{x}_{isu} \tilde{x}_{it} \right. \\ & + \tilde{x}_{iru} \tilde{x}_{is} \tilde{x}_{it}) + \frac{1}{4} q_{1i} \tilde{x}_{iu} (\tilde{x}_{ir} \tilde{x}_{ist} + \tilde{x}_{irt} \tilde{x}_{is} + \tilde{x}_{irs} \tilde{x}_{it}) - \tilde{w}_{1i} \tilde{x}_{irt} \tilde{x}_{is} \tilde{x}_{iu} \\ & + \frac{1}{4} w_{1i} (\tilde{x}_{ir} \tilde{x}_{istu} + \tilde{x}_{iru} \tilde{x}_{ist} + \tilde{x}_{irs} \tilde{x}_{itu} + \tilde{x}_{irsu} \tilde{x}_{it} + \tilde{x}_{irst} \tilde{x}_{iu}) \\ & \left. - \frac{3}{4} w_{1i} (\tilde{x}_{irt} \tilde{x}_{isu} + \tilde{x}_{irtu} \tilde{x}_{is}) \right\}. \end{aligned}$$

Invertendo a ordem das somas e rearranjando os termos, obtemos

$$\begin{aligned} \sum \lambda_{rstu} = & \sum_{i=1}^n q_{4i} \left(\sum \kappa^{rs} \tilde{x}_{ir} \tilde{x}_{is} \right) \left(\sum \kappa^{tu} \tilde{x}_{it} \tilde{x}_{iu} \right) \\ & + \sum_{i=1}^n \left(\tilde{w}_{1i} - \frac{1}{2} q_{1i} \right) \left(\sum \kappa^{rs} \tilde{x}_{ir} \tilde{x}_{is} \right) \left(\sum \kappa^{tu} \tilde{x}_{itu} \right) \\ & + \sum_{i=1}^n (\tilde{w}_{1i} - q_{1i}) \left(\sum \kappa^{rs} \kappa^{tu} \tilde{x}_{it} \tilde{x}_{isu} \tilde{x}_{ir} \right) - \frac{1}{2} w_{1i} \left(\sum \kappa^{rs} \kappa^{tu} \tilde{x}_{iru} \tilde{x}_{ist} \right) \\ & + \frac{1}{4} \sum i = 1_n w_{1i} \left(\sum \kappa^{rs} \tilde{x}_{irs} \right) \left(\sum \kappa^{tu} \tilde{x}_{itu} \right). \end{aligned}$$

Das definições dos elementos das matrizes Z , B , C e D , esta parcela se reduz a

$$\sum \lambda_{rstu} = \sum_{i=1}^n q_{4i} z_{ii}^2 + \sum_{i=1}^n \left(\tilde{w}_{1i} - \frac{1}{2} q_{1i} \right) z_{ii} d_{1i} + \sum_{i=1}^n (\tilde{w}_{1i} - q_{1i}) c_{ii} - \frac{1}{4} \sum_{i=1}^n w_{1i} (2b_{ii} - d_{1i}^2).$$

Em notação matricial, expressamos $\sum \lambda_{rstu}$ na forma

$$\sum \lambda_{rstu} = \iota^\top Z_d Q_4 Z_d \iota + \iota^\top \left(\tilde{W}_1 - \frac{1}{2} Q_1 \right) D Z_d \iota + \iota^\top (\tilde{W}_1 - Q_1) C_d \iota - \frac{1}{4} \iota^\top W_1 (2B - D^2) \iota,$$

em que ι é um vetor $n \times 1$ de uns, $\tilde{W}_1$ e W_1 são matrizes diagonais de dimensão $n \times n$, cujos elementos são $\tilde{w}_{1i}$ e w_{1i} , respectivamente, $Q_1 = \text{diag}\{q_{11}, \dots, q_{1n}\}$, $Q_2 = \text{diag}\{q_{21}, \dots, q_{2n}\}$, $Q_3 = \text{diag}\{q_{31}, \dots, q_{3n}\}$ e $Q_4 = \text{diag}\{q_{41}, \dots, q_{4n}\}$ são matrizes diagonais de dimensão $n \times n$ cujos elementos estão definidos em (4.10)–(4.13), respectivamente. De modo semelhante, obtemos a segunda parcela de ϵ_p , denotada por $\sum \lambda_{rstuvw}$, (vide Apêndice A) expressa, em

notação matricial, por

$$\begin{aligned} \sum \lambda_{rstuvw} = & \iota^\top Q_1 Z^{(3)} Q_2 \iota + \iota^\top \tilde{W}_1 Z^{(3)} \tilde{W}_1 \iota + \iota^\top Q_1 Z_d Z Z_d Q_3 \iota + \iota^\top \tilde{W}_1 Z_d Z Z_d \tilde{W}_1 \iota \\ & + \frac{1}{4} \iota^\top D W_1 Z \left[W_1 D + 4 Z_d \left(\tilde{W}_1 - \frac{1}{2} Q_1 \right) \right] \iota \\ & + tr \left\{ \left[\left(\tilde{W}_1 - Q_1 \right) C - \frac{1}{2} W_1 B \right] W_1 Z \right\}. \end{aligned}$$

A expressão de ϵ_p resultante é decomposta em

$$\epsilon_p = \epsilon_p^{(L)} + \epsilon_p^{(NL)}, \quad (4.14)$$

em que $\epsilon_p^{(L)}$ e $\epsilon_p^{(NL)}$ têm as seguintes formas matriciais:

$$\epsilon_p^{(L)} = \iota^\top Z_d Q_4 Z_d \iota + \iota^\top Q_1 Z^{(3)} Q_2 \iota + \iota^\top \tilde{W}_1 Z^{(3)} \tilde{W}_1 \iota + \iota^\top Q_1 Z_d Z Z_d Q_3 \iota + \iota^\top \tilde{W}_1 Z_d Z Z_d \tilde{W}_1 \iota, \quad (4.15)$$

que reflete a parte linear do modelo e

$$\begin{aligned} \epsilon_p^{(NL)} = & \iota^\top \left(\tilde{W}_1 - \frac{1}{2} Q_1 \right) D Z_d \iota + \iota^\top (\tilde{W}_1 - Q_1) C_d \iota - \frac{1}{4} \iota^\top W_1 (2B - D^2) \iota \\ & + \frac{1}{4} \iota^\top D W_1 Z \left[W_1 D + 4 Z_d \left(\tilde{W}_1 - \frac{1}{2} Q_1 \right) \right] \iota + tr \left\{ \left[\left(\tilde{W}_1 - Q_1 \right) C - \frac{1}{2} W_1 B \right] W_1 Z \right\}, \end{aligned} \quad (4.16)$$

que pode ser interpretado como um termo devido à não-linearidade na componente sistemática do modelo. Se $\eta(\cdot; \cdot)$ for linear, as quantidades d_{1i} , c_{ij} e b_{ij} se anulam e, consequentemente, $\epsilon_p^{(NL)} = 0$. Para os modelos pertencentes tanto à classe MSPNLG quanto à classe dos modelos não-lineares da família exponencial, vale ressaltar que a fórmula matricial dada em (4.15) coincide com a fórmula de Cordeiro (1983, p. 406, equação 4) e a fórmula dada em (4.16) coincide com a fórmula matricial dada em Cordeiro e Paula (1989, p. 97, equação 5).

Considere as matrizes $\tilde{X}_{2i} = \partial^2 \eta_i / \partial \beta_2 \partial \beta_2^\top$, $i = 1, \dots, n$, e $K_{\beta_2} = \tilde{X}_2^\top W \tilde{X}_2$. A fórmula de ϵ_{p-q} é definida analogamente à de ϵ_p dada em (4.14) com $\tilde{X}$, $\tilde{X}_i$ e K_β substituídos por $\tilde{X}_2$, $\tilde{X}_{2i}$ e K_{β_2} , respectivamente.

Os fatores de correção de Bartlett para o teste da razão de verossimilhanças discutidos na Seção 4.2 são obtidos de (4.7) com as quantidades ϵ_p e ϵ_{p-q} deduzidas de (4.15) e (4.16). É importante salientar que a fórmula dada em (4.14) somente envolve operações simples de matrizes e pode ser facilmente implementada em pacotes de computação simbólica e

linguagens que permitam executar operações simples de álgebra linear, tais como **0x**, **MAPLE**, **MATHEMATICA**, **S-Plus**, **R**, etc.

Na construção das matrizes W , W_1 , $\tilde{W}_1$, Q_1 , Q_2 , Q_3 e Q_4 , presentes nas equações (4.15)–(4.16), necessitamos da função de ligação com suas respectivas primeiras, segunda e terceira derivadas, das funções t e q com suas respectivas primeiras e segundas derivadas, das funções f , g e de variância com suas primeiras derivadas, respectivamente. Para obtermos as matrizes Z , Z_d , B , C , C_d e D precisamos da matriz modelo $\tilde{X}$ e das matrizes quadradas $\tilde{X}_i$, $i = 1, \dots, n$. Uma vez computadas as matrizes acima, o cálculo de ϵ_p na equação (4.14) é imediato. É óbvio que a expressão (4.14) depende muito do particular modelo adotado.

4.3 Resultados numéricos

Com o objetivo de avaliar o desempenho da correção de Bartlett para o teste da razão de verossimilhanças nos MSPNLGs, apresentamos, nesta seção, os resultados de simulações de Monte Carlo. Comparamos os desempenhos de três estatísticas de testes, isto é, a da razão de verossimilhanças com suas versões modificadas (LR^* e LR_1^*). Estes desempenhos são avaliados em função da proximidade das probabilidades de rejeição da hipótese nula, sendo esta verdadeira (probabilidade do erro tipo I) aos seus respectivos níveis nominais dos testes. Avaliamos também os poderes dos testes em estudo sob algumas situações.

O estudo de simulação foi baseado nas distribuições da classe dos MSPNLGs, a saber, Binomial Negativa Generalizada (BNG), Poisson Generalizada (GPO) e Consul. No caso da distribuição BNG, fixamos os parâmetros $\phi = 1$ e $\nu = 3$, já para a GPO e a Consul fixamos $\phi = 0, 2$ e $\phi = 1$, respectivamente. Este estudo foi desenvolvido utilizando a linguagem de programação matricial **0x** (Doornik, 2001) para 10000 amostras de Monte Carlo, enquanto que os gráficos foram construídos utilizando o pacote estatístico **R** na versão 2.8.0 (Venables e Ripley, 2002). As amostras consideradas foram de tamanhos $n = 20, 30, 40, 50$ e os níveis nominais considerados foram $\alpha = 1\%, 5\%$ e 10% .

Para cada tamanho da amostra e cada nível considerado, calculamos as taxas de rejeição de cada estatística de teste, isto é, estimamos via simulação $P(LR \geq \chi_{(\alpha; q)}^2)$, $P(LR^* \geq \chi_{(\alpha; q)}^2)$ e $P(LR_1^* \geq \chi_{(\alpha; q)}^2)$, em que $\chi_{(\alpha; q)}^2$ é o percentil $(1 - \alpha)$ da distribuição χ_q^2 .

4.3.1 Modelos lineares

Nesta seção, apresentaremos os resultados de simulações referentes aos modelos em séries de potências lineares generalizados, cujo preditor linear é dado por

$$\eta_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij}, \quad i = 1, \dots, n \text{ e } k = 1, \dots, 8.$$

A hipótese nula considerada foi $H_0 : \beta_5 = \beta_6 = 0$ e a variável resposta foi gerada assumindo que $\beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_7 = \beta_8 = 0,05$. As covariáveis $x_1, \dots, x_8$ foram tomadas como amostras aleatórias das seguintes distribuições: $U(0, 1)$, $F(2, 5)$, *Cauchy*, $N(0, 1)$, t_3 , $LN(0, 1)$, χ_3^2 e $F(3, 3)$.

Assumindo diversas distribuições para a variável resposta, na Tabela 4.1 temos as taxas de rejeição dos testes nos modelos com $p = 8$ e diferentes tamanhos de amostra. Podemos observar, conforme esperado, que os testes baseados nas estatísticas da razão de verossimilhanças modificadas, LR^* e LR_1^* , apresentam melhores desempenhos do que o teste baseado na estatística da razão de verossimilhanças usual, LR . Notamos ainda que este teste é bastante liberal, apresentando taxas de rejeição bastante superiores aos níveis nominais correspondentes, principalmente quando o tamanho da amostra é pequeno. Para $n = 20$ e $\alpha = 10\%$, por exemplo, o teste baseado em LR apresenta nos modelos BNG, Consul e GPO taxas iguais a 15,4%, 17,8% e 15,8%, respectivamente, enquanto que as taxas correspondentes fornecidas pelo teste baseado em LR^* são de 10,8%, 9,9% e 10,8%, respectivamente, e as do teste baseado em LR_1^* são de 9,8%, 9,4% e 10,0%, respectivamente. No entanto, conforme cresce o tamanho da amostra, as taxas de rejeição do teste usual vão se aproximando dos respectivos níveis nominais. Já para os testes modificados, quando o tamanho da amostra aumenta, as taxas de rejeição permanecem mais estáveis em relação aos respectivos níveis nominais se comparadas às taxas do teste baseado em LR .

Com o objetivo de analisar a influência do número de parâmetros de perturbação no desempenho dos testes, fixamos o tamanho da amostra em $n = 30$ e consideramos os seguintes preditores lineares:

1. $\eta_i = \beta_0 + \beta_5 x_{i5} + \beta_6 x_{i6} \quad (p = 3),$
2. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_5 x_{i5} + \beta_6 x_{i6} \quad (p = 4),$

$$3. \eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_5 x_{i5} + \beta_6 x_{i6} \quad (p = 5),$$

$\vdots$

$$8. \eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \beta_7 x_{i7} + \beta_8 x_{i8} \quad (p = 9).$$

Vale lembrar que a hipótese nula a ser testada é $H_0 : \beta_5 = \beta_6 = 0$, ou seja, temos fixo o número de parâmetros de interesse em $q = 2$. A Tabela 4.2 apresenta as taxas de rejeição dos testes dos referidos modelos. Podemos notar que para um número pequeno de parâmetros, todos os testes apresentam taxas de rejeição bem próximas aos níveis nominais correspondentes. Porém, o teste baseado na estatística de verossimilhanças LR torna-se bastante liberal à medida que aumentamos o número de parâmetros de perturbação, ou seja, suas taxas de rejeição tornam-se consideravelmente superiores ao nível de significância correspondente, enquanto para os demais testes, as taxas continuam estáveis. Este fato é mais notório no modelo Consul, em que para $p = 9$ e $\alpha = 5\%$, por exemplo, as taxas de rejeição dos testes baseados nas estatísticas LR , LR^* e LR_1^* são, respectivamente, 10,1%, 5,9% e 5,1%. Esses resultados também podem ser visualizados nos gráficos das Figuras 4.1, 4.2 e 4.3, os quais mostram as distorções dos tamanhos dos testes em relação ao nível nominal para os modelos BNG, Consul e GPO, respectivamente, e diferentes níveis nominais. Definimos como distorção do tamanho do teste a diferença entre as taxas de rejeição e o nível nominal correspondente. Através desses gráficos, observamos que o aumento do número de parâmetros faz com que o teste LR forneça tamanho estimado bastante distorcido. Notamos ainda que entre os testes corrigidos, o impacto do número de parâmetros é bem menos marcante no teste baseado na estatística LR_1^* .

Os resultados apresentados na Tabela 4.3 foram obtidos levando em consideração a hipótese alternativa $H_1 : \beta_5 = \beta_6 \neq 0$ para $n = 30$, $p = 4$, $\alpha = 5\%$ e diferentes valores de $\beta_5 = \beta_6 = \beta^{(0)}$, com $\beta^{(0)}$ variando de 0,05 a 0,35. Visto que o teste LR é bastante liberal, as simulações foram feitas com os valores críticos estimados, ou seja, com os quantis da distribuição empírica de LR , em vez dos valores tabulados para que todos os testes pudessem ter o mesmo tamanho. A partir desses resultados, observamos que o poder do teste LR é ligeiramente superior aos dos testes corrigidos, já estes apresentam poderes praticamente iguais. Para o modelo BNG, por exemplo, quando $\beta^{(0)} = 0,25$, as estimativas dos poderes dos testes LR , LR^* e LR_1^* são, respectivamente, 94,4%, 94,2% e 94,2%.

Tabela 4.1: Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ de acordo com as estatísticas dos testes LR , LR^* e LR_1^* em modelos lineares com $p = 8$ e diversos valores de n .

n	$\alpha(\%)$	Modelo BNG			Modelo Consul			Modelo GPO		
		LR	LR^*	LR_1^*	LR	LR^*	LR_1^*	LR	LR^*	LR_1^*
20	1	2,5	1,3	1,3	2,7	1,1	2,0	2,5	1,2	1,4
	5	9,0	5,7	5,3	10,9	5,1	5,4	9,0	5,8	5,4
	10	15,4	10,8	9,8	17,8	9,9	9,4	15,8	10,8	10,0
30	1	1,8	1,2	1,2	2,3	1,4	1,2	1,8	1,2	1,2
	5	7,5	5,7	5,5	8,1	5,7	5,1	7,5	5,6	5,3
	10	13,3	10,9	10,5	14,3	10,5	9,7	13,3	10,8	10,5
40	1	1,6	1,2	1,2	1,7	1,2	1,1	1,4	1,2	1,1
	5	5,9	4,9	4,8	6,9	5,2	5,0	6,1	5,0	4,9
	10	11,6	9,8	9,6	12,6	10,1	9,7	11,5	10,0	9,8
50	1	1,3	1,0	1,0	1,5	1,1	1,1	1,3	1,0	1,0
	5	5,6	4,8	4,8	6,1	5,1	5,0	5,7	5,0	5,0
	10	11,1	9,9	9,8	11,5	9,9	9,8	10,9	9,7	9,6

Tabela 4.2: Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ de acordo com as estatísticas dos testes LR , LR^* e LR_1^* em modelos lineares com $n = 30$ e diversos valores de p .

$\alpha(\%)$	p	Modelo BNG			Modelo Consul			Modelo GPO		
		LR	LR^*	LR_1^*	LR	LR^*	LR_1^*	LR	LR^*	LR_1^*
1	3	1,1	1,0	1,0	1,0	0,9	0,9	1,1	1,0	1,0
	4	1,0	1,0	1,0	1,0	0,9	0,9	1,0	0,9	0,9
	5	1,2	1,0	1,0	1,2	1,1	1,0	1,0	0,9	0,9
	6	1,3	1,1	1,1	1,3	1,0	1,1	1,3	1,0	1,0
	7	1,4	1,1	1,0	1,7	1,1	1,1	1,4	1,1	1,1
	8	1,8	1,2	1,2	2,3	1,4	1,2	1,8	1,2	1,2
	9	2,3	1,3	1,2	3,1	1,3	1,2	2,3	1,3	1,1
5	3	5,3	5,1	5,1	5,1	4,9	4,9	5,1	5,0	5,0
	4	5,3	5,1	5,1	5,1	4,7	4,7	5,3	5,2	5,2
	5	5,7	5,2	5,1	5,5	4,5	4,5	5,4	5,0	4,9
	6	5,9	5,1	5,0	6,3	5,0	4,9	5,8	5,1	5,0
	7	6,7	5,4	5,2	7,5	5,4	5,2	6,6	5,4	5,3
	8	7,5	5,7	5,5	8,1	5,7	5,1	7,5	5,6	5,3
	9	8,4	5,7	5,3	10,1	5,9	5,1	8,6	5,9	5,4
10	3	10,6	10,2	10,2	10,5	10,2	10,2	10,2	10,0	10,0
	4	10,5	10,1	10,1	10,7	10,2	10,2	10,3	10,0	10,0
	5	11,0	10,2	10,2	11,6	10,4	10,4	11,1	10,4	10,4
	6	12,0	10,9	10,8	12,4	10,7	10,5	12,0	10,8	10,8
	7	12,8	10,8	10,6	13,7	10,9	10,5	12,7	10,9	10,7
	8	13,3	10,9	10,5	14,3	10,5	9,7	13,3	10,8	10,5
	9	14,9	11,1	10,3	16,8	10,9	9,7	14,6	11,0	10,1

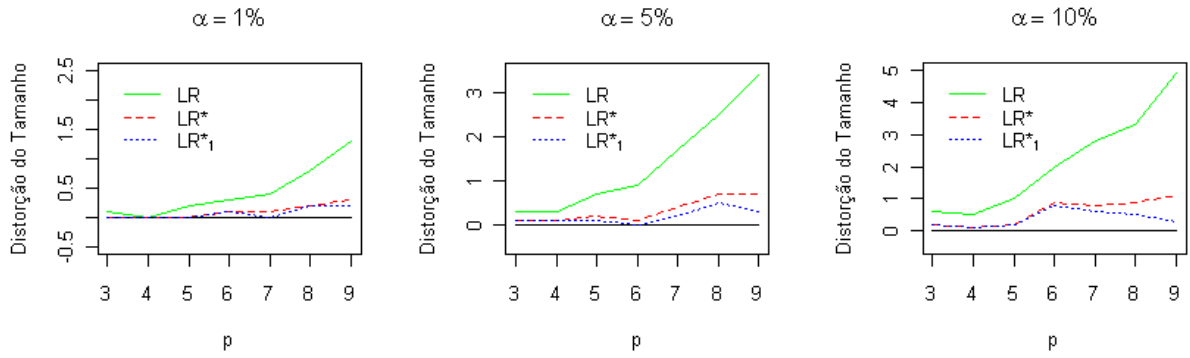


Figura 4.1: Distorção de tamanhos dos testes no modelo linear BNG, com $n = 30$.

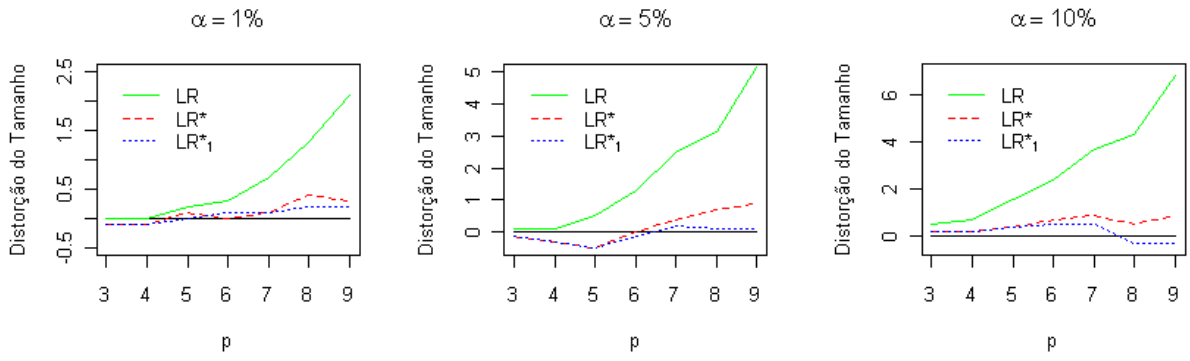


Figura 4.2: Distorção de tamanhos dos testes no modelo linear Consul, com $n = 30$.

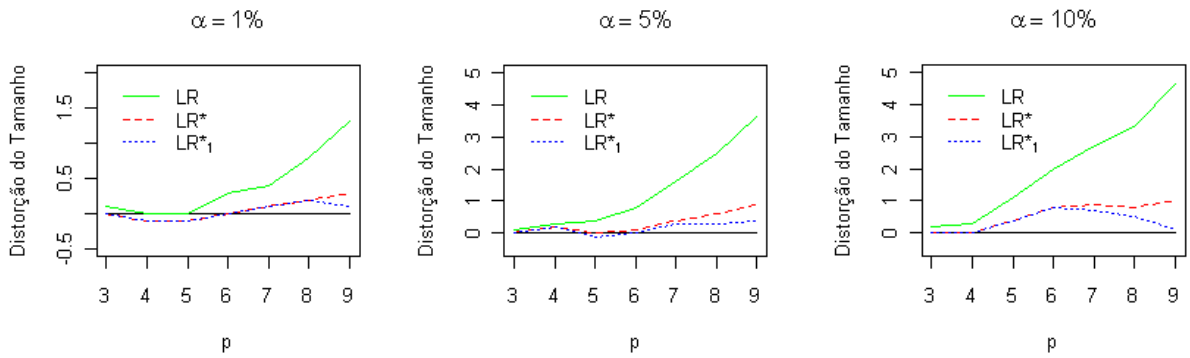


Figura 4.3: Distorção de tamanhos dos testes no modelo linear GPO, com $n = 30$.

Tabela 4.3: Poder dos testes em modelos lineares com $n = 30$, $p = 4$ e $\alpha = 5\%$.

$\beta^{(0)}$	Modelo BNG			Modelo Consul		
	LR	LR^*	LR_1^*	LR	LR^*	LR_1^*
0,05	8,3	7,9	7,9	7,1	6,9	6,8
0,10	20,6	20,2	20,2	13,8	13,2	13,2
0,15	46,8	46,4	46,4	27,0	26,2	26,1
0,20	76,9	76,6	76,6	46,4	45,3	45,3
0,25	94,4	94,2	94,2	66,9	66,0	66,0
0,30	99,3	99,3	99,3	83,3	82,7	82,7
0,35	100,0	100,0	100,0	92,4	92,3	92,3

4.3.2 Modelos não-lineares

Os resultados apresentados nesta seção são referentes aos modelos em séries de potências não-lineares generalizados, cujo preditor não-linear é dado por

$$\eta_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \exp(\beta_8 x_{i8}), \quad i = 1, \dots, n \text{ e } k = 1, \dots, 7.$$

Novamente consideramos a hipótese nula $H_0 : \beta_5 = \beta_6 = 0$ e a variável resposta gerada assumindo que $\beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_7 = \beta_8 = 0,05$. As covariáveis $x_1, \dots, x_8$ foram tomadas como amostras aleatórias das seguintes distribuições: $LN(0, 1)$, $F(2, 5)$, $Cauchy$, χ_3^2 , $Beta(2, 3)$, $N(0, 2)$, $Exp(1)$ e $U(0, 1)$.

A Tabela 4.4 mostra os resultados obtidos nos modelos em que fixamos o número de parâmetros em $p = 8$ e variamos o tamanho da amostra. Os resultados indicam que, em amostras de tamanho pequeno, o teste da razão de verossimilhanças é notavelmente liberal, uma vez que suas taxas de rejeição são maiores do que os níveis nominais correspondentes. Neste caso, os testes corrigidos apresentaram um melhor desempenho ao fornecer taxas de rejeição próximas aos níveis nominais correspondentes. Para o modelo Consul, por exemplo,

quando $n = 20$ e $\alpha = 10\%$, temos que a taxa de rejeição fornecida pelo teste baseado na estatística LR excede 15% , ao passo que as taxas de rejeição dos testes que se baseiam nas estatísticas LR^* e LR_1^* são $9,3\%$ e $10,3\%$, respectivamente. Conforme o tamanho da amostra cresce, no entanto, as taxas de rejeição das estatísticas não-modificada e modificadas se aproximam dos correspondentes níveis nominais, e as correções vão sendo cada vez menos necessárias.

Para avaliar a influência do número de parâmetros de perturbação nos desempenhos dos testes nos modelos não-lineares, fixamos o tamanho da amostra em $n = 30$ e consideramos os seguintes preditores:

1. $\eta_i = \beta_5 x_{i5} + \beta_6 x_{i6} + \exp(\beta_8 x_{i8}) \quad (p = 3),$
2. $\eta_i = \beta_0 + \beta_5 x_{i5} + \beta_6 x_{i6} + \exp(\beta_8 x_{i8}) \quad (p = 4),$
3. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_5 x_{i5} + \beta_6 x_{i6} + \exp(\beta_8 x_{i8}) \quad (p = 5),$
- $\vdots$
8. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \beta_7 x_{i7} + \exp(\beta_8 x_{i8}) \quad (p = 9).$

Os resultados desse estudo estão apresentados na Tabela 4.5. Notamos que quando o número de parâmetros de perturbação é pequeno, tanto a estatística da razão de verossimilhanças usual como as versões corrigidas têm boa aproximação pela distribuição χ^2 . Mas, de maneira análoga aos modelos lineares, o aumento no número de parâmetros de perturbação provoca um aumento nas taxas de rejeição fornecidas pelo teste baseado na estatística LR , enquanto que para os demais testes, as taxas continuam estáveis. Novamente, no modelo Consul o ganho com o uso da correção de Bartlett é mais notório. Neste modelo, quando $p = 9$ e $\alpha = 10\%$, por exemplo, as taxas de rejeição dos testes baseados nas estatísticas LR , LR^* e LR_1^* são, respectivamente, $14,3\%$, $9,6\%$ e $9,0\%$. Os gráficos das Figuras 4.4, 4.5 e 4.6 apresentam as distorções dos tamanhos dos testes para os modelos não-lineares BNG, Consul e GPO, respectivamente, e diferentes níveis nominais. A partir desses gráficos, observamos que nos modelos BNG e Consul, para os níveis nominais $\alpha = 5\%$ e $\alpha = 10\%$, os tamanhos estimados do teste LR são bastante distorcidos, independente do número de parâmetros. O mesmo é visto no modelo GPO quando $\alpha = 10\%$. Para os demais casos, o impacto com o aumento do número de parâmetro é mais marcante no teste LR .

Os resultados apresentados na Tabela 4.6 foram obtidos levando em consideração a hipótese alternativa $H_1 : \beta_5 = \beta_6 \neq 0$ para $n = 30$, $p = 4$, $\alpha = 10\%$ e diferentes valores de $\beta_5 = \beta_6 = \beta^{(0)}$, com $\beta^{(0)}$ variando de 0,05 a 0,50. Analogamente aos modelos lineares, as simulações foram feitas com os valores críticos estimados em vez dos valores tabulados. Através desses resultados, observamos que o poder do teste LR é ligeiramente superior aos dos testes corrigidos, já estes apresentam poderes praticamente iguais. Para o modelo Consul, por exemplo, quando $\beta^{(0)} = 0,50$, os poderes estimados dos testes LR , LR^* e LR_1^* são, respectivamente, 91,2%, 90,6% e 90,6%.

Tabela 4.4: Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ de acordo com as estatísticas dos testes LR , LR^* e LR_1^* em modelos não-lineares com $p = 8$ e diversos valores de n .

n	$\alpha(\%)$	Modelo BNG			Modelo Consul			Modelo GPO		
		LR	LR^*	LR_1^*	LR	LR^*	LR_1^*	LR	LR^*	LR_1^*
20	1	2,2	1,1	2,0	2,5	1,2	2,3	1,8	1,3	1,6
	5	8,2	5,3	6,4	8,6	4,7	6,0	7,3	5,1	5,6
	10	14,7	10,1	11,1	15,5	9,3	10,3	12,9	9,6	10,0
30	1	1,5	1,1	1,1	1,6	1,1	1,2	1,4	1,1	1,0
	5	6,4	5,1	5,0	7,1	4,8	4,8	6,3	5,0	5,0
	10	12,1	10,0	9,9	13,3	9,8	9,6	11,9	10,1	9,9
40	1	1,4	1,1	1,2	1,7	1,2	1,1	1,4	1,3	1,3
	5	6,1	5,1	5,1	6,8	5,2	5,0	6,1	5,3	5,2
	10	11,7	10,0	9,8	12,7	10,3	10,0	11,9	10,4	10,2
50	1	1,1	0,9	0,9	1,3	0,9	0,9	1,0	0,9	0,9
	5	5,5	4,8	4,7	5,9	4,7	4,7	5,5	4,9	4,8
	10	11,0	9,8	9,7	11,7	10,0	9,8	10,8	10,0	9,9

Tabela 4.5: Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ de acordo com as estatísticas dos testes LR , LR^* e LR_1^* em modelos não-lineares com $n = 30$ e diversos valores de p .

$\alpha(\%)$	p	Modelo BNG			Modelo Consul			Modelo GPO		
		LR	LR^*	LR_1^*	LR	LR^*	LR_1^*	LR	LR^*	LR_1^*
1	3	1,3	0,9	0,9	1,1	0,9	0,9	0,9	0,8	0,8
	4	0,9	0,8	0,8	1,0	0,8	0,8	0,9	0,9	0,9
	5	0,9	0,8	0,8	0,9	0,8	0,8	1,0	0,9	0,9
	6	1,1	0,9	0,9	1,1	0,7	0,7	1,0	0,9	0,9
	7	1,2	0,9	1,1	1,4	0,7	1,0	1,2	0,9	0,9
	8	1,5	1,1	1,1	1,6	1,1	1,2	1,4	1,1	1,0
	9	1,7	1,0	1,3	2,1	1,1	1,1	1,6	1,1	1,2
5	3	5,8	5,1	5,1	5,8	5,0	5,0	5,3	5,0	5,0
	4	5,5	5,1	5,1	5,5	4,9	4,8	5,0	4,6	4,6
	5	5,7	5,3	5,3	5,8	5,0	5,0	5,3	5,1	5,1
	6	5,4	4,9	4,8	5,9	4,9	4,8	5,3	5,0	5,0
	7	6,5	5,1	5,1	7,0	4,9	5,0	6,1	4,8	4,8
	8	6,4	5,1	5,0	7,1	4,8	4,8	6,3	5,0	5,0
	9	6,7	5,0	5,0	7,9	5,0	4,6	6,5	5,0	5,0
10	3	11,4	10,2	10,1	11,5	10,2	10,2	10,9	10,4	10,4
	4	11,0	10,4	10,4	11,2	10,2	10,1	10,7	10,2	10,2
	5	10,9	10,3	10,3	11,3	10,2	10,1	10,9	10,5	10,5
	6	10,9	10,1	10,1	11,6	10,2	10,1	10,6	10,1	10,1
	7	12,4	10,0	10,1	13,2	9,8	9,8	12,0	10,0	9,9
	8	12,1	10,0	9,9	13,3	9,8	9,6	11,9	10,1	9,9
	9	12,9	9,8	9,8	14,3	9,6	9,0	12,6	10,2	10,1

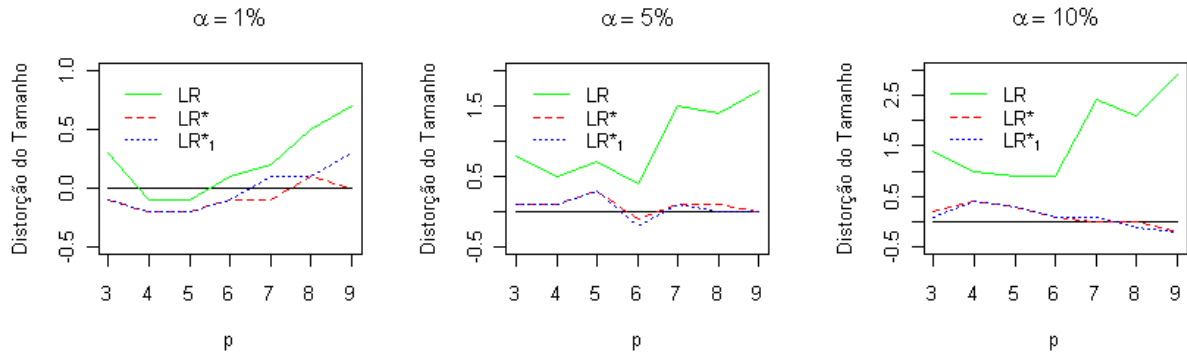


Figura 4.4: Distorção de tamanhos dos testes no modelo não-linear BNG, com $n = 30$.

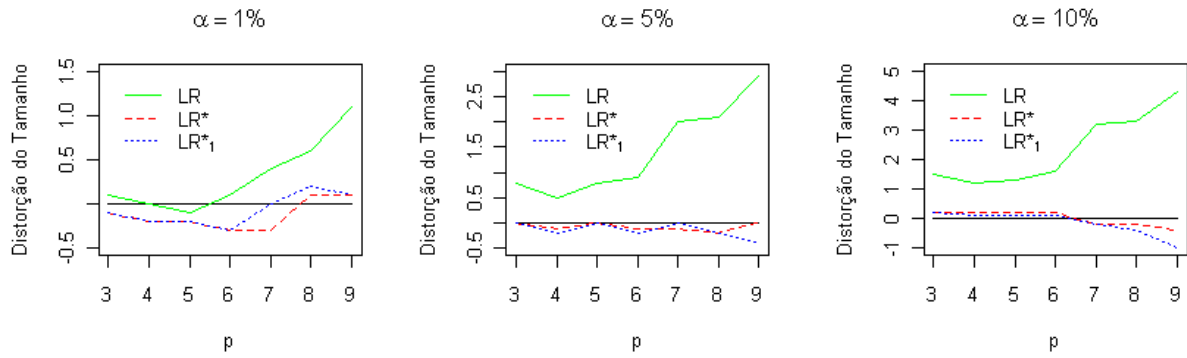


Figura 4.5: Distorção de tamanhos dos testes no modelo não-linear Consul, com $n = 30$.

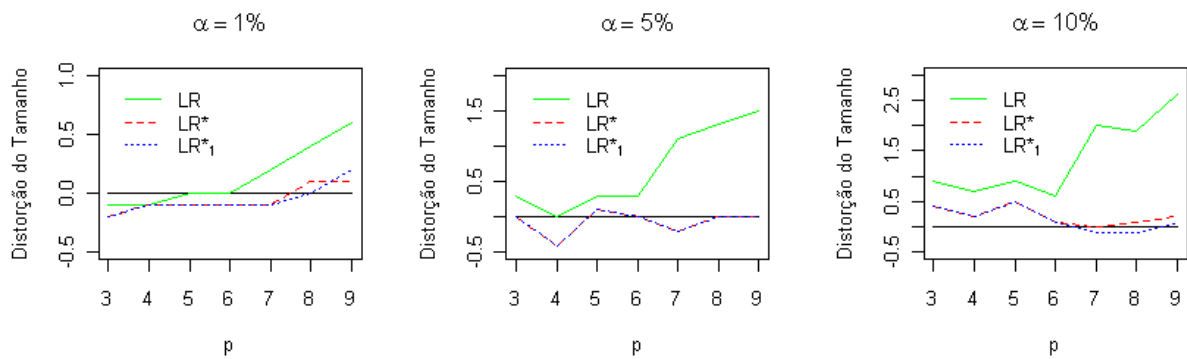


Figura 4.6: Distorção de tamanhos dos testes no modelo não-linear GPO, com $n = 30$.

Tabela 4.6: Poder dos testes em modelos não-lineares com $n = 30$, $p = 4$ e $\alpha = 10\%$.

$\beta^{(0)}$	Modelo Consul			Modelo GPO		
	LR	LR^*	LR_1^*	LR	LR^*	LR_1^*
0,05	10,4	9,6	9,5	17,6	17,6	17,6
0,10	14,1	13,1	13,1	38,1	38,3	38,3
0,15	20,4	19,0	18,9	61,3	61,4	61,4
0,20	29,3	27,8	27,8	77,9	78,0	78,0
0,25	40,0	38,5	38,4	87,1	87,1	87,1
0,30	52,1	50,6	50,5	91,4	91,3	91,4
0,35	65,0	63,5	63,5	93,6	93,3	93,4
0,40	75,8	74,7	74,6	94,9	94,1	94,2
0,45	85,1	84,2	84,1	95,7	93,5	93,8
0,50	91,2	90,6	90,6	96,0	91,6	91,8

4.4 Aplicação

Nesta seção, aplicaremos a metodologia apresentada nas seções anteriores ao conjunto de dados reais (Tabela B.1) referentes ao número de espécies de peixes em um lago (variável resposta) e o logaritmo da área do lago, em km^2 , (x) . Esses dados foram analisados inicialmente por Barbour e Brown (1974) e, posteriormente, por Rigby et al. (2008) e por Cordeiro *et al.* (2009). Estes últimos discutem a flexibilidade dos MSPNLGs em ajustar esses dados, adotando como preditores lineares

$$\eta_i = \beta_0 + \beta_1 \log(x_i) \quad (4.17)$$

e

$$\eta_i = \beta_0 + \beta_1 \log(x_i) + \beta_2 \{\log(x_i)\}^2, \quad (4.18)$$

$i = 1, \dots, 70$, sendo que $\eta_i = \log(\mu_i - m)$, em que m denota o valor mínimo do suporte da distribuição associada ao modelo. O nosso objetivo é testar a hipótese $H_0 : \beta_2 = 0$ contra

$H_1 : \beta_2 \neq 0$, ou seja, se o modelo (4.17) é mais adequado para representar o conjunto de dados do que o modelo (4.18). Com esse propósito, consideramos os modelos analisados por Cordeiro *et al.* (2009), a saber: Poisson, Binomial Negativa (BN), Poisson Generalizada (GPO), Binomial Negativa Generalizada (BNG) e a Delta Binomial (DB). A Tabela 4.7 apresenta os resultados, nesses modelos, dos testes baseados nas estatísticas LR , LR^* e LR_1^* e os respectivos níveis descritivos. Observamos que ao considerarmos o nível nominal 10%, todos os testes considerados rejeitam a hipótese nula nos modelos Poisson e BNG, mas não rejeitam H_0 nos modelos BN e GPO. Já no modelo DB, para esse mesmo nível nominal, os testes em estudo conduzem a conclusões divergentes e apenas o teste baseado na estatística LR rejeita H_0 . Pelo Critério de Informação de Akaike (AIC), dentre os modelos analisados, o modelo Delta Binomial com o preditor linear (3.10) foi o modelo mais adequado para o ajuste do número de espécies de peixe, uma vez que forneceu o menor AIC, a saber 612,1. Este resultado coincide com o resultado obtido por Cordeiro *et al.* (2009).

Tabela 4.7: Valor das estatísticas dos testes e p -valor (entre parênteses) para alguns modelos ajustados aos dados da Tabela B.1, considerando (4.18) como preditor linear.

Distribuições	Estatística		
	LR	LR^*	LR_1^*
Poisson	46,8610	46,8270	46,8270
	(0,0000)	(0,0000)	(0,0000)
BN	2,4882	2,4439	2,4431
	(0,1147)	(0,1180)	(0,1180)
GPO ($\phi = 0, 3$)	0,7701	0,7361	0,7346
	(0,3802)	(0,3909)	(0,3914)
BNG ($\phi = 1, \nu = 2, 43$)	5,6522	5,6129	5,6126
	(0,0174)	(0,0178)	(0,0178)
DB ($\phi = 3, m = 5$)	2,7447	2,5624	2,5494
	(0,0976)	(0,1094)	(0,1103)

4.5 Comentários

Neste capítulo, derivamos o fator da correção de Bartlett para a estatística da razão de verossimilhanças nos modelos em séries de potência não-lineares generalizados (MSPNLG). Abordamos o caso em que o parâmetro de dispersão é conhecido. Resultados numéricos, apresentados tanto para os modelos não-lineares como para os modelos lineares, mostraram o desempenho, em amostras finitas, dos testes baseados na estatística da razão de verossimilhanças original (LR) e nas estatísticas da razão de verossimilhanças corrigidas LR^* e LR_1^* .

Concluimos desses resultados que os testes baseados nas estatísticas corrigidas LR^* e LR_1^* tiveram um melhor desempenho do que o teste baseado na estatística da razão de verossimilhanças original. Isso porque ao utilizarmos as estatísticas LR^* e LR_1^* , obtivemos taxas de rejeição bem mais próximas do nível nominal do que os fornecidos pela estatística LR . Este, por sua vez, apresentou taxas de rejeição superiores aos níveis nominais correspondentes, principalmente quando o tamanho de amostra foi pequeno e/ou o número de parâmetros de perturbação foi grande, o que o torna liberal. Nestes casos, o ganho com a correção de Bartlett aplicada à estatística da razão de verossimilhanças ficou ainda mais evidente. O uso dessa correção atenua o efeito do tamanho da amostra e do número de parâmetros de perturbação e, conseqüentemente, as taxas de rejeição fornecidas pelos testes baseados em LR^* e LR_1^* continuaram próximas do nível nominal mesmo quando o tamanho da amostra era pequeno e/ou o número de parâmetro de perturbação aumentou. As simulações do poder do teste foram feitas usando os valores críticos estimados, ou seja, com os quantis da distribuição empírica de LR , em vez dos valores tabulados, de maneira que todos os testes pudessem ter o mesmo tamanho. Os testes baseados nas estatísticas da razão de verossimilhanças corrigidas se tornam consideravelmente menos poderosos do que o teste da razão de verossimilhanças original. Isto ilustra o fato de que em alguns caso o tamanho ajustado implica em alguma diminuição do poder.

Por fim, levando em conta as simulações envolvendo tamanho e poder dos testes nos modelos lineares e não-lineares, recomendamos o uso dos testes baseados nas estatísticas da razão de verossimilhanças corrigidas LR^* e LR_1^* em vez da estatística LR em inferências sobre os parâmetros do MSPNLG.

Capítulo 5

Conclusões

Resumimos as principais contribuições teóricas desta dissertação nos seguintes itens:

- (i) No Capítulo 3, derivamos uma expressão em forma matricial para o viés de segunda ordem de Cox & Snell (1968) para os estimadores de máxima verossimilhança dos parâmetros dos modelos em séries de potência não-lineares generalizados. Os resultados cobrem a situação em que o parâmetro de dispersão e , no caso da distribuição Binomial Negativa Generalizada, o parâmetro ν são conhecidos.
- (ii) No Capítulo 4, derivamos uma expressão em forma matricial do fator de correção de Bartlett usado para aperfeiçoar o teste baseado na estatística da razão de verossimilhanças (LR) na classe dos MSPNLGs, considerando fixos o parâmetro de dispersão e , no caso da distribuição Binomial Negativa Generalizada, o parâmetro ν .

Além dessas contribuições, estudos de simulação foram feitos com a finalidade de verificar o efeito das correções nos modelos em séries de potência lineares e não-lineares generalizados, dos quais podemos tirar as seguintes conclusões:

- (a) Os resultados das simulações de Monte Carlo para a correção de viés mostraram que, entre os estimadores em estudo, o estimador de máxima verossimilhança corrigido via Cox & Snell é o mais eficaz, em termos do viés, no caso dos modelos lineares. Para os modelos não-lineares ambos estimadores de máxima verossimilhança corrigidos via Cox & Snell e via bootstrap apresentam um bom desempenho. As correções nos estimadores

se fazem necessárias mesmo com tamanhos de amostras consideráveis, uma vez que estas produzem, em qualquer tamanho de amostra, estimadores mais eficientes tanto em termos de viés, como em termos de viés relativo e erro quadrático médio, do que o estimador de máxima verossimilhança usual.

- (b) Os resultados das simulações de Monte Carlo para avaliar o desempenho dos testes baseados na estatística LR e nas suas versões corrigidas LR^* e LR_1^* nos MSPNLG indicaram que o teste baseado na estatística da razão de verossimilhanças apresenta taxas de rejeição superiores aos níveis nominais correspondentes, ou seja, ele rejeita erroneamente a hipótese nula com uma probabilidade maior do que o nível de significância do teste. A correção de Bartlett mostra-se eficaz, produzindo testes com taxas de rejeição bem mais próximas do nível nominal, corrigindo a tendência liberal do teste original em rejeitar com maior frequência a hipótese nula. O tamanho amostral e o número de parâmetros de perturbação têm impacto considerável nas taxas de rejeição apresentadas pelo teste baseado na estatística LR . Já para os testes baseados nas estatísticas modificadas LR^* e LR_1^* , as taxas de rejeição permanecem mais estáveis. Em relação aos poderes dos testes, o teste baseado em LR apresentou uma leve vantagem. Os resultados indicam que não há nenhuma perda de poder derivada do fato de usar os fatores de correção de Bartlett obtidos nesta dissertação.

Apêndice A

Cálculo dos Momentos

Considere o modelo em séries de potência não-linear generalizado, apresentado na Seção 2.2, com o parâmetro de dispersão ϕ conhecido. Neste apêndice, apresentamos a obtenção de expressões gerais para as quantidades necessárias ao cálculo da correção de viés do estimador de máxima verossimilhança $\hat{\beta}$ do vetor de parâmetro β que indexa o referido modelo, utilizando a expressão geral de Cox & Snell (1968) apresentada na Seção 3.2.1. Apresentamos também a obtenção de expressões gerais para as quantidades necessárias ao cálculo do fator de correção de Bartlett para a estatística LR , apresentado na Seção 4.2.1.

O logaritmo da função de verossimilhança do parâmetro β , dado o vetor de observações $(y_1, \dots, y_n)^\top$, do MSPNLG tem a forma

$$l(\beta; y) = \sum_{i=1}^n \log\{a(y_i, \phi)\} + \sum_{i=1}^n [y_i \log\{g(\mu_i, \phi)\} - \log\{f(\mu_i, \phi)\}].$$

em que a função $a(y, \phi)$ é positiva e as funções analíticas $g(\mu_i, \phi)$ e $f(\mu_i, \phi)$ dos parâmetros μ_i e ϕ são positivas, finitas e duas vezes diferenciáveis. Denotando $f = f(\mu_i, \phi)$, $g = g(\mu_i, \phi)$ e o índice sobrescrito (j) indicando a j -ésima derivada em relação μ , $j = 1, 2, 3$, definimos a seguir quantidades, para $i = 1, \dots, n$, que serão de grande valia para a simplificação dos cálculos:

$$\begin{aligned}
t_i &= \frac{g_i^{(1)}}{g_i h_i^{(1)}}, \\
q_i &= \frac{f_i^{(1)}}{f_i h_i^{(1)}}, \\
d_{0i} &= y_i t_i - q_i, \\
d_{ji} &= \frac{y_i t_i^{(j)} - q_i^{(j)}}{(h_i^{(1)})^j}, \\
\varphi_{ji} &= \frac{q_i^{(1)} V_i t_i^{(j)} + q_i V_i^{(1)} t_i^{(j)} + q_i V_i t_i^{(j+1)}}{(h_i^{(1)})^j}, \\
\tilde{w}_{ji} &= \varphi_{ji} - (j-1) \frac{q_i V_i t_i^{(j)} h_i^{(2)}}{(h_i^{(1)})^{j+1}} - \frac{q_i^{(j+1)}}{(h_i^{(1)})^{j+1}} + j \frac{q_i^{(j)} h_i^{(2)}}{(h_i^{(1)})^{j+2}} \quad \text{e} \\
\tilde{w}_{ji}^* &= 2\varphi_{(j+1)i} - \frac{q_i V_i t_i^{(j+2)}}{(h_i^{(1)})^{j+1}} + \frac{t_i^{(j)} (q_i^{(2)} V_i + 2q_i^{(1)} V_i^{(1)} + q_i^{(1)} V_i^{(2)})}{(h_i^{(1)})^{j+1}} - (2j-1) \frac{h_i^{(2)} \varphi_{ji}}{(h_i^{(1)})^2} \\
&\quad + (j-1) \frac{q_i V_i t_i^{(j)}}{(h_i^{(1)})^{j+2}} \left[(j+1) \frac{(h_i^{(2)})^2}{h_i^{(1)}} - h_i^{(3)} \right] - \frac{q_i^{(j+2)}}{(h_i^{(1)})^{j+2}} + (2j+1) \frac{q_i^{(j+1)} h_i^{(2)}}{(h_i^{(1)})^{j+3}} \\
&\quad + j q_i^{(j)} \left[\frac{h_i^{(3)}}{(h_i^{(1)})^{j+3}} - (j+2) \frac{(h_i^{(2)})^2}{(h_i^{(1)})^{j+4}} \right].
\end{aligned}$$

Vale ressaltar que as quantidades acima envolvem derivadas que dependerá das formas específicas das funções f , g , h e V nas diversas distribuições pertencentes à família de série de potência. Denotando $\tilde{x}_{ir} = \partial \eta_i / \partial \beta_r$, $\tilde{x}_{irs} = \partial^2 \eta_i / \partial \beta_r \partial \beta_s$ e $\tilde{x}_{irst} = \partial^3 \eta_i / \partial \beta_r \partial \beta_s \partial \beta_t$, alguns resultados importantes são apresentados a seguir:

$$\begin{aligned}
E(d_{0i}) &= 0, \\
E(d_{ji}) &= w_{ji} = \frac{\mu_i t_i^{(j)} - q_i^{(j)}}{(h_i^{(1)})^j}, \\
\frac{\partial w_{ji}}{\partial \beta_r} &= \tilde{w}_{ji} \tilde{x}_{ir} \\
\frac{\partial \tilde{w}_{ji}}{\partial \beta_r} &= \tilde{w}_{ji}^* \tilde{x}_{ir}, \\
\frac{\partial d_{0i}}{\partial \beta_r} &= y_i \frac{\partial t_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_r} - \frac{\partial q_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_r} = y_i t_i^{(1)} \frac{1}{h_i^{(1)}} \tilde{x}_{ir} - q_i^{(1)} \frac{1}{h_i^{(1)}} \tilde{x}_{ir} = d_{1i} \tilde{x}_{ir},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial d_{ji}}{\partial \beta_r} &= \frac{(h_i^{(1)})^j [y_i t_i^{(j+1)} (h_i^{(1)})^{-1} \tilde{x}_{ir} - q_i^{(j+1)} (h_i^{(1)})^{-1} \tilde{x}_{ir}] - [y_i t_i^{(j)} - q_i^{(j)}] j (h_i^{(1)})^{j-1} h_i^{(2)} (h_i^{(1)})^{-1}}{(h_i^{(1)})^{2j}} \\
&= \frac{[y_i t_i^{(j+1)} - q_i^{(j+1)}]}{(h_i^{(1)})^{j+1}} \tilde{x}_{ir} - j \frac{[y_i t_i^{(j)} - q_i^{(j)}] h_i^{(2)}}{(h_i^{(1)})^{j+2}} \tilde{x}_{ir} = \left[d_{(j+1)i} - j \frac{d_{ji} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{ir}
\end{aligned}$$

e

$$\begin{aligned}
\frac{\partial d_{ji}}{\partial \beta_r \partial \beta_s} &= \left[d_{(j+2)i} - \frac{(j+1) d_{(j+1)i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{is} \tilde{x}_{ir} \\
&\quad - \left\{ \frac{j (h_i^{(1)})^2 [(d_{(j+1)i} - j d_{ji} h_i^{(2)} (h_i^{(1)})^{-2}) \tilde{x}_{is} h_i^{(2)} + d_{ji} h_i^{(3)} (h_i^{(1)})^{-1} \tilde{x}_{is}]}{(h_i^{(1)})^4} \right\} \tilde{x}_{ir} \\
&\quad + \left[\frac{2j d_{ji} (h_i^{(2)})^2 h_i^{(1)} (h_i^{(1)})^{-1} \tilde{x}_{is}}{(h_i^{(1)})^4} \right] \tilde{x}_{ir} + \left[d_{(j+1)i} - j \frac{d_{ji} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{irs} \\
&= \left[d_{(j+2)i} - \frac{(2j+1) d_{(j+1)i} h_i^{(2)}}{(h_i^{(1)})^2} - \frac{j d_{ji} h_i^{(3)}}{(h_i^{(1)})^3} + \frac{(j+2) j d_{ji} (h_i^{(2)})^2}{(h_i^{(1)})^4} \right] \tilde{x}_{ir} \tilde{x}_{is} \\
&\quad + \left[d_{(j+1)i} - j \frac{d_{ji} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{irs}.
\end{aligned}$$

A.1 Derivadas do logaritmo da função de verossimilhança

Por simples diferenciação em relação aos componentes do parâmetro β , temos

$$\begin{aligned}
U_r &= \frac{\partial l(\beta; y)}{\partial \beta_r} \\
&= \sum_i y_i \frac{1}{g(\mu_i, \phi)} \frac{\partial g(\mu_i, \phi)}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_r} - \sum_i \frac{1}{f(\mu_i, \phi)} \frac{\partial f(\mu_i, \phi)}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_r} \\
&= \sum_i y_i \frac{1}{g_i} g_i^{(1)} \frac{1}{h_i^{(1)}} \tilde{x}_{ir} - \sum_i \frac{1}{f_i} f_i^{(1)} \frac{1}{h_i^{(1)}} \tilde{x}_{ir} = \sum_i (y_i t_i - q_i) \tilde{x}_{ir} = \sum_i d_{0i} \tilde{x}_{ir}.
\end{aligned}$$

De forma análoga, as derivadas de segunda, terceira e quarta ordem podem ser obtidas do seguinte modo:

$$\begin{aligned}
U_{rs} &= \frac{\partial^2 l(\beta; y)}{\partial \beta_r \partial \beta_s} \\
&= \sum_i \left\{ \frac{\partial d_{0i}}{\partial \beta_s} \tilde{x}_{ir} + d_{0i} \frac{\partial \tilde{x}_{ir}}{\partial \beta_s} \right\} = \sum_i \left\{ d_{1i} \tilde{x}_{is} \tilde{x}_{ir} + d_{0i} \tilde{x}_{irs} \right\},
\end{aligned}$$

$$\begin{aligned}
U_{rst} &= \frac{\partial^3 l(\beta; y)}{\partial \beta_r \partial \beta_s \partial \beta_t} \\
&= \sum_i \left\{ \frac{\partial d_{1i}}{\partial \beta_t} \tilde{x}_{is} \tilde{x}_{ir} + d_{1i} \frac{\partial \tilde{x}_{is}}{\partial \beta_t} \tilde{x}_{ir} + d_{1i} \tilde{x}_{is} \frac{\partial \tilde{x}_{ir}}{\partial \beta_t} + \frac{\partial d_{0i}}{\partial \beta_t} \tilde{x}_{irs} + d_{0i} \frac{\partial \tilde{x}_{irs}}{\partial \beta_t} \right\} \\
&= \sum_i \left\{ \left[d_{2i} - \frac{d_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + d_{1i} \tilde{x}_{ist} \tilde{x}_{ir} + d_{1i} \tilde{x}_{is} \tilde{x}_{irt} + d_{1i} \tilde{x}_{it} \tilde{x}_{irs} + d_{0i} \tilde{x}_{irst} \right\} \\
&= \sum_i \left\{ \left[d_{2i} - \frac{d_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + d_{1i} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs}) + d_{0i} \tilde{x}_{irst} \right\}
\end{aligned}$$

e

$$\begin{aligned}
U_{rstu} &= \frac{\partial^4 l(\beta; y)}{\partial \beta_r \partial \beta_s \partial \beta_t \partial \beta_u} \\
&= \sum_i \left\{ \frac{\partial^2 d_{1i}}{\partial \beta_t \partial \beta_u} \tilde{x}_{is} \tilde{x}_{ir} + \left[d_{2i} - \frac{d_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{it} (\tilde{x}_{isu} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{iru}) \right. \\
&\quad + \frac{\partial d_{1i}}{\partial \beta_u} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs}) \\
&\quad + d_{1i} (\tilde{x}_{istu} \tilde{x}_{ir} + \tilde{x}_{ist} \tilde{x}_{iru} + \tilde{x}_{isu} \tilde{x}_{irt} + \tilde{x}_{is} \tilde{x}_{irtu} + \tilde{x}_{itu} \tilde{x}_{irs} + \tilde{x}_{it} \tilde{x}_{irsu}) \\
&\quad \left. + \frac{\partial d_{0i}}{\partial \beta_u} \tilde{x}_{irst} + d_{0i} \tilde{x}_{irstu} \right\} \\
&= \sum_i \left\{ \left[d_{3i} - \frac{3d_{2i} h_i^{(2)}}{(h_i^{(1)})^2} - \frac{d_{1i} h_i^{(3)}}{(h_i^{(1)})^3} + \frac{3d_{1i} (h_i^{(2)})^2}{(h_i^{(1)})^4} \right] \tilde{x}_{iu} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + \left[d_{2i} - \frac{d_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{itu} \tilde{x}_{is} \tilde{x}_{ir} \right. \\
&\quad + \left[d_{2i} - \frac{d_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{it} (\tilde{x}_{isu} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{iru}) + \left[d_{2i} - \frac{d_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{iu} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs}) \\
&\quad + d_{1i} (\tilde{x}_{istu} \tilde{x}_{ir} + \tilde{x}_{ist} \tilde{x}_{iru} + \tilde{x}_{isu} \tilde{x}_{irt} + \tilde{x}_{is} \tilde{x}_{irtu} + \tilde{x}_{itu} \tilde{x}_{irs} + \tilde{x}_{it} \tilde{x}_{irsu}) + d_{1i} \tilde{x}_{iu} \tilde{x}_{irst} \\
&\quad \left. + d_{0i} \tilde{x}_{irstu} \right\} \\
&= \sum_i \left\{ \left[d_{3i} - \frac{3d_{2i} h_i^{(2)}}{(h_i^{(1)})^2} - \frac{d_{1i} h_i^{(3)}}{(h_i^{(1)})^3} + \frac{3d_{1i} (h_i^{(2)})^2}{(h_i^{(1)})^4} \right] \tilde{x}_{iu} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} \right. \\
&\quad + \left[d_{2i} - \frac{d_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] [\tilde{x}_{itu} \tilde{x}_{is} \tilde{x}_{ir} + \tilde{x}_{it} \tilde{x}_{isu} \tilde{x}_{ir} + \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{iru} + \tilde{x}_{iu} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs})] \\
&\quad + d_{1i} (\tilde{x}_{istu} \tilde{x}_{ir} + \tilde{x}_{ist} \tilde{x}_{iru} + \tilde{x}_{isu} \tilde{x}_{irt} + \tilde{x}_{is} \tilde{x}_{irtu} + \tilde{x}_{itu} \tilde{x}_{irs} + \tilde{x}_{it} \tilde{x}_{irsu} + \tilde{x}_{iu} \tilde{x}_{irst}) \\
&\quad \left. + d_{0i} \tilde{x}_{irstu} \right\}.
\end{aligned}$$

A.2 Cálculo de cumulantes

Tomando as esperanças nas derivadas acima, obtemos os seguintes cumulantes:

$$\begin{aligned}
\kappa_{rs} &= \sum_i w_{1i} \tilde{x}_{is} \tilde{x}_{ir}, \\
\kappa_{rst} &= \sum_i \left\{ \left[w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + w_{1i} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs}) \right\} \quad \text{e} \\
\kappa_{rstu} &= \sum_i \left\{ \left[w_{3i} - \frac{3w_{2i} h_i^{(2)}}{(h_i^{(1)})^2} - \frac{w_{1i} h_i^{(3)}}{(h_i^{(1)})^3} + \frac{3w_{1i} (h_i^{(2)})^2}{(h_i^{(1)})^4} \right] \tilde{x}_{iu} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} \right. \\
&\quad + \left[w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] [\tilde{x}_{itu} \tilde{x}_{is} \tilde{x}_{ir} + \tilde{x}_{it} \tilde{x}_{isu} \tilde{x}_{ir} + \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{iru} + \tilde{x}_{iu} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs})] \\
&\quad \left. + w_{1i} (\tilde{x}_{istu} \tilde{x}_{ir} + \tilde{x}_{ist} \tilde{x}_{iru} + \tilde{x}_{isu} \tilde{x}_{irt} + \tilde{x}_{is} \tilde{x}_{irtu} + \tilde{x}_{itu} \tilde{x}_{irs} + \tilde{x}_{it} \tilde{x}_{irsu} + \tilde{x}_{iu} \tilde{x}_{irst}) \right\}.
\end{aligned}$$

A.2.1 Derivadas dos cumulantes

Calculando as derivadas das expressões da Seção A.2 em relação aos componentes de β , obtemos:

$$\begin{aligned}
\kappa_{rs}^{(t)} &= \frac{\partial \kappa_{rs}}{\partial \beta_t} = \sum_i \left\{ \tilde{w}_{1i} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + w_{1i} \tilde{x}_{ist} \tilde{x}_{ir} + w_{1i} \tilde{x}_{is} \tilde{x}_{irt} \right\}, \\
\kappa_{rs}^{(tu)} &= \frac{\partial^2 \kappa_{rs}}{\partial \beta_t \partial \beta_u} = \sum_i \left\{ \tilde{w}_{1i}^* \tilde{x}_{iu} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + \tilde{w}_{1i} \tilde{x}_{itu} \tilde{x}_{is} \tilde{x}_{ir} + \tilde{w}_{1i} \tilde{x}_{it} \tilde{x}_{isu} \tilde{x}_{ir} + \tilde{w}_{1i} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{iru} \right. \\
&\quad \left. + \tilde{w}_{1i} \tilde{x}_{iu} \tilde{x}_{ist} \tilde{x}_{ir} + w_{1i} \tilde{x}_{istu} \tilde{x}_{ir} + w_{1i} \tilde{x}_{ist} \tilde{x}_{iru} + \tilde{w}_{1i} \tilde{x}_{iu} \tilde{x}_{is} \tilde{x}_{irt} + w_{1i} \tilde{x}_{isu} \tilde{x}_{irt} + w_{1i} \tilde{x}_{is} \tilde{x}_{irtu} \right\}
\end{aligned}$$

e

$$\begin{aligned}
\kappa_{rst}^{(u)} &= \frac{\partial \kappa_{rst}}{\partial \beta_u} = \sum_i \left\{ \left[\tilde{w}_{2i} \tilde{x}_{iu} - \frac{(h_i^{(1)})^2 (\tilde{w}_{1i} \tilde{x}_{iu} h_i^{(2)} + w_{1i} h_i^{(3)} (h_i^{(1)})^{-1} \tilde{x}_{iu})}{(h_i^{(1)})^4} \right. \right. \\
&\quad \left. \left. + \frac{2h_i^{(1)} h_i^{(2)} (h_i^{(1)})^{-1} \tilde{x}_{iu} w_{1i} h_i^{(2)}}{(h_i^{(1)})^4} \right] \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} \right. \\
&\quad + \left[w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] (\tilde{x}_{itu} \tilde{x}_{is} \tilde{x}_{ir} + \tilde{x}_{it} \tilde{x}_{isu} \tilde{x}_{ir} + \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{iru}) \\
&\quad + \tilde{w}_{1i} \tilde{x}_{iu} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs}) \\
&\quad \left. + w_{1i} (\tilde{x}_{istu} \tilde{x}_{ir} + \tilde{x}_{ist} \tilde{x}_{iru} + \tilde{x}_{isu} \tilde{x}_{irt} + \tilde{x}_{is} \tilde{x}_{irtu} + \tilde{x}_{itu} \tilde{x}_{irs} + \tilde{x}_{it} \tilde{x}_{irsu}) \right\}
\end{aligned}$$

$$\begin{aligned}
&= \sum_i \left[\tilde{w}_{2i} - \frac{\tilde{w}_{1i} h_i^{(2)}}{(h_i^{(1)})^2} - \frac{w_{1i} h_i^{(3)}}{(h_i^{(1)})^3} + 2 \frac{w_{1i} (h_i^{(2)})^2}{(h_i^{(1)})^4} \right] \tilde{x}_{iu} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} \\
&+ \sum_i \left[w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right] (\tilde{x}_{itu} \tilde{x}_{is} \tilde{x}_{ir} + \tilde{x}_{it} \tilde{x}_{isu} \tilde{x}_{ir} + \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{iru}) \\
&+ \sum_i \tilde{w}_{1i} \tilde{x}_{iu} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs}) \\
&+ \sum_i w_{1i} (\tilde{x}_{istu} \tilde{x}_{ir} + \tilde{x}_{ist} \tilde{x}_{iru} + \tilde{x}_{isu} \tilde{x}_{irt} + \tilde{x}_{is} \tilde{x}_{irtu} + \tilde{x}_{itu} \tilde{x}_{irs} + \tilde{x}_{it} \tilde{x}_{irsu}).
\end{aligned}$$

A.3 Cálculo de $\sum \lambda_{rstuvw}$

Para obtenção, em forma matricial, do segundo termo $\sum \lambda_{rstuvw}$ de ϵ_p , sendo este definido em (4.3), consideramos $\tilde{x}_i^\top$ como sendo a i -ésima linha da matriz $\tilde{X}$, $\tilde{\tilde{X}}_i$ uma matriz $p \times p$ cujo elemento (r, s) é $\tilde{x}_{irs}$, $i = 1, \dots, n$, K_β^{-1} a inversa da matriz de informação de Fisher K_β dada pela equação (4.9), a matriz

$$Z = \tilde{X} (\tilde{X}^\top W \tilde{X})^{-1} \tilde{X}^\top$$

de dimensão $n \times n$ positiva semi-definida de posto p com elementos $z_{ij} = \tilde{x}_i^\top K_\beta^{-1} \tilde{x}_j$, as matrizes quadradas B e C de dimensão $n \times n$, com elementos $b_{ij} = \text{tr}(K_\beta^{-1} \tilde{\tilde{X}}_i K_\beta^{-1} \tilde{\tilde{X}}_j)$ e $c_{ij} = \tilde{x}_i^\top K_\beta^{-1} \tilde{\tilde{X}}_j K_\beta^{-1} \tilde{x}_i$, respectivamente, e a matriz diagonal $D = \text{diag}\{d_1, \dots, d_n\}$ com $d_i = \text{tr}(K_\beta^{-1} \tilde{\tilde{X}}_i)$. Utilizamos a notação Z_d , B_d e C_d para representar matrizes diagonais formadas pelos correspondentes elementos das diagonais das matrizes Z , B e C , respectivamente. Denotamos $Z^{(3)} = Z^{(2)} \odot Z$, $Z^{(2)} = Z \odot Z$, em que $\odot$ denota o produto de Hadamard (Rao, 1973, p. 30), ou seja, o elemento (i, j) de $Z^{(3)}$ é z_{ij}^3 . Adicionalmente, definimos as matrizes diagonais Q_1 , Q_2 e Q_3 de dimensão $n \times n$, cujos elementos estão definidos em 4.10, 4.11 e 4.12, respectivamente.

Substituindo os valores de κ 's, encontrados para os modelos MSPNLG na Seção A.2,

temos que

$$\begin{aligned}
\frac{1}{6}\kappa_{suw} - \kappa_{sw}^{(u)} &= \sum_i \left\{ \left[\frac{1}{6}w_{2i} - \frac{1}{6} \frac{w_{1i}h_i^{(2)}}{(h_i^{(1)})^2} - \tilde{w}_{1i} \right] \tilde{x}_{iw}\tilde{x}_{iu}\tilde{x}_{is} - \frac{5}{6}w_{1i}(\tilde{x}_{iuw}\tilde{x}_{is} + \tilde{x}_{iw}\tilde{x}_{isu}) \right. \\
&\quad \left. + \frac{1}{6}w_{1i}\tilde{x}_{iu}\tilde{x}_{isw} \right\} \text{ e} \\
\frac{1}{4}\kappa_{svw} - \kappa_{sw}^{(v)} &= \sum_i \left\{ \left[\frac{1}{4}w_{2i} - \frac{1}{4} \frac{w_{1i}h_i^{(2)}}{(h_i^{(1)})^2} - \tilde{w}_{1i} \right] \tilde{x}_{iw}\tilde{x}_{iv}\tilde{x}_{is} - \frac{3}{4}w_{1i}(\tilde{x}_{ivw}\tilde{x}_{is} + \tilde{x}_{iw}\tilde{x}_{isv}) \right. \\
&\quad \left. + \frac{1}{4}w_{1i}\tilde{x}_{iv}\tilde{x}_{isw} \right\}.
\end{aligned}$$

Das definições dos elementos q_{1i} , q_{2i} e q_{3i} na Seção 4.2.1, temos que

$$\begin{aligned}
\kappa_{rtv} \left(\frac{1}{6}\kappa_{suw} - \kappa_{sw}^{(u)} \right) + \kappa_{rt}^{(v)} \kappa_{sw}^{(u)} &= \sum_{i,j} \left\{ (q_{1i}q_{2j} + \tilde{w}_{1i}\tilde{w}_{1j})\tilde{x}_{iv}\tilde{x}_{it}\tilde{x}_{ir}\tilde{x}_{jw}\tilde{x}_{ju}\tilde{x}_{js} \right. \\
&\quad + \left(-\frac{5}{6}q_{1i} + \tilde{w}_{1i} \right) w_{1j}\tilde{x}_{iv}\tilde{x}_{it}\tilde{x}_{ir}(\tilde{x}_{juw}\tilde{x}_{js} + \tilde{x}_{jw}\tilde{x}_{jsu}) + \frac{1}{6}q_{1j}w_{1i}\tilde{x}_{jw}\tilde{x}_{ju}\tilde{x}_{js}(\tilde{x}_{itv}\tilde{x}_{ir} + \tilde{x}_{it}\tilde{x}_{irv}) \\
&\quad + \frac{1}{6}w_{1j}w_{1i}(\tilde{x}_{itv}\tilde{x}_{ir} + \tilde{x}_{it}\tilde{x}_{irv})(\tilde{x}_{juw}\tilde{x}_{js} + \tilde{x}_{jw}\tilde{x}_{jsu} + \tilde{x}_{ju}\tilde{x}_{jsw}) \\
&\quad + w_{1i}\tilde{x}_{iv}\tilde{x}_{irt} \left[q_{2j}\tilde{x}_{jw}\tilde{x}_{ju}\tilde{x}_{js} - \frac{5}{6}w_{1j}(\tilde{x}_{juw}\tilde{x}_{js} + \tilde{x}_{jw}\tilde{x}_{jsu}) \right] \\
&\quad \left. + \frac{1}{6}w_{1j}\tilde{x}_{ju}\tilde{x}_{jsw}(q_{1i}\tilde{x}_{iv}\tilde{x}_{it}\tilde{x}_{ir} + w_{1i}\tilde{x}_{iv}\tilde{x}_{irt}) \right\} \tag{A.1}
\end{aligned}$$

e

$$\begin{aligned}
\kappa_{rtu} \left(\frac{1}{4}\kappa_{svw} - \kappa_{sw}^{(v)} \right) + \kappa_{rt}^{(u)} \kappa_{sw}^{(v)} &= \sum_{i,j} \left\{ (q_{1i}q_{3j} + \tilde{w}_{1i}\tilde{w}_{1j})\tilde{x}_{iu}\tilde{x}_{it}\tilde{x}_{ir}\tilde{x}_{jw}\tilde{x}_{jv}\tilde{x}_{js} \right. \\
&\quad + \left(-\frac{3}{4}q_{1i} + \tilde{w}_{1i} \right) w_{1j}\tilde{x}_{iu}\tilde{x}_{it}\tilde{x}_{ir}(\tilde{x}_{jvw}\tilde{x}_{js} + \tilde{x}_{jw}\tilde{x}_{jsv}) + \frac{1}{4}q_{1j}w_{1i}\tilde{x}_{jw}\tilde{x}_{jv}\tilde{x}_{js}(\tilde{x}_{itu}\tilde{x}_{ir} + \tilde{x}_{it}\tilde{x}_{iru}) \\
&\quad + \frac{1}{4}w_{1j}w_{1i}(\tilde{x}_{itu}\tilde{x}_{ir} + \tilde{x}_{it}\tilde{x}_{iru})(\tilde{x}_{jvw}\tilde{x}_{js} + \tilde{x}_{jw}\tilde{x}_{jsv} + \tilde{x}_{jv}\tilde{x}_{jsw}) \\
&\quad + w_{1i}\tilde{x}_{iu}\tilde{x}_{irt} \left[q_{3j}\tilde{x}_{jw}\tilde{x}_{jv}\tilde{x}_{js} - \frac{3}{4}w_{1j}(\tilde{x}_{jvw}\tilde{x}_{js} + \tilde{x}_{jw}\tilde{x}_{jsv}) \right] \\
&\quad \left. + \frac{1}{4}w_{1j}\tilde{x}_{jv}\tilde{x}_{jsw}(q_{1i}\tilde{x}_{iu}\tilde{x}_{it}\tilde{x}_{ir} + w_{1i}\tilde{x}_{iu}\tilde{x}_{irt}) \right\}. \tag{A.2}
\end{aligned}$$

O próximo passo será multiplicar (A.1) e (A.2) por $\kappa^{rs}\kappa^{tu}\kappa^{vw}$, em que $-\kappa^{rs} = \kappa^{r,s}$ representa o elemento (r, s) da matriz K_β^{-1} , $r, s = 1, \dots, p$, e aplicar o somatório sobre todos os componentes do vetor β . Invertendo a ordem das somas e rearranjando os termos, obtemos

$$\begin{aligned}
& \sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \left[\kappa_{rtv} \left(\frac{1}{6} \kappa_{suw} - \kappa_{sw}^{(u)} \right) + \kappa_{rt}^{(v)} \kappa_{sw}^{(u)} \right] \\
&= \sum_{i,j} \left(q_{1i} q_{2j} + \tilde{w}_{1i} \tilde{w}_{1j} \right) \left(\sum \kappa^{rs} \tilde{x}_{ir} \tilde{x}_{js} \right) \left(\sum \kappa^{tu} \tilde{x}_{it} \tilde{x}_{ju} \right) \left(\sum \kappa^{vw} \tilde{x}_{iv} \tilde{x}_{jw} \right) \\
&\quad + \sum_{i,j} w_{1j} \left(-\frac{5}{6} q_{1i} + \tilde{w}_{1i} \right) \left(\sum \kappa^{vw} \tilde{x}_{iv} \tilde{x}_{jw} \right) \left(\sum \kappa^{rs} \kappa^{tu} \tilde{x}_{it} \tilde{x}_{ir} \tilde{x}_{jsu} \right) \\
&\quad + \sum_{i,j} w_{1j} \left(-\frac{5}{6} q_{1i} + \tilde{w}_{1i} \right) \left(\sum \kappa^{rs} \tilde{x}_{ir} \tilde{x}_{js} \right) \left(\sum \kappa^{tu} \kappa^{vw} \tilde{x}_{iv} \tilde{x}_{it} \tilde{x}_{juw} \right) \\
&\quad + \frac{1}{6} \sum_{i,j} q_{1j} w_{1i} \left(\sum \kappa^{rs} \tilde{x}_{ir} \tilde{x}_{js} \right) \left(\sum \kappa^{tu} \kappa^{vw} \tilde{x}_{ju} \tilde{x}_{jw} \tilde{x}_{itv} \right) \\
&\quad + \frac{1}{6} \sum_{i,j} q_{1j} w_{1i} \left(\sum \kappa^{tu} \tilde{x}_{ju} \tilde{x}_{it} \right) \left(\sum \kappa^{rs} \kappa^{vw} \tilde{x}_{js} \tilde{x}_{jw} \tilde{x}_{irv} \right) \\
&\quad + \frac{1}{6} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{rs} \tilde{x}_{ir} \tilde{x}_{js} \right) \left(\sum \kappa^{tu} \kappa^{vw} \tilde{x}_{itv} \tilde{x}_{juw} \right) \\
&\quad + \frac{1}{6} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \tilde{x}_{itv} \tilde{x}_{ir} \tilde{x}_{jsu} \tilde{x}_{jw} \right) \\
&\quad + \frac{1}{6} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \tilde{x}_{itv} \tilde{x}_{ir} \tilde{x}_{jsw} \tilde{x}_{ju} \right) \\
&\quad + \frac{1}{6} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \tilde{x}_{it} \tilde{x}_{irv} \tilde{x}_{juw} \tilde{x}_{js} \right) \\
&\quad + \frac{1}{6} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \tilde{x}_{it} \tilde{x}_{irv} \tilde{x}_{jsu} \tilde{x}_{jw} \right) \\
&\quad + \frac{1}{6} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{tu} \tilde{x}_{ju} \tilde{x}_{it} \right) \left(\sum \kappa^{rs} \kappa^{vw} \tilde{x}_{irv} \tilde{x}_{jsw} \right) \\
&\quad + \sum_{i,j} w_{1i} q_{2j} \left(\sum \kappa^{vw} \tilde{x}_{iv} \tilde{x}_{jw} \right) \left(\sum \kappa^{rs} \kappa^{tu} \tilde{x}_{irt} \tilde{x}_{ju} \tilde{x}_{js} \right) \\
&\quad - \frac{5}{6} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \tilde{x}_{iv} \tilde{x}_{irt} \tilde{x}_{juw} \tilde{x}_{js} \right) \\
&\quad - \frac{5}{6} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{vw} \tilde{x}_{iv} \tilde{x}_{jw} \right) \left(\sum \kappa^{rs} \kappa^{tu} \tilde{x}_{irt} \tilde{x}_{jsu} \right) \\
&\quad + \frac{1}{6} \sum_{i,j} w_{1j} q_{1i} \left(\sum \kappa^{tu} \tilde{x}_{ju} \tilde{x}_{it} \right) \left(\sum \kappa^{rs} \kappa^{vw} \tilde{x}_{jsw} \tilde{x}_{iv} \tilde{x}_{ir} \right) \\
&\quad + \frac{1}{6} w_{1i} w_{1j} \left(\sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \tilde{x}_{jsw} \tilde{x}_{ju} \tilde{x}_{irt} \tilde{x}_{iv} \right)
\end{aligned} \tag{A.3}$$

e

$$\begin{aligned}
& \sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \left[\kappa_{rtu} \left(\frac{1}{4} \kappa_{svw} - \kappa_{sw}^{(v)} \right) + \kappa_{rt}^{(u)} \kappa_{sw}^{(v)} \right] \\
&= \sum_{i,j} \left(q_{1i} q_{3j} + \tilde{w}_{1i} \tilde{w}_{1j} \right) \left(\sum \kappa^{rs} \tilde{x}_{ir} \tilde{x}_{js} \right) \left(\sum \kappa^{tu} \tilde{x}_{it} \tilde{x}_{iu} \right) \left(\sum \kappa^{vw} \tilde{x}_{jv} \tilde{x}_{jw} \right) \\
&+ \sum_{i,j} w_{1j} \left(-\frac{3}{4} q_{1i} + \tilde{w}_{1i} \right) \left(\sum \kappa^{rs} \tilde{x}_{ir} \tilde{x}_{js} \right) \left(\sum \kappa^{tu} \tilde{x}_{it} \tilde{x}_{iu} \right) \left(\sum \kappa^{vw} \tilde{x}_{jvw} \right) \\
&+ \sum_{i,j} w_{1j} \left(-\frac{3}{4} q_{1i} + \tilde{w}_{1i} \right) \left(\sum \kappa^{tu} \tilde{x}_{it} \tilde{x}_{iu} \right) \left(\sum \kappa^{rs} \kappa^{vw} \tilde{x}_{ir} \tilde{x}_{jw} \tilde{x}_{jv} \right) \\
&+ \frac{1}{4} \sum_{i,j} q_{1j} w_{1i} \left(\sum \kappa^{rs} \tilde{x}_{ir} \tilde{x}_{js} \right) \left(\sum \kappa^{vw} \tilde{x}_{jv} \tilde{x}_{jw} \right) \left(\sum \kappa^{tu} \tilde{x}_{itu} \right) \\
&+ \frac{1}{4} \sum_{i,j} q_{1j} w_{1i} \left(\sum \kappa^{vw} \tilde{x}_{jv} \tilde{x}_{jw} \right) \left(\sum \kappa^{rs} \kappa^{tu} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{iru} \right) \\
&+ \frac{1}{4} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{rs} \tilde{x}_{ir} \tilde{x}_{js} \right) \left(\sum \kappa^{tu} \tilde{x}_{itu} \right) \left(\sum \kappa^{vw} \tilde{x}_{jvw} \right) \\
&+ \frac{1}{4} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{tu} \tilde{x}_{itu} \right) \left(\sum \kappa^{rs} \kappa^{vw} \tilde{x}_{ir} \tilde{x}_{jw} \tilde{x}_{jv} \right) \\
&+ \frac{1}{4} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{vw} \tilde{x}_{jvw} \right) \left(\sum \kappa^{rs} \kappa^{tu} \tilde{x}_{it} \tilde{x}_{iru} \tilde{x}_{js} \right) \\
&+ \frac{1}{4} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{jv} \tilde{x}_{jw} \right) \\
&+ \frac{1}{4} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{tu} \tilde{x}_{itu} \right) \left(\sum \kappa^{rs} \kappa^{vw} \tilde{x}_{ir} \tilde{x}_{jw} \tilde{x}_{jv} \right) \\
&+ \frac{1}{4} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \tilde{x}_{it} \tilde{x}_{iru} \tilde{x}_{jw} \tilde{x}_{jv} \right) \\
&+ \sum_{i,j} w_{1i} q_{3j} \left(\sum \kappa^{vw} \tilde{x}_{jv} \tilde{x}_{jw} \right) \left(\sum \kappa^{rs} \kappa^{tu} \tilde{x}_{iu} \tilde{x}_{irt} \tilde{x}_{js} \right) \\
&+ \frac{3}{4} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{vw} \tilde{x}_{jvw} \right) \left(\sum \kappa^{rs} \kappa^{tu} \tilde{x}_{iu} \tilde{x}_{irt} \tilde{x}_{js} \right) \\
&- \frac{3}{4} \sum_{i,j} w_{1i} w_{1j} \left(\sum \kappa^{vw} \kappa^{rs} \kappa^{tu} \tilde{x}_{iu} \tilde{x}_{irt} \tilde{x}_{jv} \tilde{x}_{jw} \right) \\
&+ \frac{1}{4} \sum_{i,j} q_{1i} w_{1j} \left(\sum \kappa^{tu} \tilde{x}_{iu} \tilde{x}_{it} \right) \left(\sum \kappa^{rs} \kappa^{vw} \tilde{x}_{jw} \tilde{x}_{jv} \tilde{x}_{ir} \right) \\
&+ \frac{1}{4} \sum_{i,j} w_{1j} w_{1i} \left(\sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \tilde{x}_{iu} \tilde{x}_{irt} \tilde{x}_{jw} \tilde{x}_{jv} \right). \tag{A.4}
\end{aligned}$$

Adicionalmente, denotamos os termos $\sum \kappa^{rs} \kappa^{tu} \tilde{x}_{ir} \tilde{x}_{jt} \tilde{x}_{jsu}$, $-\sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{jsw} \tilde{x}_{ju}$ e $-\sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \tilde{x}_{it} \tilde{x}_{iru} \tilde{x}_{jsv} \tilde{x}_{jw}$ encontrados em (A.3) e (A.4) por a_{ij} , f_{ij} e g_{ij} , respectivamente. Uma vez que os demais termos, a saber $-\sum \tilde{x}_{ir} \kappa^{rs} \tilde{x}_{js}$, $\sum^{rs} \tilde{x}_{jsu} \kappa^{ut} \tilde{x}_{itr}$, $\sum \tilde{x}_{it} \kappa^{tu} \tilde{x}_{jus} \kappa^{sr} \tilde{x}_{ir}$ e $-\sum \kappa^{rs} \tilde{x}_{irs}$, correspondem aos elementos das matrizes Z , B , C e D , respectivamente, as parcelas (A.3) e (A.4) reduzem-se a

$$\begin{aligned}
& \sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \left[\kappa_{rtv} \left(\frac{1}{6} \kappa_{suw} - \kappa_{sw}^{(u)} \right) + \kappa_{rt}^{(v)} \kappa_{sw}^{(u)} \right] \\
&= \sum_{i,j} \left\{ - (q_{1i} q_{2j} + \tilde{w}_{1i} \tilde{w}_{1j}) z_{ij}^3 - 2 \left(-\frac{5}{6} q_{1i} + \tilde{w}_{1i} \right) w_{1j} z_{ij} c_{ij} - \frac{1}{3} q_{1j} w_{1i} z_{ij} c_{ji} \right. \\
&\quad - \frac{1}{6} w_{1i} w_{1j} z_{ij} b_{ij} - \frac{4}{6} w_{1i} w_{1j} f_{ij} - \frac{1}{6} w_{1i} w_{1j} z_{ij} b_{ij} - w_{1i} q_{2j} z_{ij} c_{ji} + \frac{5}{6} w_{1j} w_{1i} f_{ij} \\
&\quad \left. + \frac{5}{6} w_{1j} w_{1i} z_{ij} b_{ij} - \frac{1}{6} w_{1j} q_{1i} z_{ij} c_{ij} - \frac{1}{6} w_{1i} w_{1j} f_{ij} \right\} \\
&= \sum_{i,j} \left\{ - (q_{1i} q_{2j} + \tilde{w}_{1i} \tilde{w}_{1j}) z_{ij}^3 - \left(\tilde{w}_{1i} - q_{1i} \right) w_{1j} z_{ij} c_{ij} + \frac{1}{2} w_{1i} w_{1j} z_{ij} b_{ij} \right\}
\end{aligned}$$

e

$$\begin{aligned}
& \sum \kappa^{rs} \kappa^{tu} \kappa^{vw} \left[\kappa_{rtu} \left(\frac{1}{4} \kappa_{svw} - \kappa_{sw}^{(v)} \right) + \kappa_{rt}^{(u)} \kappa_{sw}^{(v)} \right] \\
&= - \sum_{i,j} (q_{1i} q_{3j} + \tilde{w}_{1i} \tilde{w}_{1j}) z_{ij} z_{ii} z_{jj} - \sum_{i,j} \left(-\frac{3}{4} q_{1i} + \tilde{w}_{1i} \right) w_{1j} z_{ij} z_{ii} d_j \\
&\quad - \sum_{i,j} \left(-\frac{3}{4} q_{1i} + \tilde{w}_{1i} \right) w_{1j} z_{ii} a_{ij} - \frac{1}{4} \sum_{i,j} q_{1j} w_{1i} z_{ij} z_{jj} d_i - \frac{1}{4} \sum_{i,j} q_{1j} w_{1i} z_{jj} a_{ji} \\
&\quad - \frac{1}{4} \sum_{i,j} w_{1i} w_{1j} z_{ij} d_i d_j - \frac{1}{4} \sum_{i,j} w_{1i} w_{1j} d_i a_{ij} - \frac{2}{4} \sum_{i,j} w_{1i} w_{1j} d_j a_{ji} - \frac{2}{4} \sum_{i,j} w_{1i} w_{1j} g_{ij} \\
&\quad - \sum_{i,j} w_{1i} q_{3j} z_{jj} a_{ji} + \frac{3}{4} \sum_{i,j} w_{1i} w_{1j} d_j a_{ji} + \frac{3}{4} \sum_{i,j} w_{1i} w_{1j} g_{ij} - \frac{1}{4} \sum_{i,j} w_{1j} q_{1i} z_{ii} a_{ij} \\
&\quad - \frac{1}{4} \sum_{i,j} w_{1j} w_{1i} g_{ij} \\
&= \sum_{i,j} \left\{ - (q_{1i} q_{3j} + \tilde{w}_{1i} \tilde{w}_{1j}) z_{ii} z_{ij} z_{jj} - \left(\tilde{w}_{1i} - \frac{1}{2} q_{1i} \right) w_{1j} z_{ij} z_{ii} d_j - \frac{1}{4} w_{1i} w_{1j} z_{ij} d_i d_j \right\}.
\end{aligned}$$

Assim, a parcela $\sum \lambda_{rstuvw}$ é dada por

$$\begin{aligned}
\sum \lambda_{rstuvw} &= \kappa^{rs} \kappa^{tu} \kappa^{vw} \left[\kappa_{rtv} (\kappa_{suw}/6 - \kappa_{sw}^{(u)}) + \kappa_{rtu} (\kappa_{svw}/4 - \kappa_{sw}^{(v)}) + \kappa_{rt}^{(v)} \kappa_{sw}^{(u)} + \kappa_{rt}^{(u)} \kappa_{sw}^{(v)} \right] \\
&= \sum_{i,j} \left\{ - (q_{1i} q_{2j} + \tilde{w}_{1i} \tilde{w}_{1j}) z_{ij}^3 - \left(\tilde{w}_{1i} - q_{1i} \right) w_{1j} z_{ij} c_{ij} + \frac{1}{2} w_{1i} w_{1j} z_{ij} b_{ij} \right. \\
&\quad \left. - (q_{1i} q_{3j} + \tilde{w}_{1i} \tilde{w}_{1j}) z_{ii} z_{ij} z_{jj} - \left(\tilde{w}_{1i} - \frac{1}{2} q_{1i} \right) w_{1j} z_{ij} z_{ii} d_j - \frac{1}{4} w_{1i} w_{1j} z_{ij} d_i d_j \right\}.
\end{aligned}$$

Em notação matricial, expressamos $\sum \lambda_{rstuvw}$ na forma

$$\begin{aligned}
\sum \lambda_{rstuvw} &= -\mathbf{1}^\top Q_1 Z^{(3)} Q_2 \iota - \iota^\top \tilde{W}_1 Z^{(3)} \tilde{W}_1 \iota - \iota^\top Q_1 Z_d Z Z_d Q_3 \iota - \iota^\top \tilde{W}_1 Z_d Z Z_d \tilde{W}_1 \iota \\
&\quad - \frac{1}{4} \iota^\top D W_1 Z \left[W_1 D + 4 Z_d \left(\tilde{W}_1 - \frac{1}{2} Q_1 \right) \right] \iota - tr \left\{ \left[\left(\tilde{W}_1 - Q_1 \right) C - \frac{1}{2} W_1 B \right] W_1 Z \right\}.
\end{aligned}$$

Apêndice B

Conjuntos de dados

Tabela B.1: Número de espécies de peixe em um lago (y) e o logaritmo da área do lago, em km^2 , (x).

y	x	y	x	y	x	y	x	y	x	y	x	y	x
10	2	10	4	68	8	18	8	11	9	24	6	48	7
37	4	14	0	93	10	214	10	48	10	12	10	21	5
60	5	39	5	13	7	177	11	14	3	26	10	46	7
113	10	14	1	53	8	17	11	28	9	13	6	14	7
99	11	14	4	17	8	50	10	17	1	19	6	7	5
13	0	67	11	245	10	5	10	17	11	19	7	5	2
30	4	36	4	88	8	22	7	21	5	22	4	40	9
114	11	30	0	24	3	156	13	13	8	15	4	18	9
112	10	19	2	37	9	74	13	14	5	9	3	20	6
17	2	46	9	22	8	13	5	21	9	23	5	17	6

Tabela B.2: Média dos erros cometidos pelos ratos.

Ratos isquêmicos (lesionados)											
Rato	Bl ₁	Bl ₂	Bl ₃	Bl ₄	Bl ₅	Rato	Bl ₁	Bl ₂	Bl ₃	Bl ₄	Bl ₅
1	0,5	1,2	0,3	0,5	0,3	14	3,2	2,6	2,7	0,8	0,3
2	0,5	1,0	0,6	0,0	0,0	15	2,2	1,9	0,3	0,0	0,0
3	0,3	0,5	0,2	0,6	0,1	16	1,9	0,3	0,4	0,2	0,2
4	0,3	0,4	0,0	0,0	0,0	17	0,9	1,8	1,1	1,0	0,6
5	0,8	0,1	0,0	0,0	0,1	18	1,8	1,9	0,8	0,2	0,6
6	1,3	1,0	0,5	0,0	0,6	19	3,2	2,6	2,7	0,8	0,3
7	0,8	0,1	0,2	0,0	0,0	20	2,2	1,9	0,3	0,0	0,0
8	2,3	0,4	0,0	0,0	0,0	21	1,9	0,3	0,4	0,2	0,2
9	0,6	0,8	0,0	0,0	0,0	22	0,9	1,8	1,1	1,0	0,6
10	0,4	0,0	0,0	0,0	0,0	23	1,8	1,9	0,8	0,2	0,6
11	0,1	0,2	0,6	0,0	0,0	24	0,8	1,0	0,0	0,3	0,0
12	0,3	0,2	0,8	0,0	0,0	25	0,2	0,5	0,2	0,2	0,1
13	0,2	0,0	0,9	0,1	0,0						
Ratos não-lesionados											
Rato	Bl ₁	Bl ₂	Bl ₃	Bl ₄	Bl ₅	Rato	Bl ₁	Bl ₂	Bl ₃	Bl ₄	Bl ₅
1	0,8	0,0	0,0	0,0	0,0	14	0,9	0,0	0,1	0,0	0,0
2	0,3	0,4	0,2	0,0	0,0	15	2,8	1,8	1,0	0,0	0,1
3	0,7	0,0	0,0	0,0	0,0	16	0,9	0,7	0,0	0,0	0,0
4	0,7	0,8	0,0	0,0	0,0	17	0,8	0,4	0,2	0,0	0,0
5	0,8	0,4	0,1	0,0	0,0	18	0,4	0,2	0,1	0,0	0,0
6	0,9	0,3	0,1	0,0	0,0	19	0,7	0,2	0,0	0,0	0,0
7	0,6	0,1	0,0	0,0	0,0	20	3,3	1,8	0,7	0,2	0,0
8	1,1	1,2	0,3	0,2	0,0	21	2,4	1,3	0,0	0,0	0,0
9	0,2	0,6	0,3	0,0	0,0	22	3,3	1,8	0,7	0,2	0,0
10	1,1	0,0	0,0	0,0	0,0	23	2,4	1,3	0,0	0,0	0,0
11	0,2	0,1	0,0	0,0	0,0	24	0,6	0,3	0,3	0,0	0,0
12	0,7	0,1	0,0	0,0	0,1	25	0,4	0,0	0,2	0,2	0,0
13	0,1	0,0	0,1	0,0	0,0	26	1,2	0,1	0,0	0,0	0,0

Referências Bibliográficas

- [1] Barbour, C.D., Brown, J.H., 1974. Fish species diversity in lakes. *The American Naturalist*, **108**, 473–489.
- [2] Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society A*, **160**, 268–282.
- [3] Bartlett, M. S. (1953). Aproximate confidence intervals II. *Biometrika*, **40**, 306–317.
- [4] Box, M. J. (1971). Bias in nonlinear estimation (with discussion). *Journal of the Royal Statistical Society B*, **33**, 171–201.
- [5] Cameron, A.C., Trivedi, P.K. (1998). Regression Analysis of Count Data. Cambridge University Press, New York, 434p.
- [6] Chesher, A., Smith, R. (1995). Bartlett corrections to likelihood ratio tests. *Biometrika*, **82**, 433–436.
- [7] Consul, P.C. (1990). New class of location-parameter discrete probability distributions and their characterizations. *Communications in Statistics, Theory and Methods*. **19**, 4653–4666.
- [8] Cook, R., Tsai, C. e Wei, B. (1986). Bias in nonlinear regression. *Biometrika*, **73**, 615–623.
- [9] Cordeiro, G.M. (1982). Improved likelihood ratio statiscs for generalized linear models. London. 227p. Tese (Doutorado)-Imperial Colege of Science and Technology-University of London.

- [10] Cordeiro, G.M. (1983). Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistical Society B*, **45**, 404–413.
- [11] Cordeiro, G.M. (1987). On the corrections to the likelihood ratio statistics. *Biometrika*, **74**, 265–274.
- [12] Cordeiro, G. M. (1999). Introdução à Teoria Assintótica. *IMPA Instituto de Matemática Pura e Aplicada*, Rio de Janeiro.
- [13] Cordeiro, G.M. (2004). Corrected likelihood ratio tests in symmetric nonlinear regression models. *Journal of Statistical Computation and Simulation*, **74**, 600–620.
- [14] Cordeiro, G. M., Andrade, M. G., de Castro, M. (2009). Power series generalized non-linear models. *Computational Statistics and Data Analysis*, **53**, 1155–1166.
- [15] Cordeiro, G. M., Barroso, L. P. (2007). A Third-order Bias Corrected Estimate in Generalized Linear Models. *Test*, **16**, 76–89.
- [16] Cordeiro, G.M., Cribari-Neto, F., Aubin, E.C.Q., Ferrari, S.L.P. (1995). Bartlett correction for one-parameter exponential family models. *Journal of Statistical Computation and Simulation*, **53**, 211–231.
- [17] Cordeiro, G. M., Cysneiros, A.H.M.A., Cysneiros, F.J.A. (2009). Bias-Corrected Maximum Likelihood Estimators in Nonlinear Heteroscedastic Models. *Communications in Statistics - Theory and Methods*, **87**, 1–26.
- [18] Cordeiro, G. M., Demétrio, C. G. B. (2008). Corrected estimators in extended quasi-likelihood methods. *Communications in Statistics. Theory and Methods*, **37**, 873–880.
- [19] Cordeiro, G. M., Ferrari, S. L. P., Uribe-Opazo, M. A., Vasconcellos, K. L. P. (2000). Corrected maximum-likelihood estimation in a class of symmetric nonlinear regression models. *Statistics and Probability Letters*, **46**, 317–328.
- [20] Cordeiro, G. M., McCullagh, P. (1991). Bias correction in generalized linear models. *Journal of the Royal Statistical Society B*, **53**, 629–643.

- [21] Cordeiro, G.M., Paula, G.A. (1989). Improved likelihood ratio statistics for exponential family nonlinear models. *Biometrika* 76, 93–100.
- [22] Cordeiro, G. M., Udo, M. C. T. (2008). Bias correction in generalized nonlinear models with dispersion covariates. *Communications in Statistics. Theory and Methods*, **37**, 2219–2225, 2008.
- [23] Cordeiro, G. M., Vasconcellos, K. L. P. (1997). Bias correction for a class of multivariate nonlinear regression models. *Statistics and Probability Letters*, **35**, 155–164.
- [24] Cordeiro, G. M. , Vasconcellos, K. L. P. (1999). Second–order biases of the maximum likelihood estimates in Von Mises regression models. *Australian and New Zealand Journal of Statistics*, **41**, 901–910.
- [25] Cox, D. R., Snell, E. (1968). A general definition of residuals. *Journal of the Royal Statistical Society B*, **30**, 248–275.
- [26] Cribari-Neto, F., Ferrari, S.L.P. (1995). Bartlett corrected tests for heteroskedastic linear models. *Economics Letters*, **48**, 113–118.
- [27] Cribari-Neto, F., Vasconcellos, K. L. P. (2002). Nearly unbiased maximum likelihood estimation for the beta distribution. *Journal of Statistical Computation and Simulation*, **72**, 107–118.
- [28] Cribari-Neto, F., Zarkos, S.G. (1995). Improved test statistics for multi- variate regression. *Economics Letters*, **49**, 113–120.
- [29] Cribari-Neto, F., Zarkos, S.G. (1999). R: yet another econometric programming environment. *Journal of Applied Econometrics*, **14**, 319–329.
- [30] Cribari-Neto, F., Zarkos, S.G. (2003). Econometric and statistical computing using Ox. *Computational Economics*, **21**, 277–295.
- [31] Cysneiros, A.H.M.A. (1997). Correção de Bartlett e tipo-Bartlett em modelos lineares generalizados, 80p., Dissertação (Mestrado em Estatística)–Instituto de Matemática e Estatística, USP, São Paulo.

- [32] Cysneiros, F.J.A., Cordeiro, G. M. e Cysneiros, A.H.M.A. (2009). Corrected maximum likelihood estimators in heteroscedastic symmetric nonlinear models . *Journal of Statistical Computation and Simulation* , **79**, 1–11.
- [33] Cysneiros, A.H.M.A., Ferrari, S.L.P. (2006). An improved likelihood ratio test for varying dispersion in exponential family nonlinear models. *Statistics & Probability Letters*, **76**, 255–265.
- [34] Doornik, J. A. (2001). *Ox: An Object-Oriented Matrix Language*. 4th ed. Timberlake Consultants Press, London; Oxford, <http://www.doornik.com>.
- [35] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, **7**, 1–26.
- [36] Ferrari, S.L.P., Arellano-Valle, R. B. (1996). Modified likelihood ratio and score tests in regression models using the t distribution. *Brazilian Journal of Probability and Statistics*, **10**, 15–33.
- [37] 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 Inferences*, **57**, 353–361.
- [38] Ferrari, S.L.P., Uribe-Opazo, M.A. (2001). Corrected likelihood ratio test in class of symmetric linear regression models. *Brazilian Journal of Probability and Statistics*, **15**, 49–67.
- [39] Frydenberg, M., Jensen, J. L. (1989). Is the “improved likelihood ratio statistic” really improved in the discrete case? *Biometrika*, **76**, 655–61.
- [40] Gupta, R.C. (1974). Modified power series distribution and some of its applications. *Sankhyā* **B36**, 288–298.
- [41] Haldane, J. B. S. (1953). The estimation of two parameters from a sample. *Sankhyā*, **12**, 313–320.
- [42] Haldane, J. B. S., Smith, S. M. (1956). The sampling distribution of a maximum likelihood estimate. *Biometrika*, **43**, 96–103.

- [43] Hayakawa, T. (1977). The likelihood ratio criterion and the asymptotic expansion of its distribution. *Annals of the Institute of Statistical Mathematics*, **29**, 359–378.
- [44] Ihaka, R., Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of Computational Graphics and Statistics*, **5**, 299–314.
- [45] Lawley, D. (1956). A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika*, **43**, 295–303.
- [46] Lemonte, A. J., Cribari Neto, F., Vasconcellos, K.L. (2007). Improved statistical inference for the two-parameter Birnbaum Saunders distribution. *Computational Statistics & Data Analysis*, **51**, 4656–4681.
- [47] Mittelbach, F., Goossens, M., Braams, J., Carlisle, D., Rowley, C. (2004). The LATEX Companion. Tools and Techniques for Computer Typesetting. *Addison Wesley*, Boston.
- [48] Montenegro, L.C.C., Cordeiro, G.M. (2002). Bartlett corrected likelihood ratio tests in location-scale nonlinear models. *Communications in Statistics, Theory and Methods*, **30**, 1353–1372.
- [49] Ospina, R., Cribari-Neto, F., Vasconcellos, K. L. P. (2006). Improved point and interval estimation for beta regression model. *Computational Statistics and Data Analysis*, **51**, 960–981.
- [50] Previdelli, I. T. S. (2005). Estimadores corrigidos para modelos não-lineares superdispersados. 144p. Tese (Doutorado em Engenharia de Produção) – Programa de Pós-Graduação em Engenharia de Produção, UFSC, Florianópolis - SC.
- [51] Rigby, R.A., Stasinopoulos, D.M., Akantziliotou, C. (2008). A framework for modelling overdispersed count data, including the Poisson-shifted generalized inverse Gaussian distribution. *Computational Statistics and Data Analysis*, **53**, 381–393.
- [52] Vasconcellos, K. L. P., Cordeiro, G. M. (2000). Bias corrected estimates in multivariate Student t regression models. *Communications in Statistics - Theory and Methods*, **29**, 797–822.

- [53] Vasconcellos, K. L. P., Silva, S. G. (2005). Corrected estimates for Student t regression models with unknown degrees of freedom. *Journal of Statistical Computation and Simulation*, **75**, 409–423.
- [54] Venables, W. N, Ripley, B. D. (2002). Modern Applied Statistics with S-Plus. *Springer Erlag*, New York.
- [55] Young, D., Bakir, S. (1987). Bias correction for a generalized log-gamma regression model. *Technometrics*, **29**, 183–191.