APERFEIÇOAMENTO DE TESTES NOS MODELOS SÉRIES DE POTÊNCIA NÃO-LINEARES GENERALIZADOS

MARIA DO CARMO SOARES DE LIMA

Orientadora: Profª Drª 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 2012

Catalogação na fonte
Bibliotecária Jane Souto Maior, CRB4-571

Lima, Maria do Carmo Soares de
Aperfeiçoamento de testes nos modelos séries de potência não-lineares generalizados / Maria do Carmo Soares de Lima - Recife: O Autor, 2012.
viii, 57 folhas: tab.

Orientador: Audrey Helen Mariz de Aquino Cysneiros.
Dissertação (mestrado) - Universidade Federal de Pernambuco. CCEN, Estatística, 2012.

Inclui bibliografia e apêndice.

1. Estatística Matemática. 2. Teoria assintótica. I. Cysneiros, Audrey Helen Mariz de Aquino (orientador). II. Título.

519.9 CDD (23. ed.) MEI2012 – 027

Universidade Federal de Pernambuco
Pós-Graduação em Estatística

27 de fevereiro de 2012

Nós recomendamos que a dissertação de mestrado de autoria de

Maria do Carmo Soares de Lima

intitulada

"Aperfeiçoamento de testes nos modelos 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:

Audrey Helen Mariz de Aquino Cysneiros orientadora

Betsabé Grimalda Blas Achic

Michelli Karinne Barros da Silva (UFCG)

Este documento será anexado à versão final da dissertação.

Dedico esse trabalho àqueles que estiveram comigo nessa longa caminhada, feita de dias cheios de sucesso e também de fracassos. À minha mãe Edlúcia, meu irmão Douglas, minha avó Iracema, meu companheiro Flávio e ao meu anjinho Letícia. Amo vocês! “Construí amigos, enfrentei derrotas, venci obstáculos, bati na porta da vida e disse-lhe: Não tenho medo de vivê-la.” (Augusto Cury)

Agradecimentos

Primeiramente a Deus, que me guiou por esse longo caminho até mais uma conquista. Espero que ele esteja presente em muitas outras que virão. Ao longo desses anos de estudo, meu foco sempre foi a área de Matemática e, por algum motivo que só Ele deve saber, acabei conhecendo essa área maravilhosa que é a Estatística, em que posso aplicar meus conhecimentos matemáticos e obter novos, antes temidos por mim, por assim dizer.

À minha orientadora Audrey Helen Mariz de Aquino Cysneiros, pelos conselhos e orientações fornecidas ao longo do trabalho. Foram muitas horas de trabalho árduo que me tomaram o sono e me leveram a muitos questionamentos. Apesar dos problemas ao longo desse ano, consegui com apoio da professora elevar meus níveis de conhecimento e também ter afeição. Por isso, muito obrigada pelos momentos inesquecíveis de alegrias, risos, choro e também desespero, que me fizeram crescer como pessoa.

À minha família, em especial à minha mãe Edlúcia, meu irmão Douglas e minha vó Iracema, pelo apoio e credibilidade que depositaram em mim. Agradeço do fundo do coração pela educação que minha mãe me proporcionou, sempre com muita garra, como a mulher guerreira que ela é. Devo tudo o que sou hoje a essa pessoa admirável que me amou desde o momento que soube que eu estava dentro dela, sem nem mesmo me conhecer. Obrigada também ao meu irmão, que sempre esteve comigo, mesmo que em pensamento, sempre torcendo pela minha vitória. Quanto à minha avó, digo que ela é minha segunda mãe, apoiando-me em tudo que faço. Assim, agradeço a todos da minha família pelo apoio e credibilidade no meu trabalho.

Ao meu companheiro, Flávio, que me deu sempre muito apoio e incentivo nas difíceis e intermináveis horas no computador. Além de companheiro, posso dizer que foi como um pai que nunca tive. Portanto, obrigada pelos sermões e abraços quando precisei. Talvez não tenha sido tão precisa como deveria, mas agora o faço: assim como sei que você me admira, eu também o admiro demais como companheiro, pai e pessoa.

A todos os professores e funcionários do Departamento de Estatística da UFPE pelo apoio e confiança em mim, em especial à Valéria, que sempre me auxiliou em tudo. Agora entendo porque me disseram a seguinte frase quando entrei em Estatística: "Qualquer problema, procure Valéria".

À professora Cristina por ter me mostrado como essa área de estudo pode ser tão encantadora e a todos os outros professores que fizeram parte dessa história, inclusive aos professores do Departamento de Matemática (minha origem).

Aos estudantes de Estatística que tive o prazer de conhecer, principalmente à Priscila que me ajudou em vários programas durante a realização dessa dissertação.

Ao CNPq pelo apoio financeiro oferecido.

Resumo

Essa dissertação tem dois objetivos. O primeiro consiste na obtenção de um fator de correção tipo-Bartlett para a estatística escore nos modelos em série de potência não-lineares generalizados. O segundo é comparar o desempenho dos testes baseados nas seguintes estatísticas, a saber: escore e suas versões corrigidas, razão de verossimilhanças e suas versões corrigidas e a gradiente, no que tange ao tamanho e ao poder. Finalmente apresentamos uma aplicação a dados reais.

Palavras-chave: Correção tipo-Bartlett, modelos em série de potência não-lineares generalizados, estatística escore, estatística da razão de verossimilhanças, estatística gradiente.

Abstract

This dissertation has two purposes. The first involves obtaining a correction factor for the Bartlett-type statistical scoring models in power series nonlinear generalized. The second is to compare the performance of the test score and their corrected versions tests with the gradient of the likelihood ratio and its corrected versions with respect to the size and power. Finally we present an application to real data.

Keywords: Bartlett-type correction, models in power series nonlinear generalized, test score, likelihood ratio test, gradient test.

Índice

Página

Lista de Tabelas

Página

Capítulo 1

Introdução

Dados na forma de contagens ocorrem nas mais diferentes áreas do conhecimento. Os modelos de Poisson e binomial negativo são amplamente utilizados para analisar dados desse tipo, sendo comum encontrar situações em que a variância da variável resposta é maior do que ou menor do que a média, sendo essas denominadas de sobredispersão e subdispersão, respectivamente. Estes fenômenos são bem conhecidos na literatura estatística e podem ser provocados por uma ou mais de uma causa. Existem estudos sobre o efeito da sobredispersão quanto à inferência realizada a partir de um modelo de Poisson e modelos alternativos foram propostos com o objetivo de acomodar a sobredispersão.

Cordeiro et al. (2009) propuseram uma nova classe de modelos em séries de potências para representar a variável resposta, adotando como componente sistemático uma função de ligação não-linear entre a média da variável resposta e a estrutura não-linear do modelo. Esta classe de modelos é descrita pela sigla MSPNLGs (modelos em séries de potência não-lineares generalizados). Desta forma, engloba-se vários modelos discretos importantes em uma única estrutura conceitual. Esta família de distribuições discretas tem uma estrutura bastante flexível para a modelagem de dados discretos. Além disso, esta classe de modelos engloba modelos tradicionais tais como os modelos log-lineares, binomial e binomial negativo.

Silva (2010) derivou expressões em formas matriciais para o fator de correção de Bartlett à estatística da razão de verossimilhanças nos modelos MSPNLG, supondo o parâmetro de dispersão conhecido. Além disso, obteve a correção do viés de segunda ordem, via Cox & Snell (1968), dos estimadores de máxima verossimilhança dos parâmetros nessa classe de modelos.

Dando continuidade aos trabalhos recentemente desenvolvidos a respeito do tema, a principal contribuição dessa dissertação é a obtenção de um fator de correção tipo-Bartlett para a estatística escore no modelo em questão. Essa correção visa melhorar a qualidade de aproximação da estatística escore pela distribuição qui-quadrado, dada por um polinômio na própria estatística escore. Desta forma, obtemos uma estatística escore modificada tendo distribuição qui-quadrado até ordem $O(n^{-1})$, sob a hipótese nula. Outra contribuição dada nesse trabalho, foi avaliar numericamente o desempenho dos testes baseados nas estatísticas, a saber: razão de verossimilhanças e escore, bem como nas versões corrigidas via Bartlett e tipo-Bartlett, respectivamente e a gradiente.

No Capítulo 2, procuramos revisar os principais resultados teóricos relacionados com os MSPNLGs. Em particular, mostramos como ficam os testes baseados nas estatísticas da razão de verossimilhanças, escore e gradiente nessa classe de modelos.

No Capítulo 3, revisamos a obtenção do fator de correção de Bartlett para a estatística da razão de verossimilhanças e apresentamos a obtenção do fator de correção tipo-Bartlett para a estatística escore nos MSPNLGs. As fórmulas das correções são dadas em notação matricial e podem ser implementadas em sistema de computação algébrica. Vale salientar aqui que o Capítulo 3 será desenvolvido de maneira não-relacionada com os outros.

No Capítulo 4, apresentamos resultados de simulação para avaliar o efeito da correção para testes em MSPNLGs. Além disso, resultados numéricos sobre o comportamento em amostras finitas de diferentes testes nos MSPNLGs são apresentados em relação ao tamanho e poder. Por fim, uma aplicação a dados reais é apresentada e considerações são feitas no final do capítulo.

Todos os procedimentos inferenciais derivados dos resultados teóricos obtidos na dissertação são avaliados e comparados através de simulações Monte Carlo as quais foram desenvolvidas a partir de programas construídos usando a linguagem de programação matricial Ox (Doornik, 2006). As maximizações não-lineares necessárias para o cálculo das estimativas de máxima verossimilhança foram feitas utilizando o algoritmo quasi-Newton BFGS (Nocedal e Wright, 1999, capítulo 8), disponíveis em funções pré-definidas da linguagem Ox.

Capítulo 2

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

2.1 Introdução

Cordeiro et al. (2009) propuseram uma nova classe de modelos em séries de potências não-lineares generalizados (MSPNLGs) que são definidos por um conjunto de variáveis aleatórias independentes pertencentes à família de distribuições em séries de potência e admitem que uma função monótona da média da variável resposta seja definida por um preditor não-linear envolvendo regressores e parâmetros desconhecidos. Nesta classe de modelos, abordamos a situação em que o parâmetro de dispersão é conhecido.

Neste capítulo, apresentamos a definição dos modelos em séries de potência não-lineares generalizados e algumas propriedades. Além disso, apresentamos aspectos inferenciais tais como, a estimação dos parâmetros de regressão e mostramos como ficam os testes baseados nas estatísticas da razão de verossimilhanças, escore e gradiente.

2.2 Definição e propriedades

Consideremos n variáveis aleatórias discretas independentes $Y_1, \ldots, Y_n$ tais que Y_i pertence à uma família de distribuições com parâmetros de média $\mu_i > 0$ e parâmetro de dispersão $\phi > 0$ (que assumimos ser conhecido), cada uma com função de probabilidade na forma

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

em que o suporte de Y_i é um subconjunto A_ϵ dos inteiros $\{\epsilon, \epsilon + 1, \ldots\}$, $\epsilon \geq 0$ e não depende de parâmetros desconhecidos. As funções $a(y, \phi)$,$g(\mu_i, \phi)$ e $f(\mu_i, \phi)$ são positivas, sendo as duas últimas funções analíticas dos parâmetros μ_i e ϕ, finitas e duas vezes diferenciáveis, em que $f(\mu, \phi)$ é tal que

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

Desde que seja satisfeita a suposição de que o parâmetro $\phi \geq 0$, que assumimos ser conhecido, a variável aleatória Y tem uma distribuição de probabilidade completamente determinada por sua função de variância.

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

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

em que $f = f(\mu_i, \phi)$, $g = g(\mu_i, \phi)$ e o símbolo ("$\prime$") significa que a diferenciação é realizada em relação a μ. Note que a função de variância, dada na segunda equação em (2.2) depende apenas da função $g(\mu, \phi)$ e de sua primeira derivada. Observe também que há uma relação entre a média de Y_i e o componente sistemático, feita através de uma função de ligação da forma

$$h(\mu_i) = \eta_i = \eta(x_i; \boldsymbol{\beta}), \quad i = 1, \ldots, n. \tag{2.3}$$

Em (2.3), $h(\cdot)$ representa uma função de ligação conhecida e duas vezes diferenciável tal que $\boldsymbol{\beta = (\beta_1, \ldots, \beta_p)^T}$, $(p < n)$, é um vetor de parâmetros desconhecidos a serem estimados, $\boldsymbol{x_i = (x_{i1}, \ldots, x_{ik})^T}$ são 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 $\boldsymbol{\beta}$. Além disso, $\tilde{X} = \tilde{X}(\boldsymbol{\beta}) = \partial\eta/\partial\boldsymbol{\beta}^T$ é a matriz de derivadas, com $\boldsymbol{\eta = (\eta_1, \ldots, \eta_n)^T}$, tal que o posto é p para todo $\boldsymbol{\beta}$, com elementos que são geralmente funções do vetor de parâmetros $\boldsymbol{\beta}$ desconhecidos.

A família (2.1) contém diversas distribuições conhecidas, tais como: Poisson, Binomial, Binomial Negativa, Poisson Generalizada, Binomial Negativa Generalizada, Borel, Consul, Borel-Tanner, Geeta-m e Haight. Algumas características de tais distribuições podem ser encontradas na Tabela 2.1.

O logaritmo da função de verossimilhança dos parâmetros do modelo, considerando um MSPNLG definido por (2.1) e (2.3), dado o vetor de observações $\boldsymbol{(y_1, \ldots, y_n)^T}$, pode ser expresso na forma:

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

Assumimos que $\ell(\boldsymbol{\beta}; \boldsymbol{y})$ seja regular com respeito às derivadas em relação aos componentes de β até quarta ordem. Detalhes sobre as condições de regularidade podem ser encontradas em Cox & Hinkley (1974, Capítulo 9).

2.3 Estimação dos parâmetros de regressão

A função escore para o vetor de parâmetros β é expressa por:

$$U_r = \frac{\partial \ell(\boldsymbol{\beta}; \boldsymbol{y})}{\partial \boldsymbol{\beta}_r} = \sum_{i=1}^{n} \left[y_i \frac{g_i^{'}}{g_i} - \frac{f_i^{'}}{f_i}\right] \frac{1}{h_i^{'}} \tilde{x}_{ir}.$$

em que $r = 1, \ldots, p$.

Matricialmente, a função escore fica dada da seguinte forma:

$$U_{\boldsymbol{\beta}} = \frac{\partial \ell(\boldsymbol{\beta}; \boldsymbol{y})}{\partial \boldsymbol{\beta}} = \tilde{X}^T (T\boldsymbol{y} - \boldsymbol{Q}),$$

em que T =diag$\{t_i\}$ $(i = 1, \ldots, n)$ é uma matriz diagonal $n \times n$ com elementos definidos por $t_i = \frac{g_i^{'}}{g_i h_i^{'}}$ e $\boldsymbol{Q = (q_1, \ldots, q_n)}$ é um vetor $n \times 1$ com o i-ésimo elemento dado por $q_i = \frac{f_i^{'}}{f_i h_i^{'}}$. As funções t e q para algumas distribuições podem ser encontradas na Tabela 2.2.

A matriz de informação para β é expressa por:

$$K_{\boldsymbol{\beta}} = E\left\{-\frac{\partial^2 \ell(\boldsymbol{\beta}; \boldsymbol{y})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T}\right\} = \tilde{X}^T W \tilde{X}, \tag{2.5}$$

em que W =diag$\{w_i\}$ $(i = 1, \ldots, n)$ é uma matriz diagonal $n \times n$ de pesos definidos por:

$$w_i = \left(q_i^{'} - \frac{f_i^{'} g_i}{f_i g_i^{'}} t_i^{'}\right) \frac{1}{h_i^{'}}.$$

Quanto à inferência baseada no método de razão de verossimilhanças sobre o parâmetro β, esta pode ser realizada maximizando numericamente o logaritmo da função de verossimilhança do modelo. Alternativamente, poderíamos utilizar o processo iterativo de Newton-Raphson com o intuito de obter a estimativa de $\boldsymbol{\beta}$. O processo iterativo escoring de Fisher é definido como

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

.

Vale ressaltar que t_i e q_i podem ser reescritos como $t_i = (V_i h_i')^{-1}$ e $q_i = \mu_i (V_i h_i')^{-1}$. Assim, a matriz T fica dada por $(VL)^{-1}$ e o vetor Q por $(VL)^{-1}\mu$, em que $\mu = (\mu_1, \ldots, \mu_n)^\top$ é um vetor $n \times 1$. Esse processo iterativo pode ser reescrito como um processo de mínimos quadrados reponderados, mais detalhes ver Silva (2010).

2.3.1 Testes baseados nas estatísticas da razão de verossimilhanças, escore e gradiente em MSPNLGs

Consideremos o vetor $\boldsymbol{Y} = (\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n})^{\boldsymbol{T}}$ representando n observações independentes cujo logaritmo da função de verossimilhança $l(\boldsymbol{\beta}; \boldsymbol{y})$, dado por (2.4), depende do parâmetro desconhecido $\boldsymbol{\beta} = (\boldsymbol{\beta_1}, \ldots, \boldsymbol{\beta_p})^{\boldsymbol{T}}$.

Assuma que o vetor de parâmetros $\boldsymbol{\beta}$ pode ser decomposto como $\boldsymbol{\beta} = (\boldsymbol{\beta_1^T}, \boldsymbol{\beta_2^T})^{\boldsymbol{T}}$, sendo $\boldsymbol{\beta_1} = (\boldsymbol{\beta_1}, \ldots, \boldsymbol{\beta_q})^{\boldsymbol{T}}$ o vetor de parâmetros de interesse e $\boldsymbol{\beta_2} = (\boldsymbol{\beta_{q+1}}, \ldots, \boldsymbol{\beta_p})^{\boldsymbol{T}}$ o vetor de parâmetros de perturbação. Da mesma forma, consideremos a decomposição da matriz modelo $X = (X_1, X_2)$ e da função escore $U(\boldsymbol{\beta}) = (U_1(\boldsymbol{\beta})^T, U_2(\boldsymbol{\beta})^T)^T$. Sejam os estimadores de máxima verossimilhança $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_{\boldsymbol{1}}^{\boldsymbol{T}}, \hat{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{T}})^{\boldsymbol{T}}$ e $\tilde{\boldsymbol{\beta}} = (\boldsymbol{\beta}_{\boldsymbol{1}}^{\boldsymbol{(0)T}}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{2}}^{\boldsymbol{T}})^{\boldsymbol{T}}$ os estimadores de máxima verossimilhança irrestrito e restrito de $\boldsymbol{\beta}$, respectivamente. A decomposição de $\boldsymbol{\beta}$ induz às seguintes partições da matriz de informação de Fisher, dada na sua forma matricial em (2.5), e sua inversa:

$$K(\boldsymbol{\beta}) = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \quad e \quad K(\boldsymbol{\beta})^{-1} = \begin{bmatrix} K^{11} & K^{12} \\ K^{21} & K^{22} \end{bmatrix},$$

sendo $K_{11} = X_1^T W X_1$, $K_{22} = X_2^T W X_2$, $K_{12} = K_{21}^T = X_1^T W X_2$, $K^{11} = (R^T W R)^{-1}$, $K^{22} = (X_2^T W X_2)^{-1} + C(R^T W R)^{-1} C^T$ (covariância assintótica de $\hat{\beta}_2$) e $K^{12} = K^{21} = (R^T W R)^{-1} C^T$, em que $R = X_1 - X_2 C$ e $C = (X_2^T W X_2)^{-1} X_2^T W X_1$ é uma matriz $n \times q$, sendo suas colunas vetores dos coeficientes de regressão linear das colunas de X_1 sobre X_2, tal que o peso é W.

Em muitas situações, há interesse em testar hipótese sobre uma parte do vetor de parâmetros $\boldsymbol{\beta}$, digamos $H_0 : \boldsymbol{\beta_1} = \boldsymbol{\beta_1^{(0)}}$ contra $H_1 : \boldsymbol{\beta_1} \neq \boldsymbol{\beta_1^{(0)}}$, em que $\boldsymbol{\beta_1^{(0)}}$ é um vetor especificado de dimensão q ($q \leq p$).

A estatística da razão de verossimilhanças para o teste de H_0 é definida como

$$LR = 2[l(\boldsymbol{\hat{\beta}_1}; \boldsymbol{\hat{\beta}_2}) - l(\boldsymbol{\beta_1^{(0)}}; \boldsymbol{\tilde{\beta}_2})], \tag{2.6}$$

em que o estimador de máxima verossimilhança irrestrito de $\boldsymbol{\beta}$ é $\boldsymbol{\hat{\beta} = (\hat{\beta}_1^T, \hat{\beta}_2^T)}$ e $\boldsymbol{\tilde{\beta} = (\beta_1^{(0)T}, \tilde{\beta}_2^T)^T}$ é o estimador de máxima verossimilhança restrito sob a hipótese nula, sendo $\boldsymbol{\tilde{\beta}_2}$ o estimador de máxima verossimilhança restrito de $\boldsymbol{\beta_2}$ sob a hipótese nula. A estatística da razão de verossimilhanças tem, sob a hipótese nula, distribuição assintótica χ^2_q, em que q é o número de restrições impostas sob H_0. Rejeitamos a hipótese nula, ao nível de significância α, se $LR > \chi^2_{(\alpha;q)}$, em que $\chi^2_{(\alpha;q)}$ é o percentil $(1-\alpha)$ da distribuição χ^2_q.

A estatística escore (Rao, 1948) para testar H_0 contra H_1 é dada por:

$$S_R = U(\tilde{\boldsymbol{\beta}})^T K(\tilde{\boldsymbol{\beta}})^{-1} U(\tilde{\boldsymbol{\beta}}),$$

em que $U(\boldsymbol{\beta})$ é a função escore, $K(\boldsymbol{\beta})$ é a matriz de informação de Fisher e $\boldsymbol{\tilde{\beta} = (\beta_1^{(0)}, \tilde{\beta}_2)^T}$ sendo $\boldsymbol{\tilde{\beta}_2}$ é o estimador de máxima verossimilhança restrito de $\boldsymbol{\beta_2}$ e $\boldsymbol{\beta_1^{(0)}}$ um vetor especificado de dimensão q. A estatística escore tem, sob a hipótese nula, distribuição assintótica χ^2_q. Rejeitamos a hipótese nula, ao nível de significância α, se $S_R > \chi^2_{(\alpha;q)}$.

A estatística gradiente (Terrel, 2002) para testar H_0 contra H_1 é definida como:

$$S_T = U_1(\tilde{\boldsymbol{\beta}})^T (\boldsymbol{\widehat{\beta}_1} - \boldsymbol{\beta_1^{(0)}}), \tag{2.7}$$

em que $\hat{\boldsymbol{\beta}}$ é o estimador de máxima verossimilhança irrestrito de $\boldsymbol{\beta}$. Vale salientar que a estatística gradiente definida em (2.7) é uma espécie de combinação das estatísticas escore e Wald modificada (ver Hayakawa e Puri, 1985). A estatística Wald modificada, denotada aqui por W_1, para testar a hipótese nula composta H_0:$\boldsymbol{\beta_1} = \boldsymbol{\beta_1^{(0)}}$ contra a hipótese alternativa bilateral H_1:$\boldsymbol{\beta_1} \neq \boldsymbol{\beta_1^{(0)}}$ é definida como:

$$W_1 = (\widehat{\boldsymbol{\beta}_1} - \boldsymbol{\beta_1^{(0)}})^T K^{11}(\tilde{\boldsymbol{\beta}})^{-1}(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta_1^{(0)}})$$

em que K^{11} é a matriz de covariância assintótica de $\hat{\boldsymbol{\beta}}_{\mathbf{1}}$ obtida de $K(\boldsymbol{\beta})^{-1}$.

Em problemas regulares, sabe-se que as estatísticas da razão de verossimilhanças, escore e Wald tem assintoticamente e, sob a hipótese nula, distribuição χ^2_q . Terrel (2002) mostra que a estatística gradiente tem a mesma distribuição assintótica que as estatísticas escore e Wald. Portanto, a estatística gradiente tem, sob a hipótese nula, distribuição assintótica χ^2_q. Rejeitamos a hipótese nula, ao nível de significância α, se $S_T > \chi^2_{(\alpha;q)}$. Note que nesta dissertação, não daremos ênfase ao estudo da estatística Wald. Lemonte (2009) comparou o poder local do teste gradiente com o poder local dos testes baseados nas estatísticas da razão de verossimilhanças, Wald e escore no modelo Birnbaum-Saunders.

Sabe-se que o cálculo da estatística da razão de verossimilhanças requer estimação dos parâmetros sob ambas as hipóteses (nula e alternativa), enquanto que o cálculo da estatística escore envolve a estimação dos parâmetros apenas sob a hipótese nula. Nos casos em que a estimação sob a hipótese alternativa é complicada, a estatística escore pode apresentar uma vantagem computacional em relação a estatística da razão de verossimilhanças. Uma vantagem computacional da estatística gradiente em relação às demais estatísticas é quando o cálculo da matriz de informação de Fisher for muito complicado.

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	$\frac{1}{y!}$	$\{0,1,2,\ldots\}$
2. Binomial	$\left(1+\frac{\mu}{m-\mu}\right)^{m}$	$\frac{\mu}{m-\mu}$	$\binom{m}{y}$	$\{0,1,2,\ldots,m\}$
3. Binomial negativa	$\left(1-\frac{\mu}{\mu+\phi}\right)^{-\phi}$	$\frac{\mu}{\mu+\phi}$	$\frac{\Gamma(\phi+y)}{y!\Gamma(\phi)}$	$\{0,1,2,\ldots\}$
4. Poisson generalizada	$e^{\mu(1+\mu\phi)^{-1}}$	$\frac{\mu e^{-\mu\phi(1+\mu\phi)^{-1}}}{1+\mu\phi}$	$\frac{(1+\phi y)^{y-1}}{y!}$	$\{0,1,2,\ldots\}$
5. Borel	$1-\frac{1}{\mu}$	$\left(1-\frac{1}{\mu}\right)e^{-1+1/\mu}$	$\frac{y^{y-2}}{(y-1)!}$	$\{1,2,\ldots\}$
6. Consul	$\frac{\mu-1}{\mu(\phi-1)+1}$	$\phi^{-\phi}(1-\mu^{-1})(\phi-1+\mu^{-1})^{\phi-1}$	$\frac{\Gamma(\phi y+1)}{y!\Gamma(\phi y-y+2)}$	$\{1,2,\ldots\}$
7. Binomial negativa generalizada	$\left(\frac{\phi-1+\nu/\mu}{\phi+\nu/\mu}\right)^{-\nu}$	$\frac{1}{\phi+\nu/\mu}\left(\frac{\phi-1+\nu/\mu}{\phi+\nu/\mu}\right)^{\phi-1}$	$\frac{\nu\Gamma(\phi y+\nu+1)}{(\phi y+\nu)y!\Gamma(\phi y-y+\nu+1)}$	$\{0,1,2,\ldots\}$
8. Borel–Tanner	$\left(1-\frac{m}{\mu}\right)^{m}$	$\left(1-\frac{m}{\mu}\right)e^{-1+m/\mu}$	$\frac{my^{y-m-1}}{(y-m)!}$	$\{m,m+1,\ldots\}$
9. Delta binomial	$\left\{\frac{\mu-m}{\mu(\phi-1)+m}\right\}^{m}$	$\frac{1}{\phi^{\phi}}\left(1-\frac{m}{\mu}\right)\left(\phi-1+\frac{m}{\mu}\right)^{\phi-1}$	$\frac{m\Gamma(\phi y+1)}{y(y-m)!\Gamma(\phi y-y+m+1)}$	$\{m,m+1,\ldots\}$
10. Geeta	$\frac{\mu-1}{\phi\mu-1}$	$\frac{\mu-1}{\phi\mu-1}\left\{\frac{(\phi-1)\mu}{\phi\mu-1}\right\}^{\phi-1}$	$\frac{\Gamma(\phi y-1)}{y!\Gamma(\phi y-y)}$	$\{1,2,\ldots\}$
11. Geeta-m	$\left(\frac{\mu-m}{\phi\mu-m}\right)^{m}$	$\frac{\mu-m}{\phi\mu-m}\left\{\frac{(\phi-1)\mu}{\phi\mu-m}\right\}^{\phi-1}$	$\frac{m\Gamma(\phi y-m)}{y(y-m)!\Gamma(\phi y-y)}$	$\{m,m+1,\ldots\}$
12. Haigth	$\frac{\mu-1}{2\,\mu-1}$	$\frac{\mu\,(\mu-1)}{(2\mu-1)^2}$	$\frac{(2y-2)!}{y!(y-1)!}$	$\{1,2,\ldots\}$

Tabela 2.2: Funções t e q para algumas distribuições da família (2.1).

Distribution	t	q
1. Poisson	$\frac{1}{\mu h'}$	$\frac{1}{h'}$
2. binomial	$\frac{m}{\mu(m-\mu)h'}$	$\frac{m}{(m-\mu)h'}$
3. generalized Poisson	$\frac{1}{\mu(1+\phi\mu)^2h'}$	$\frac{1}{(1+\phi\mu)^2h'}$
4. Consul	$\frac{\phi}{\mu(\mu-1)\left\{\mu(\phi-1)+1\right\}h'}$	$\frac{\phi}{(\mu-1)\left\{\mu(\phi-1)+1\right\}h'}$
5. generalized negative binomial	$\frac{\nu^2}{\mu(\nu+\mu\phi)\left\{\nu+\mu(\phi-1)\right\}h'}$	$\frac{\nu^2}{(\nu+\mu\phi)\left\{\nu+\mu(\phi-1)\right\}h'}$
6. delta binomial	$\frac{m^2\phi}{\mu\left\{\mu(\phi-1)+m\right\}(\mu-m)h'}$	$\frac{m^2\phi}{\left\{\mu(\phi-1)+m\right\}(\mu-m)h'}$

Capítulo 3

Correção de Bartlett e tipo-Bartlett

3.1 Introdução

Sabe-se que, em problemas regulares, as estatísticas da razão de verossimilhanças e escore são assintoticamente equivalentes, uma vez que possuem a mesma distribuição de referência qui-quadrado (χ^2), sob H_0. Tendo em vista a dificuldade em se determinar as distribuições exatas das estatísticas razão de verossimilhanças (LR) e escore (S_R), os testes realizados têm sido construídos baseados em resultados assintóticos. Entretanto, em amostras de tamanho pequeno ou até mesmo moderado, a aproximação dessas estatísticas pela distribuição χ^2 pode não ser aceitável.

Assim, precisava-se de uma solução que viesse a melhorar a qualidade das aproximações das distribuições das estatísticas razão de verossimilhanças e escore pela distribuição χ^2 de referência. Bartlett (1937) propôs a modificação da estatística de razão de verossimilhanças, tendo como objetivo gerar uma estatística modificada, com o primeiro momento equivalente ao da distribuição χ^2 de referência. Bartlett (1937) mostrou que, sob a hipótese nula e em problemas regulares, a média da estatística da razão de verossimilhanças pode ser expandida de tal forma que seu cálculo envolve uma constante d de ordem $O(n^{-1})$ em que n o tamanho da amostra. Dessa forma a média da estatística corrigida, $LR^* = LR/(1 + d)$, em que LR é dada em (2.6), se torna mais próxima da média da distribuição χ^2_q, quando comparada a estatística usual. Na fórmula de LR^*, o fator $1/(1 + d)$ é conhecido como fator de correção de Bartlett.

Hayakawa (1977) derivou uma expansão assintótica para a distribuição de razão de verossimilhanças sob a hipótese nula, mostrando que, se a hipótese nula for simples, a estatística da razão de verossimilhanças modificada tem distribuição χ^2 até ordem $O(n^{-1})$.

Cordeiro (1982) obteve fatores de correção para a estatística da razão de verossimilhanças em modelos multinomiais, mostrando que as taxas de rejeição dos testes baseados na estatística corrigida estão mais próximas dos seus respectivos níveis nominais que as taxas de rejeição do teste baseado na estatística não corrigida.

Cordeiro (1983 e 1987) obteve o fator de correção de Bartlett para estatística de razão de verossimilhanças nos MLGs considerando o fator de escala conhecido e desconhecido, respectivamente, e mostrou que a correção representou um considerável aperfeiçoamento. Outras referências sobre trabalhos que tratam de correção de Bartlett em outros modelos podem ser encontradas em Cribari-Neto e Cordeiro (1997) e Cordeiro (1999).

Ferrari e Uribe - Opazo (2001) obtiveram um fator de correção de Bartlett na classe dos modelos de regressão lineares simétricos.

Cysneiros e Ferrari (2006) obtiveram refinamentos de um teste de heteroscedasticidade baseado em verossimilhança perfilada nos modelos não-lineares da família exponencial.

Cordeiro e Ferrari (1991) mostraram que uma estatística escore aperfeiçoada seguindo distribuição χ^2 até ordem $O(n^{-1})$, sob a hipótese nula, pode ser sempre encontrada e envolve um polinômio de segundo grau na própria estatística escore. As quantidades a, b e c possuem fórmulas fechadas e envolvem funções de cumulantes conjuntos de derivadas do logaritmo da função de verossimilhança e o número de restrições impostas sob a hipótese nula. As fórmulas de tais quantidades serão mostradas ainda nesse capítulo.

Cordeiro et al. (1998) apresentaram uma fórmula para a estatística escore aperfeiçoada, sendo essa uma transformação monótona da estatística escore. Ainda nesse capítulo, falaremos com mais detalhes sobre o fator de correção tipo-Bartlett.

Braga (2007) apresentou fatores de correção tipo-Bartlett em modelos de regressão não-lineares simétricos visando melhorar a estatística escore.

Brito (2009) obteve a expressão do fator de correção Bartlett para a estatística de razão de verossimilhanças nos modelos não-lineares simétricos heteroscedásticos com parâmetro de escala desconhecido. Lemonte (2009) obteve um fator de correção de Bartlett para a estatística da razão de verossimilhanças no modelo de regressão Birnbaum Saunders.

Silva (2010) obteve um fator de correção de Bartlett à estatística da razão de verossimilhanças na classe de modelos em séries de potência não-lineares generalizados. Nascimento (2010) desenvolveu um fator de correção tipo-Bartlett para a estatística escore nos modelos não-lineares simétricos heteroscedásticos, usando quaisquer funções de ligação para a média e para o parâmetro de dispersão.

Uribe - Opazo (1997) obteve fatores de correção de Bartlett e tipo-Bartlett para as estatísticas da razão de verossimilhanças e escore nos modelos lineares simétricos homoscedásticos, respectivamente.

Cavalcanti (2009) desenvolveu fatores de correção tipo-Bartlett para o teste escore em modelos não-lineares da família exponencial considerando covariáveis para modelar o parâmetro de dispersão.

Lemonte e Ferrari (2009) obtiveram um fator de correção tipo-Bartlett para a estatística escore no modelo de regressão Birnbaum Saunders.

Vale ressaltar que não é possível garantir que, para variáveis aleatórias discretas, a estatística da razão de verossimilhança corrigida pelo fator de correção de Bartlett irá produzir um melhoramento na aproximação assintótica da distribuição da estatística do teste pela distribuição qui-quadrado. Frydenberg e Jensen (1989) mostraram que, em situações particulares, a correção de Bartlett nem sempre é viável, isto é, nem sempre melhora a aproximação assintótica da distribuição da estatística do teste pela distribuição qui-quadrado. No entanto, Cordeiro (1982) e Cysneiros (1997) mostraram, via simulação, que os testes baseados nas estatísticas da razão de verossimilhanças e escore corrigidas apresentam taxas de rejeição mais próximas dos seus respectivos níveis nominais, especialmente nos modelos discretos, evidenciando que as correções foram eficazes.

3.2 Correção de Bartlett e tipo-Bartlett nos MSPNLGs

Considere $Y_1, \ldots, Y_n$ variáveis aleatórias discretas independentes, com função de probabilidade dada por:

$$\pi(y; \mu_i, \phi) = \frac{a(y, \phi) g(\mu_i, \phi)^y}{f(\mu_i, \phi)}, \qquad y \in A_\epsilon, \tag{3.1}$$

em que o suporte de Y_i é um subconjunto A_ϵ dos inteiros $\{\epsilon, \epsilon+1, \ldots\}$, $\epsilon \geq 0$, não dependendo 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; $\phi > 0$ (que assumimos ser conhecido) e $\mu_i > 0$ são chamados de parâmetros de dispersão e de média, respectivamente.

Para a família de distribuições dada em (3.1), as seguintes relações são válidas:

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

em que o índice sobrescrito ($'$) indica a primeira diferenciação em relação a μ. Os modelos em séries de potência não-lineares generalizados são definidos pela função de probabilidade e pelo componente sistemático

$$h(\mu_i) = \eta_i = \eta(x_i; \boldsymbol{\beta}), \quad i = 1, \ldots, n$$

em que $h(\cdot)$ é uma função de ligação conhecida e duplamente diferenciável, $\boldsymbol{\beta = (\beta_1, \ldots, \beta_p)^T}$ é um vetor de p $(p < n)$ parâmetros desconhecidos a serem estimados, $\boldsymbol{x_i = (x_{i1}, \ldots, x_{ik})^T}$ 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 de tal forma que a matriz de derivadas $\tilde{X} = \tilde{X}(\beta) = \partial\eta/\partial\beta^T$, com $\boldsymbol{\eta = (\eta_1, \ldots, \eta_n)^T}$, tem posto p para todo β. A matriz $\tilde{X}$ tem elementos que são, geralmente, funções do vetor de parâmetros $\boldsymbol{\beta}$ desconhecidos. Sejam $\tilde{x}_i'$ a i-ésima linha da matriz $\tilde{X}$ e $\tilde{\tilde{X}}$ uma matriz de dimensão $p \times p$ cujo elemento (r,s) é $\tilde{x}_{irs} = \partial^2\eta_i/\partial\beta_r\partial\beta_s$, $i = 1, \ldots, n$.

Dado o vetor de observações $\boldsymbol{(y_1, \ldots, y_n)^T}$, o logaritmo da função de verossimilhança do vetor de parâmetros $\boldsymbol{\beta}$, dos MSPNLGs pode ser expresso na forma

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

Assumimos que $l(\beta)$ seja regular com respeito às derivadas em relação aos componentes de β até quarta ordem.

Nosso interesse aqui é testar a hipótese nula H$_0$: $\boldsymbol{\beta_1} = \boldsymbol{\beta_1^{(0)}}$ contra a hipótese alternativa H$_1$: $\boldsymbol{\beta_1} \neq \boldsymbol{\beta_1^{(0)}}$, em que $\boldsymbol{\beta = (\beta_1^T, \beta_2^T)^T}$, sendo $\boldsymbol{\beta_1 = (\beta_1, \ldots, \beta_q)^T}$ um vetor de parâmetros de interesse, q dimensional, $\boldsymbol{\beta_2 = (\beta_{q+1}, \ldots, \beta_p)^T}$ um vetor de parâmetros de perturbação, $p - q$ dimensional e $\boldsymbol{\beta_1^{(0)}}$ é um vetor de constantes conhecidas.

Seja $U(\boldsymbol{\beta}) = \partial l(\boldsymbol{\beta})/\partial\boldsymbol{\beta} = (U_1(\boldsymbol{\beta})^T, U_2(\boldsymbol{\beta})^T)^T$ a função escore total de $\boldsymbol{\beta}$. Considere o estimador de máxima verossimilhança irrestrito de $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_1^T, \hat{\boldsymbol{\beta}}_2^T)^T$ e $\tilde{\boldsymbol{\beta}} = (\boldsymbol{\beta}_1^{(0)T}, \tilde{\boldsymbol{\beta}}_2^T)^T$ o estimador de máxima verossimilhança restrito sob a hipótese nula. A matriz de informação de Fisher e a sua inversa são decompostas como a seguir:

$$K(\boldsymbol{\beta}) = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \quad e \quad K(\boldsymbol{\beta})^{-1} = \begin{bmatrix} K^{11} & K^{12} \\ K^{21} & K^{22} \end{bmatrix},$$

respectivamente, sendo $K_{11} = X_1^T W X_1$, $K_{22} = X_2^T W X_2$, $K_{12} = K_{21}^T = X_1^T W X_2$, $K^{11} = (R^T W R)^{-1}$, $K^{22} = (X_2^T W X_2)^{-1} + C(R^T W R)^{-1} C^T$ (covariância assintótica de $\hat{\beta}_2$) e $K^{12} = K^{21} = (R^T W R)^{-1} C^T$, em que $R = X_1 - X_2 C$ e $C = (X_2^T W X_2)^{-1} X_2^T W X_1$ é uma matriz $n \times q$, sendo suas colunas vetores dos coeficientes de regressão linear das colunas de X_1 sobre X_2, tal que o peso é W.

Definimos os escalares t_i, q_i, w_{ji}, $\tilde{w}_{ji}$ e $\tilde{w}_{1i}^*$, respectivamente, por:

$$t_i = \frac{g_i^{\{1\}}}{g_i h_i^{\{1\}}} \quad , \quad q_i = \frac{f_i^{\{1\}}}{f_i h_i^{\{1\}}},$$

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

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 $i = 1, \ldots, n$, em que o índice sobrescrito $\{j\}$ indica a j-ésima derivada em relação a μ. Vale salientar 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.

Além disso, as matrizes diagonais Q_1, Q_2, Q_3 e Q_4 são necessárias ao cálculo do fator de correção de Bartlett, e as matrizes Q_5, Q_6, Q_7, Q_8 e Q_9 são usadas para o cálculo do fator de correção tipo-Bartlett, todas de dimensão $n \times n$. Além disso, vale ressaltar dois aspectos: as matrizes foram denotadas dessa forma por simplicidade, uma vez que essa é a notação utilizada por Silva (2010) para algumas matrizes que surgem nos cálculos dos cumulantes; e que essas matrizes nada têm a ver com o vetor $\boldsymbol{Q}$ definido na Seção (2.3). Os elementos de tais matrizes são definidos, respectivamente, por:

$$
\begin{aligned}
q_{1i} &= w_{2i} - \frac{w_{1i} h_i^{\{2\}}}{(h_i^{\{1\}})^2}, \\
q_{2i} &= \frac{1}{6}\left(w_{2i} - \frac{w_{1i} h_i^{\{2\}}}{(h_i^{\{1\}})^2} \right) - \tilde{w}_{1i}, \\
q_{3i} &= \frac{1}{4}\left(w_{2i} - \frac{w_{1i} h_i^{\{2\}}}{(h_i^{\{1\}})^2} \right) - \tilde{w}_{1i} \\
q_{4i} &= \frac{1}{4} w_{3i} - \frac{3}{4}\frac{w_{2i} h_i^{\{2\}}}{(h_i^{\{1\}})^2} + \frac{3}{4}\frac{w_{1i} h_i^{\{3\}}}{(h_i^{\{1\}})^3} - \frac{5}{4}\frac{w_{1i} (h_i^{\{2\}})^2}{(h_i^{\{1\}})^4} + \frac{\tilde{w}_{1i} h_i^{\{2\}}}{(h_i^{\{1\}})^2} - \tilde{w}_{2i} + \tilde{w}_{1i}^{*}, \\
q_{5i} &= \frac{w_{1i} h_i^{\{1\}} + q_i^{\{1\}}}{\mu_i (h_i^{\{1\}})^2}, \\
q_{6i} &= \frac{q_i}{\mu_i h_i^{\{1\}}}, \\
q_{7i} &= w_{3i} - \frac{3 w_{2i} h_i^{\{2\}}}{(h_i^{\{1\}})^2} - \frac{w_{1i} h_i^{\{2\}}}{(h_i^{\{1\}})^3} + \frac{3 w_{1i} (h_i^{\{2\}})^2}{(h_i^{\{1\}})^4}, \\
q_{8i} &= \tilde{w}_{2i} - \frac{w_{1i} h_i^{\{2\}}}{(h_i^{\{1\}})^2} - \frac{w_{1i} h_i^{\{3\}}}{(h_i^{\{1\}})^3} + \frac{2 w_{1i} (h_i^{\{2\}})^2}{(h_i^{\{1\}})^4} \qquad \text{e} \\
q_{9i} &= \frac{V_i w_{1i}^2}{\mu_i} + \frac{2 V_i w_{1i} q_i^{(1)}}{\mu_i^2 h_i^{\{1\}}} + \frac{V_i (q_i^{\{1\}})^2}{\mu_i^2 (h_i^{\{1\}})^2}.
\end{aligned}
$$

A estatística da razão de verossimilhanças para o teste de H_0 contra H_1 pode ser reescrita como:

$$
LR = 2\left[\left\{ \ell(\hat{\boldsymbol{\beta}}_{\mathbf{1}}, \hat{\boldsymbol{\beta}}_{\mathbf{2}}) - l(\boldsymbol{\beta}_{\mathbf{1}}, \boldsymbol{\beta}_{\mathbf{2}}) \right\} - \left\{ \ell(\boldsymbol{\beta}_{\mathbf{1}}^{\mathbf{(0)}}, \tilde{\boldsymbol{\beta}}_{\mathbf{2}}) - l(\boldsymbol{\beta}_{\mathbf{1}}, \boldsymbol{\beta}_{\mathbf{2}}) \right\} \right].
$$

Dessa forma, uma aproximação da estatística da razão de verossimilhanças pela distribuição χ^2 pode ser melhorada trocando LR pela estatística modificada LR^* ou, equivalentemente, LR_1^*, dadas por:

$$LR^* = \frac{LR}{1+d} \quad e \quad LR_1^* = LR(1-d)$$

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

$$d = \frac{\epsilon_p - \epsilon_{p-q}}{q}.$$

Silva (2010) obteve o fator de correção de Bartlett para a estatística da razão de verossimilhanças nos MSPNLGs, cuja expressão de ϵ_p foi decomposta da seguinte forma:

$$\epsilon_p = \epsilon_p^{(L)} + \epsilon_p^{(NL)}. \tag{3.2}$$

em que

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

é a parte linear do modelo e

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

é a parte associada à não-linearidade na componente sistemática do modelo, sendo Z, B, C e D, matrizes de dimensão $n \times n$ cujos elementos são, respectivamente, dados por: $z_{ij} = \tilde{x}_i^\top K_{\boldsymbol{\beta}}^{-1} \tilde{x}_j$, $b_{ij} = tr(K_{\boldsymbol{\beta}}^{-1} \tilde{\tilde{X}}_i K_{\boldsymbol{\beta}}^{-1} \tilde{\tilde{X}}_j)$, $c_{ij} = \tilde{x}_i^\top K_{\boldsymbol{\beta}}^{-1} \tilde{\tilde{X}}_j K_{\boldsymbol{\beta}}^{-1} \tilde{x}_i$ e $d_{1i} = tr(K_{\boldsymbol{\beta}}^{-1} \tilde{\tilde{X}}_i)$. Além disso, Z_d, B_d e C_d representam matrizes diagonais formadas pelos correspondentes elementos das diagonais das matrizes Z, B e C, respectivamente e denotamos $Z^{(3)} = Z^{(2)} \odot Z$, $Z^{(2)} = Z \odot Z$, em que $\odot$ é o produto de Hadamard (Rao, 1973, p. 30), ou seja, o elemeto (i,j) de $Z^{(3)}$ é z_{ij}^3. Observe que aqui a notação $Z^{(3)}$ não indica a terceira derivada com relação a μ de Z.

Considere as matrizes $\tilde{\tilde{X}}_{2i} = \partial^2 \eta_i / \partial \boldsymbol{\beta_2} \partial \boldsymbol{\beta_2}^\top$, $i = 1, \ldots, n$, e $K(\boldsymbol{\beta_2}) = \tilde{X}_2^\top W \tilde{X}_2$. De forma análoga à definição feita em (3.2), podemos obter a fórmula de ϵ_{p-q} com $\tilde{X}$, $\tilde{\tilde{X}}_i$ e $K(\boldsymbol{\beta})$ substituídos por $\tilde{X}_2$, $\tilde{\tilde{X}}_{2i}$ e $K(\boldsymbol{\beta_2})$, respectivamente.

A estatística escore (Rao, 1948) para testar H_0 contra H_1 é dada por:

$$S_R = U(\tilde{\boldsymbol{\beta}})^T K(\tilde{\boldsymbol{\beta}})^{-1} U(\tilde{\boldsymbol{\beta}}).$$

Cordeiro e Ferrari (1991) mostraram que é sempre possível obter uma estatística escore modificada possuindo distribuição χ^2 até ordem $O(n^{-1})$. Tal estatística é dada por:

$$S_R^* = S_R \left\{1 - \left(c + bS_R + aS_R^2\right)\right\} \tag{3.3}$$

em que o fator entre chaves é chamado fator de correção tipo-Bartlett e envolve as quantidades a, b e c que são termos de ordem $O(n^{-1})$ e dados por:

$$\begin{aligned} a &= \frac{A_3}{12q\,(q+2)\,(q+4)}, \\ b &= \frac{A_2 - 2A_3}{12q\,(q+2)}, \\ c &= \frac{A_1 - A_2 + A_3}{12q} \end{aligned}$$

sendo q o número de restrições impostas sob a hipótese nula e A_1, A_2 e A_3 funções de cumulantes conjuntos de derivadas do logaritmo da função de verossimilhança e expressas por:

$$\begin{aligned} A_1 &= 3\sum(\kappa_{ijk} + 2\kappa_{i,jk})(\kappa_{rst} + 2\kappa_{r,st})a_{ij}a_{st}m_{kr} \\ &- 6\sum(\kappa_{ijk} + 2\kappa_{i,jk})\kappa_{r,s,t}a_{ij}a_{kr}m_{st} \\ &+ 6\sum(\kappa_{i,jk} - \kappa_{i,j,k})(\kappa_{rst} + 2\kappa_{rs,t})a_{js}a_{kt}m_{ir} \\ &- 6\sum(\kappa_{i,j,k,r} + \kappa_{i,j,kr})a_{kr}m_{ij} \\ &= A_{11} + A_{12} + A_{13} + A_{14}, \end{aligned}$$

$$\begin{aligned} A_2 &= -3\sum \kappa_{i,j,k}\kappa_{r,s,t}a_{kr}m_{ij}m_{st} \\ &+ 6\sum(\kappa_{ijk} + 2\kappa_{i,jk})\kappa_{r,s,t}a_{ij}m_{kr}m_{st} \\ &- 6\sum \kappa_{i,j,k}\kappa_{r,s,t}a_{kt}m_{ir}m_{js} \\ &+ 3\sum \kappa_{i,j,k,r}m_{ij}m_{kr} \\ &= A_{21} + A_{22} + A_{23} + A_{24}, \end{aligned}$$

$$
\begin{aligned}
A_3 &= 3\sum \kappa_{i,j,k}\kappa_{r,s,t}m_{ij}m_{kr}m_{st}\\
&+ 2\sum \kappa_{i,j,k}\kappa_{r,s,t}m_{ir}m_{js}m_{kt}\\
&= A_{31}+A_{32},
\end{aligned}
$$

em que a_{ij} e m_{ij} são os elementos (i,j), respectivamente, das matrizes:

$$
A=\begin{bmatrix} 0 & 0\\ 0 & K_{22}^{-1}\end{bmatrix} \quad \text{e} \quad M=K(\boldsymbol{\beta})^{-1}-A.
$$

Aqui, K_{22}^{-1} é a matriz de covariância assintótica de $\tilde{\boldsymbol{\beta}}_{\mathbf{2}}$ e $\sum$ denota o somatório sobre todos os componentes de $\boldsymbol{\beta}$, isto é, sobre os p-parâmetros. Além disso, vale salientar que os $A's$ são de ordem $O(n^{-1})$ e avaliados sob H_0. Vale salientar que, caso os $A's$ envolvam parâmetros desconhecidos, eles poderão ser substituídos por seus respectivos estimadores de máxima verossimilhança sob H_0, não afetando a ordem de aproximação da correção.

A fórmula da estatística escore modificada S_R^*, dada em (3.3), está fortemente relacionada com os quantis modificados de S_R, que foram obtidos por Harris (1985). Além disso, ele mostrou que P$(S_R \leq z_\alpha)$ =P$(\chi_q^2 \leq x_\alpha) = 1-\alpha$, quando desprezados os termos de ordem inferior a n^{-1}. Nessa equação: $z_\alpha = x_\alpha\left\{1+b\left(x_\alpha, A_1, A_2, A_3, q\right)\right\}$. Assim, desprezando termos de ordem inferior a n^{-1}, temos que P$(S_R^* \leq x_\alpha)$=P$(S_R \leq z_\alpha)$ =P$(\chi_q^2 \leq x_\alpha) = 1-\alpha$, sob a hipótese nula. Desta forma, um teste escore aperfeiçoado pode ser conduzido das seguintes maneiras: obtém-se a estatística escore aperfeiçoada e utiliza-se a distribuição χ_q^2 como referência, ou obtém-se a estatística escore usual e utilizam-se os quantis modificados z_α como referência.

A estatística modificada S_R^* pode ser função não-monótona de S_R. Como solução a esse problema, Kakizawa (1996) propôs uma transformação monótona dada por $S_{R1}^* = S_R^* + P(S_R)$ em que $P(S_R)$ é dada da seguinte forma:

$$
P(S_R)=\frac{1}{4}\left\{c^2S_R+2bcS_R^2+\left(2ac+\frac{4}{3}b^2\right)S_R^3+3abS_R^4+\frac{9}{5}S_R^5\right\}.
$$

Além disso, Cordeiro, Ferrari e Cysneiros (1998) obtiveram uma estatística escore modificada, S_{R2}^* assintoticamente equivalente a S_R^* com a propriedade de ser monótona em S_R.

Tal estatística é dada por:

$$S_{R2}^{*} = \begin{cases} \sqrt{\frac{\pi}{3a}} exp\left(\frac{b^2}{3a} - c\right)\left\{\phi\left(\sqrt{6}aS_R + \sqrt{\frac{2}{3a}}b\right) - \phi\left(\sqrt{\frac{2}{3a}}b\right)\right\} & , \quad \text{se} \quad \text{a>0}, \\ \dfrac{1}{2b} exp(-c)\left\{1 - exp(-2bS_R)\right\} & , \quad \text{se} \quad \text{a} = 0 \quad \text{e} \quad \text{b} \neq 0, \end{cases}$$

sendo a sempre não negativo.

3.2.1 Obtenção dos A's

A obtenção dos $A's$ é feita da seguinte maneira: basta inserir os cumulantes ($\kappa's$) de derivadas do logaritmo da função de verossimilhança nas equações de Harris (1985). Desta forma, pode-se obter expressões matriciais simples para A_1, A_2 e A_3. Os detalhes algébricos não são inseridos aqui uma vez que constituem cálculos enfadonhos; no entanto, os cálculos seguem de forma similar aos encontrados em Cordeiro et al. (1993). Detalhes sobre a obtenção dos $A's$ encontram-se no Apêndice A. A seguir, são apresentadas as expressões obtidas para o modelo em questão que, em notação matricial, ficam dadas por:

$$\begin{aligned} A_{11} &= 6\iota^T(2\tilde{W}_1 - Q_1)Z_{2d}(Z - Z_2)Z_{2d}\tilde{W}_1\iota + 3\iota^T(Q_1 - 2\tilde{W}_1)Z_{2d}(Z - Z_2)Z_{2d}Q_1\iota \\ &+ 3\iota^T W_1 E_2(Z - Z_2)E_2 W_1\iota + 3\iota^T(2\tilde{W}_1 - Q_1)Z_{2d}(Z - Z_2)E_2 W_1\iota \\ &+ 3\iota^T W_1 E_2 Z_{2d}(Z - Z_2)(2\tilde{W}_1 - Q_1)\iota \end{aligned}$$

$$A_{12} = -6\iota^T(-Q_1 + 2\tilde{W}_1)Z_{2d}(Z - Z_2)_d Z_2(2Q_1 - 3\tilde{W}_1)\iota - 6\iota^T W_1 C_2(Z - Z_2)_d(2Q_1 - 3\tilde{W}_1)\iota,$$

$$\begin{aligned} A_{13} &= 6\iota^T(-3Q_1 + 4\tilde{W}_1)Z_2(Z - Z_2)Z_2(-Q_1 + 2\tilde{W}_1)\iota + 6\iota^T(-3Q_1 + 4\tilde{W}_1)C_2(Z - Z_2)W_1\iota \\ &+ 30\iota^T W_1 C_2(Z - Z_2)(-Q_1 + 2\tilde{W}_1)\iota + 30\iota^T W_1 J_2(Z - Z_2)W_1\iota, \end{aligned}$$

$$\begin{aligned} A_{14} &= 6\iota^T(-2Q_9 - 6Q_8 + 2Q_7 + 5W_1^*)Z_{2d}(Z - Z_2)_d\iota \\ &+ 6\iota^T(7\tilde{W}_1 - 6Q_1 - 4Q_5)(Z - Z_2)_d E_2\iota, \end{aligned}$$

$$A_{21} = -3\iota^T(2Q_1 - 3\tilde{W}_1)(Z - Z_2)_d Z_2 (Z - Z_2)_d (2Q_1 - 3\tilde{W}_1)\iota$$

$$\begin{aligned} A_{22} &= 6\iota^T(-Q_1 + 2\tilde{W}_1)(Z - Z_2)_d Z_{2d} (Z - Z_2)(2Q_1 - 3\tilde{W}_1)\iota \\ &+ 6\iota^T W_1 E_2 (Z - Z_2)(Z - Z_2)_d (2Q_1 - 3\tilde{W}_1)\iota \end{aligned}$$

$$A_{23} = -6\iota^T(2Q_1 - 3\tilde{W}_1)(Z - Z_2) Z_2 (Z - Z_2)(2Q_1 - 3\tilde{W}_1)\iota$$

$$A_{24} = 3\iota^T(8Q_8 - 3Q_7 + 3Q_9 - 6\tilde{W}_1^*)(Z - Z_2)_d (Z - Z_2)_d \iota,$$

$$\begin{aligned} A_{31} &= 6\iota^T Q_1 (Z - Z_2)_d (Z - Z_2)(Z - Z_2)_d (2Q_1 - 3W_1)\iota \\ &+ 9\iota^T W_1 (Z - Z_2)_d (Z - Z_2)(Z - Z_2)_d (3W_1 - 2Q_1)\iota, \end{aligned}$$

$$\begin{aligned} A_{32} &= 4\iota^T Q_1 (Z - Z_2)(Z - Z_2)(Z - Z_2)(2Q_1 - 3\tilde{W}_1)\iota \\ &- 6\iota^T \tilde{W}_1 (Z - Z_2)(Z - Z_2)(Z - Z_2)(2Q_1 - 3\tilde{W}_1)\iota, \end{aligned}$$

em que $Z = X(X^TWX)^{-1}X^T$, $Z_2 = X_2(X_2^TWX_2)^{-1}X_2^T$, Z_d e Z_{2d} contém os elementos da diagonal da matriz Z e Z_2, respectivamente. Além disso, as matrizes C_2, E_2 e J_2 possuem elementos dados, respectivamente, por $c_2 = \sum \tilde{x}_{li} a_{ij} \tilde{x}_{ljk} a_{kr} \tilde{x}_{mr}$, $e_2 = \sum \tilde{x}_{lij} a_{ij}$ e $j_2 = \sum \tilde{x}_{ljk} a_{js} \tilde{x}_{mst} a_{kt}$.

Capítulo 4

Resultados numéricos

Nesse capítulo apresentaremos alguns resultados de simulações Monte Carlo para avaliar a eficácia da correção tipo-Bartlett nos MSPNLGs. Para tanto, comparamos o desempenho dos testes baseados nas seguintes estatísticas, a saber: razão de verossimilhanças usual (LR), suas versões corrigidas (LR^* e LR_1^*), escore (S_R), suas versões corrigidas (S_R^*, $S_{R_1}^*$ e $S_{R_2}^*$) e gradiente (S_T).

O estudo de simulação foi baseado nas seguintes distribuições: Binomial Negativa Generalizada (BNG), Poisson Generalizada (GPO) e Consul. No caso da distribuição BNG, os parâmetros foram fixados em $\phi = 1$ e $\nu = 3$. Para o caso das distribuições GPO e Consul fixou-se os parâmetros em $\phi = 0,2$ e $\phi = 1,0$, respectivamente.

As simulações realizadas são baseadas no seguinte preditor:

$$\eta_i \quad = \quad \sum_{j=1}^{k} \beta_j x_{ij} + \exp(\beta_7 x_{i7}), i = 1, \ldots, n \quad \text{e} \quad k = 0, \ldots, 6.$$

Duas hipóteses serão consideradas neste estudo de simulação: (i) $H_0 : \beta_7 = 0$ contra $H_0 : \beta_7 \neq 0$ e (ii) $H_0 : \beta_5 = \beta_6 = 0$ contra $H_0 : \beta_5 = \beta_6 \neq 0$. A variável resposta foi gerada assumindo que todos os parâmetros de perturbação são iguais a 0.05, para cada caso. As covariadas $x_1, \ldots, x_7$ foram geradas como amostras aleatórias das distribuições $U(0,1)$, $F(2,5)$, Cauchy, $N(0,2)$, $LN(0,1)$, χ_3^2, $Exp(1)$ e $Beta(2,3)$. O número de réplicas Monte Carlo foi fixado em 10.000 e foram considerados os seguintes níveis nominais $\alpha = 1\%, 3\%, 5\%$

e 10%. As simulações foram realizadas usando a linguagem de programação matricial Ox (Doornik, 2006).

Para cada tamanho amostral e cada nível considerado estimamos via simulação $P(LR \geq \chi^2_{(\alpha;q)})$,$P(LR^* \geq \chi^2_{(\alpha;q)})$,$P(LR^*_1 \geq \chi^2_{(\alpha;q)})$, $P(S_R \geq \chi^2_{(\alpha;q)})$, $P(S^*_R \geq \chi^2_{(\alpha;q)})$, $P(S^*_{R_1} \geq \chi^2_{(\alpha;q)})$, $P(S^*_{R_2} \geq \chi^2_{(\alpha;q)})$ e $P(S_T \geq \chi^2_{(\alpha;q)})$, em que $\chi^2_{(\alpha;q)}$ é o percentil $(1-\alpha)$ da distribuição χ^2_q. Todas as entradas das Tabelas apresentadas são porcentagens.

Nas Tabelas 4.1 e 4.2, fixamos o número de parâmetros de parâmetros em $p = 4$ e variamos o tamanho amostral em $n = 30; 40; 50$ e 100, para as distribuições Consul e GPO, respectivamente. Consideramos inicialmente a hipótese dada em (i). Observamos, nas duas tabelas, que o teste da razão de verossimilhanças, em geral, apresentou taxas de rejeição mais próximas dos níveis nominais correspondentes que os testes baseados nas estatísticas da razão de verossimilhanças corrigidas que, por sua vez são conservativos, pois apresentaram taxas de rejeição menores que os níveis nominais considerados. Por exemplo, para $n = 30$, $p = 4$ e $\alpha = 5\%$ as taxas de rejeição dos testes baseados nas estatísticas LR, LR^* e LR^*_1 são, respectivamente, 4.9%, 4.5% e 4.4% para o modelo Consul e 4.9%, 4.8% e 4.8%, para o modelo GPO. No entanto, observamos que os testes baseados nas estatísticas escore, suas versões corrigidas e gradiente (S_R S^*_R, $S^*_{R_1}$, $S^*_{R_2}$ e S_T, respectivamente) são bastante liberais apresentando taxas de rejeição acima dos níveis nominais correspondentes. Novamente, quando $n = 30$, $p = 4$ e $\alpha = 5\%$ as taxas de rejeição destes testes são, respectivamente, 8.3%, 7.7%, 8.3%, 8.4% e 8.2% para o modelo Consul e 7.5%, 7.1%, 7.3%, 7.4% e 7.3% para o modelo GPO. Observa-se ainda que o fator de correção tipo-Bartlett tende a corrigir a tendência liberal do teste original. Entre as versões corrigidas da estatística escore, o teste baseado na estatística S^*_R foi o que apresentou melhor desempenho. Vale salientar aqui que os testes baseados nas estatísticas S_R e S_T apresentaram piores desempenhos em ambos os cenários. À medida em que o tamanho amostral vai aumentando, as taxas de rejeição de todos os testes se aproximam dos seus respectivos níveis nominais, como era de se esperar.

Os resultados encontrados na Tabela 4.3 foram obtidos levando em consideração a hipótese alternativa $H_1 : \beta_7 \neq 0$ para $n = 30$, $p = 4$, $\alpha = 5\%$ e diversos valores de $\beta_7 = \beta^{(0)}$, tal que $0,05 \leq \beta^{(0)} \leq 0,30$ e considerando o modelo Consul. Observe que estas simulações de poder correspondem ao cenário abordado na Tabela 4.1 para $n = 30$. Além disso, as simulações do poder foram feitas usando valores críticos estimados e não valores tabulados. Adotamos este procedimento porque a maioria dos testes são anti-conservativos. Dessa forma, as simulações foram ajustadas para que todos os testes tenham o mesmo tamanho.

Analisando a Tabela 4.3, podemos observar que entre os testes baseados nas estatísticas usuais, o teste escore tem melhor desempenho do que o teste da razão de verossimilhanças. Entre os testes baseados nas versões corrigidas, o teste baseado na estatística $S^*_{R_1}$ foi o que apresentou melhor desempenho seguido pelos testes baseados nas estatísticas $S^*_{R_2}$, S^*_R, LR^* e LR^*_1. Vale salientar aqui que o teste baseado na estatística gradiente apresentou pior desempenho.

Tabela 4.1: Tamanho dos testes - Modelo não-linear Consul - p=4 e diferentes valores de n - ($H_0 : \beta_7 = 0$).

		Modelo Consul							
n	$\alpha(\%)$	LR	LR^*	LR_1^*	S_R	S_R^*	$S_{R_1}^*$	$S_{R_2}^*$	S_T
30	5	4.9	4.5	4.4	8.3	7.7	8.3	8.4	8.2
	10	9.8	9.1	9.0	13.2	11.0	11.6	11.8	13.2
40	5	4.7	4.4	4.4	7.6	7.2	7.5	7.6	7.5
	10	9.5	9.0	8.9	12.6	10.9	11.2	11.3	12.6
50	5	4.8	4.4	4.4	7.2	6.9	7.1	7.1	6.9
	10	9.2	8.7	8.7	11.9	10.7	10.9	10.9	11.8
100	5	4.9	4.7	4.7	6.1	5.9	5.9	5.9	6.1
	10	9.7	9.6	9.6	11.2	10.4	10.4	10.5	11.3

Tabela 4.2: Tamanho dos testes - Modelo não-linear GPO - p=4 e diferentes valores de n - ($H_0 : \beta_7 = 0$).

		Modelo GPO							
n	$\alpha(\%)$	LR	LR^*	LR_1^*	S_R	S_R^*	$S_{R_1}^*$	$S_{R_2}^*$	S_T
30	5	4.9	4.8	4.8	7.5	7.1	7.3	7.4	7.3
	10	9.7	9.5	9.5	12.3	10.9	11.2	11.2	12.3
40	5	4.7	4.6	4.6	7.0	6.7	6.9	6.9	6.9
	10	9.4	9.4	9.4	11.5	10.6	10.6	10.7	11.5
50	5	4.6	4.6	4.6	6.5	6.3	6.4	6.4	6.2
	10	9.2	9.2	9.2	11.3	10.4	10.4	10.5	11.2
100	5	5.0	4.9	4.9	5.8	5.7	5.7	5.7	5.9
	10	9.9	9.9	9.9	10.9	10.5	10.5	10.5	10.9

Tabela 4.3: Poder dos testes - Modelo não-linear Consul com n = 30, p = 4, α= 5%.

	Modelo Consul							
$\beta^{(0)}$	LR	LR^*	LR_1^*	S_R	S_R^*	$S_{R_1}^*$	$S_{R_2}^*$	S_T
0,05	7.7	7.1	7.9	32.6	32.2	34.5	32.0	15.0
0,10	16.4	15.4	16.2	47.5	46.8	49.2	46.6	21.6
0,15	31.7	30.2	30.7	64.2	62.7	65.7	62.4	28.7
0,20	52.2	49.8	50.1	78.7	75.6	79.6	76.2	35.6
0,25	71.7	68.9	68.8	89.5	85.4	90.0	86.4	41.7
0,30	86.9	83.6	82.9	95.5	89.9	95.6	91.6	46.9

Nas Tabelas 4.4 e 4.5, fixamos o número de parâmetros em $p = 7$ e variamos o tamanho amostral em $n = 30, 40, 50$ e 100, para as distribuições BNG e Consul, respectivamente. Consideramos agora a hipótese dada em (ii). Na Tabela 4.4, os resultados indicam que no modelo BNG, o teste da razão de verossimilhanças usual é liberal enquanto que os testes baseados nas versões corrigidas da estatística da razão de verossimilhanças apresentaram taxas de rejeição mais próximas dos níveis nominais. A correção tipo-Bartlett tende a atenuar a tendência liberal do teste escore, de modo que os testes baseados nas versões corrigidas da estatística escore apresentaM menor distorção de tamanho do que o teste baseado em S_R. Por exemplo, quando $n = 30$ e $\alpha = 10\%$, temos que as taxas de rejeição dos testes baseados nas estatísticas LR, LR^*, LR_1^*, S_R, S_R^*, $S_{R_1}^*$, $S_{R_2}^*$ são, respectivamente, 11.1%, 10.3%, 10.2%, 11.9%, 10.0%, 10.1%, 9.1% e 11.4%. Na Tabela 4.5, notamos que no modelo Consul, os testes baseados nas estatísticas usuais, apresentaram taxas superiores aos níveis nominais correspondentes para tamanho amostral pequeno ou mesmo moderado. No entanto, os testes baseados nas versões corrigidas da estatística da razão de verossimilhanças apresentaram melhores desempenhos quando comparados às versões corrigidas da estatística escore. Vale salientar aqui que os testes baseados nas versões corrigidas da estatística escore corrigem a tendência liberal do teste escore usual apresentando, em geral, taxas de rejeição ligeiramente superiores aos níveis nominais. À medida que o tamanho amostral cresce, todos os testes apresentam taxas de rejeição próximas aos respectivos níveis nominais considerados. Além disso, o teste baseado na estatística gradiente (S_T), apresentou o pior desempenho em relação às versões corrigidas dos testes baseados nas estatísticas da razão de verossimilhanças e escore.

Os resultados encontrados na Tabela 4.6 foram obtidos para o modelo BNG, levando em consideração a hipótese alternativa $H_0 : \beta_5 = \beta_6 = 0$ para $n = 30$, $p = 7$, $\alpha = 5\%$ e diversos valores de $\beta_5 = \beta_6 = \beta$, tal que $0.10 \leq \beta \leq 0,80$. Observe que estas simulações de poder correspondem ao cenário abordado na Tabela 4.4 para $n = 30$. Além disso, as simulações do poder foram feitas usando valores críticos estimados e não valores tabulados. Adotamos este procedimento porque a maioria dos testes são anti-conservativos. Dessa forma, as simulações foram ajustadas para que todos os testes tenham o mesmo tamanho.

Analisando a Tabela 4.6, podemos observar que entre os testes baseados nas estatísticas da razão de verossimilhanças, escore e gradiente, o teste da razão de verossimilhanças teve um melhor desempenho do que os testes gradiente e escore. Entre os testes baseados nas versões corrigidas, o teste baseado na estatística LR_1^* foi o que apresentou melhor desempenho seguido pelos testes baseados nas estatísticas LR^*, $S_{R_2}^*$, $S_{R_1}^*$, S_R^*.

Tabela 4.4: Tamanho dos testes - Modelo não-linear BNG - p=7 e diferentes valores de n - ($H_0 : \beta_5 = \beta_6 = 0$).

		Modelo BNG							
n	$\alpha(\%)$	LR	LR^*	LR_1^*	S_R	S_R^*	$S_{R_1}^*$	$S_{R_2}^*$	S_T
30	5	5.5	5.1	5.1	6.4	5.7	5.7	5.1	5.9
	10	11.1	10.3	10.2	11.9	10.0	10.1	9.1	11.4
40	5	5.7	5.3	5.2	6.4	5.7	5.8	5.2	5.8
	10	11.1	10.7	10.7	11.6	10.4	10.4	9.4	11.4
50	5	5.5	5.2	5.2	6.0	5.4	5.5	5.0	5.7
	10	10.7	10.2	10.2	11.1	10.0	10.1	9.1	10.9
100	5	5.2	5.0	5.0	5.4	5.1	5.1	4.8	5.3
	10	10.3	10.0	10.0	10.7	9.9	9.9	9.5	10.3

Tabela 4.5: Tamanho dos testes - Modelo não-linear Consul - p=7 e diferentes valores de n - ($H_0 : \beta_5 = \beta_6 = 0$).

		Modelo Consul							
n	$\alpha(\%)$	LR	LR^*	LR_1^*	S_R	S_R^*	$S_{R_1}^*$	$S_{R_2}^*$	S_T
30	5	6.4	5.3	5.2	7.4	5.9	6.2	5.5	6.8
	10	12.2	10.3	10.1	12.5	9.0	9.4	8.5	12.5
40	5	6.0	5.1	5.0	6.9	5.7	6.0	5.2	6.2
	10	11.8	10.4	10.3	12.2	9.3	9.6	8.5	12.0
50	5	5.5	4.8	4.8	6.4	5.5	5.6	5.0	5.6
	10	11.1	10.0	10.0	11.6	9.4	9.5	8.4	11.1
100	5	5.2	4.9	4.9	5.7	5.1	5.1	4.6	5.3
	10	10.4	9.8	9.7	11.1	9.3	9.4	8.6	10.4

Tabela 4.6: Poder dos testes em modelos não lineares com $n = 30$, $p = 7$, $\alpha = 5\%$ e considerando o modelo BNG.

	Modelo BNG							
$\beta^{(0)}$	LR	LR^*	LR_1^*	S_R	S_R^*	$S_{R_1}^*$	$S_{R_2}^*$	S_T
0,10	8.9	8.2	8.2	7.6	7.0	7.1	6.9	8.4
0,20	22.4	21.4	21.4	18.5	17.8	17.9	18.0	20.7
0,30	45.3	43.7	43.7	36.2	35.5	35.6	35.8	42.6
0,40	70.3	69.0	69.0	55.4	54.1	54.2	54.4	67.8
0,50	87.9	87.1	87.1	72.8	72.0	71.2	71.3	86.9
0,60	98.8	98.7	98.7	90.1	91.3	91.4	91.3	98.8
0,70	99.1	99.0	99.0	93.7	92.4	92.5	92.5	99.1
0,80	99.8	99.8	99.8	97.8	97.0	97.2	97.2	99.8

4.1 Aplicação

Essa seção tem por objetivo ilustrar a metodologia apresentada anteriormente através do conjunto de dados reais (Tabela B.1). Tal tabela se refere ao número de espécies de peixes em um lago (variável resposta) e o logaritmo da área do lado, dada em Km^2,(x). Tais dados foram analisados primeiramente por Barbour e Brown (1974) e, muitos anos depois, por Rigby et al. (2008) e por Cordeiro et al. (2009). O trabalho de Barbour e Brown (1974) foi um dos primeiros que tomou lagos como ilhas, considerando um contexto biogeográfico. Eles examinaram uma diversidade de peixes em lagos americanos e estudaram trabalhos como os de Sepkoski& Rex (1974), que considerava a dispersão de mexilhões; Lassen (1975); Aho (1978), que considerava gastrópodes e Browne (1981), que considerava zooplâncton. Diante de vários trabalhos na área, algumas conclusões foram tomadas, como a do fato de que lagos maiores e mais profundos podem abrigar um potencial número de espécies, já que apresentam uma área de interações maior.

Cordeiro et al. (2009), discutiram a flexibilidade do uso dos MSPNLGs com o objetivo de ajustar os dados. Dentre os modelos analisados, o mais adequado para ajustar o número de espécies de peixes, segundo o critério de Informação de Akaike (AIC), foi o modelo Delta Binomial (DB), que forneceu AIC igual a 612.1, menor AIC dentre os modelos analisados. Tal resultado coincide com o encontrado por Cordeiro et al. (2009) e Silva (2010).

O preditor utilizado é dado por:

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

$i = 1, \ldots, 70$, em que $\eta_i = \log(\mu_i - m)$ e m representa o valor mínimo do suporte da distribuição associada ao modelo. Aqui, testaremos a hipótese $H_0 : \beta_2 = 0$ contra $H_1 : \beta_2 \neq 0$, isto é, queremos averiguar qual o modelo mais adequado para os dados em questão. Seguindo a mesma linha de raciocínio de Cordeiro et al. (2009) e a de Silva (2010), tomamos os mesmos modelos por eles analisados: Poisson, Binomial Negativa (BN), Poisson Generalizada (GPO), Binomial Negativa Generalizada (BNG) e Delta Binomial (DB). Na Tabela 4.8, apresentamos os resultados dos testes LR, LR^*, LR_1^*, S_R, S_R^*, S_T, $S_{R_1}^*$ e $S_{R_2}^*$. O comportamento das três primeiras estatísticas foi analisado por Silva (2010), que concluiu que ao nível nominal de 10% e considerando os modelos Poisson e BNG, todos os testes rejeitam a hipótese nula. Já para os modelos BN e GPO, os testes considerados não rejeitam a hipótese nula considerando o mesmo nível nominal.

Analisaremos agora os testes S_R, S_R^* e S_T. Note que, ao nível de 10% e considerando os modelos Poisson e BNG, todos os testes rejeitam a hipótese nula, como no caso anterior. Para o caso em que consideramos o modelo BN e GPO, todos os testes não rejeitam a hipótese nula ao nível de 10%. Enquanto que para o modelo DB, só há não-rejeição da hipótese nula para o teste S_R^*. Quanto aos testes $S_{R_1}^*$ e $S_{R_2}^*$, podemos observar que, ao nível nominal de 10% e considerando os modelos Poisson, BNG e DB, os dois testes rejeitam a hipótese nula. Enquanto que considerando o modelo GPO, eles não rejeitam H_0 ao nível nominal de 10%. Já no modelo BN, apenas o teste $S_{R_1}^*$ rejeita a hipótese nula a 10%.

Tabela 4.7: Valor das estatísticas do teste e p-valor - dados reais.

	Estatísticas							
Distribuições	LR	LR^*	LR_1^*	S_R	S_R^*	$S_{R_1}^*$	$S_{R_2}^*$	S_T
Poisson	46,8610	46,8270	46,8270	50,838	46,765	46,926	47,067	45,123
	(0,0000)	(0,0000)	(0,0000)	(0,0000)	(0,0000)	(0,0000)	(0,0000)	(0,0000)
BN	2,4882	2,4439	2,4431	2,6835	2,4080	2,4187	2,4282	2,5723
	(0,1147)	(0,1180)	(0,1180)	(0,10139)	(0,12072)	(0.11990)	(0,11917)	(0,10875)
GPO ($\phi = 0,3$)	0,7701	0,7361	0,7346	0,89139	0,079259	0,26858	0,36191	0,81674
	(0,3802)	(0,3909)	(0,3914)	(0,34510)	(0,77830)	(0,60429)	(0,54745)	(0,36613)
BNG ($\phi = 1$, $\nu = 2,43$)	5,6522	5,6129	5,6126	6,0578	6,0983	6,1023	6,1062	5,8375
	(0,0174)	(0,0178)	(0,0178)	(0,013845)	(0,013531)	(0,015688)	(0,013470)	(0,015688)
DB ($\phi = 3$, $m = 5$)	2,7447	2,5624	2,5494	3,1757	2,7049	2,7359	2,7614	2,8985
	(0,0976)	(0,1094)	(0,1103)	(0,074740)	(0,10004)	(0,098116)	(0,096566)	(0,088663)

4.2 Comentários

Nesse capítulo fizemos uma revisão sobre a correção de Bartlett e apresentamos o fator de correção de Bartlett para a estatística da razão de verossimilhanças nos modelos em série de potência não-lineares generalizados (MSPNLGs), obtido por Silva (2010). Obtivemos, então, o fator de correção tipo-Bartlett para a estatística escore nos modelos já mencionados. Com o objetivo de avaliar o desempenho, em amostras finitas, dos testes baseados na estatística da razão de verossimilhanças usual e suas versões corrigidas (LR, LR^* e LR_1^*, respectivamente), nas estatísticas escore e suas versões corrigidas (S_R, S_R^*, S_{R1}^* e S_{R2}^*, respectivamente) e na estatística gradiente (S_T), resultados numéricos foram apresentados para os modelos não-lineares.

Concluímos que os testes baseados nas estatísticas da razão de verossimilhanças corrigidas, apresentaram melhor desempenho que o teste baseado na estatística da razão de verossimilhanças usual, LR, na maioria dos cenários considerados. Ou seja, quando utilizamos as estatísticas LR^* e LR_1^*, obtivemos taxas de rejeição mais próximas dos seus níveis nominais correspondentes que quando usamos a estatística original, LR. Notou-se que as taxas de rejeição dos testes baseados nas estatísticas corrigidas permaneceram mais próximas dos níveis nominais correspondentes, mesmo quando o tamanho amostral era pequeno.

No que diz respeito ao poder dos testes, os testes baseados nas estatísticas da razão de verossimilhanças corrigidas (LR^* e LR_1^*) foram menos poderosos do que o teste baseado na estatística da razão de verossimilhanças usual (LR), dando indícios de que, possivelmente, ajustar o tamanho implica em alguma diminuição de poder, para algum caso.

Concluímos ainda que entre os testes baseados nas estatísticas corrigidas da estatística escore, o que apresentou melhor desempenho foi o teste baseado na estatística S_R^*. Podemos afirmar isso uma vez que ao usarmos a estatística corrigida via correção tipo-Bartlett, obtivemos taxas de rejeição mais próximas dos seus respectivos níveis nominais, principalmente quando o tamanho amostral. Assim, o ganho com uso da correção tipo-Bartlett aplicada à estatística escore é mais claro. Dessa forma, o uso de tal correção é eficaz, exibindo taxas de rejeição mais próximas dos seus níveis nominais correspondentes. Por consequência, as taxas de rejeição dos testes baseados na estatística corrigida S_R^* continuam próximos dos seus níveis nominais correspondentes, mesmo que o tamanho amostral. Quanto ao poder dos testes baseados nas estatísticas S_R e suas versões corrigidas, notou-se que dentre os testes baseados nas estatísticas corrigidas o que apresentou melhor desempenho foi aquele baseado na estatística $S^*_{R_1}$.

Finalmente, levando em consideração todas as simulações realizadas envolvendo tamanho e poder nos modelos não-lineares trabalhados, aconselhamos o uso dos testes baseados nas estatísticas da razão de verossimilhanças corrigidas LR^* e LR^*_1, no lugar da estatística usual LR, como foi recomendado também por Silva (2010). Além disso, aconselhamos também o uso dos testes baseados na estatítica escore corrigida via correção tipo-Bartlett S^*_R no lugar das estatísticas S_R, S^*_{R1} e S^*_{R2} em inferências sobre os parâmetros nos MSPNLGs.

Capítulo 5

Conclusões

Nossa contribuição com essa dissertação foi apresentar o fator de correção tipo-Bartlett usado para aperfeiçoar o teste baseado na estatística escore (S_R) para os modelos em série de potência não-lineares generalizados, considerando fixo o parâmetro de dispersão e, para a distribuição Binomial Negativa Generalizada, consideramos fixo o parâmetro ν.

Além disso, estudos de simulação foram realizados com o intuito de analisar o efeito das correções feitas nos modelos em questão. De tais estudos podemos concluir que:

- Os resultados de simulação de Monte Carlo para avaliar o desempenho dos testes baseados nas estatísticas da razão de verossimilhanças e suas versões corrigidas (LR, LR^* e LR_1^*, respectivamente) nos MSPNLGs nos mostraram que o teste baseado na estatística da razão de verossimilhanças usual rejeita com maior frequência a hipótese nula. A correção de Bartlett aplicada mostrou-se eficaz, uma vez que produziu taxas de rejeição mais próximas dos seus respectivos níveis nominais, corrigindo a tendência liberal para o teste das hipóteses $H_0 : \beta_5 = \beta_6 = 0$ contra $H_1 : \beta_5 = \beta_6 \neq 0$ e mantendo-se estável para o teste das hipóteses $H_0 : \beta_7 = 0$ contra $H_1 : \beta_7 \neq 0$. Além disso, o impacto do tamanho amostral é mais visível no teste baseado na estatística LR, enquanto que as taxas para os testes baseados nas versões corrigidas permanecem mais estáveis. Quanto aos poderes dos testes, o teste baseado na estatística usual apresentou poder levemente maior com relação à suas versões corrigidas. Tais resultados mostram que não houve nenhum tipo de perda de poder derivada do fato da utilização da correção de Bartlett;

- Os resultados de simulação de Monte Carlo para avaliar o desempenho dos testes baseados nas estatísticas escore (S_R) e suas versões corrigidas (S_R^*, S_{R1}^* e S_{R2}^*) nos MSPNLGs, indicaram que o teste baseado nessas estatísticas foram bastante liberais, ou seja, apresentaram taxas de rejeição superiores aos seus respectivos níveis nominais, rejeitando a hipótese nula com probabilidade maior que o nível de significânica do teste, com o teste baseado na estatística S_R^* apresentando melhor desempenho na maioria das situações colocadas. Assim, a correção tipo-Bartlett mostra-se eficaz, uma vez que produziu taxas de rejeição mais próximas dos níveis nominais correspondentes. Consequentemente, tal correção corrigiu a tendência liberal do teste baseado na estatística original, S_R. Podemos dizer ainda que o tamanho amostral apresentou impacto notável no teste baseado na estatística S_R. Quanto aos poderes dos testes, pudemos notar que o poder do teste baseado na estatística $S_{R_1}^*$ apresentou vantagem quanto aos outros testes baseados nas versões corrigidas da estatística escore. Os resultados mostraram, em suma, que não houve qualquer perda de poder com o uso do fator de correção tipo-Bartlett;

- Os resultados de simulação de Monte Carlo para avaliar o desempenho dos testes baseados nas estatísticas gradiente (S_T) indicam que em todos dos cenários considerados, o teste baseado nessa estatística foi bastante liberal, ou seja, rejeitou a hipótese nula com probabilidade maior que o nível de significânica do teste. Além disso, há um impacto considerável quando aumentamos o tamanho amostral nas taxas de rejeição produzidas pelo teste baseado nessa estatística. No que diz respeito ao poder do teste, quando comparado aos outros testes, ele apresentou pior desempenho.

Apêndice A

Cálculo dos Momentos

Neste apêndice, apresentamos a obtenção de expressões gerais para as quantidades necessárias para a realização do cálculo do fator de correção de tipo-Bartlett para a estatística escore, apresentado na Seção (3.2). Ainda, tendo em vista que algumas expressões foram cruciais para a simplificação dos cálculos, consideramos necessário incluí-las também nesse apêndice.

Para tanto, considere o modelo em séries de potência não-linear generalizado, apresentado na Seção (2.2), com o parâmetro de dispersão ϕ conhecido e o logaritmo da função de verossimilhança do parâmetro β, dado o vetor de observações $(y_1, \ldots, y_n)^\top$, do MSPNLG apresentado na Seção (2.2). Seja $f = f(\mu_i, \phi)$, $g = g(\mu_i, \phi)$ e o índice sobrescrito $\{j\}$ indicando a j-ésima derivada em relação μ, $j = 1, 2, 3$, denotamos as seguintes quantidades, para $i = 1, \ldots, n$, as quais serão utilizadas no cálculo dos $A's$:

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

$$
\begin{aligned}
\varphi_{ji} &= \frac{q_i^{\{1\}} V_i t_i^{\{j\}} + q_i V_i^{\{1\}} t_i^{\{j\}} + q_i V_i t_i^{\{j+1\}}}{(h_i^{\{1\}})^j}, \\
\tilde{w}_{ji} &= \varphi_{ji} - (j-1)\frac{q_i V_i t_i^{\{j\}} h_i^{\{2\}}}{(h_i^{\{1\}})^{j+1}} - \frac{q_i^{\{j+1\}}}{(h_i^{\{1\}})^{j+1}} + j\frac{q_i^{\{j\}} h_i^{\{2\}}}{(h_i^{\{1\}})^{j+2}} \qquad \text{e} \\
\tilde{w}_{ji}^* &= 2\varphi_{(j+1)i} - \frac{q_i V_i t_i^{\{j+2\}}}{(h_i^{\{1\}})^{j+1}} + \frac{t_i^{\{j\}}(q_i^{\{2\}} V_i + 2q_i^{\{1\}} V_i^{\{1\}} + q_i^{\{1\}} V_i^{\{2\}})}{(h_i^{\{1\}})^{j+1}} - (2j-1)\frac{h_i^{\{2\}}\varphi_{ji}}{(h_i^{\{1\}})^2} \\
&+(j-1)\frac{q_i V_i t_i^{\{j\}}}{(h_i^{\{1\}})^{j+2}}\Big[(j+1)\frac{(h_i^{\{2\}})^2}{h_i^{\{1\}}} - h_i^{\{3\}}\Big] - \frac{q_i^{\{j+2\}}}{(h_i^{\{1\}})^{j+2}} + (2j+1)\frac{q_i^{\{j+1\}} h_i^{\{2\}}}{(h_i^{\{1\}})^{j+3}} \\
&+j q_i^{\{j\}}\Big[\frac{h_i^{\{3\}}}{(h_i^{\{1\}})^{j+3}} - (j+2)\frac{(h_i^{\{2\}})^2}{(h_i^{\{1\}})^{j+4}}\Big].
\end{aligned}
$$

No entanto, devemos lembrar que as quantidades dependem das distribuições pertencentes à família de série de potência que estará sendo utilizada, uma vez que envolvem formas específicas das funções f, g, h e V. Considere $\tilde{x}_{ir} = \partial\eta_i/\partial\boldsymbol{\beta}_r$, $\tilde{x}_{irs} = \partial^2\eta_i/\partial\boldsymbol{\beta}_r\partial\boldsymbol{\beta}_s$ e $\tilde{x}_{irst} = \partial^3\eta_i/\partial\boldsymbol{\beta}_r\partial\boldsymbol{\beta}_s\partial\boldsymbol{\beta}_t$, alguns resultados importantes são dados a seguir:

$$
E(d_{0i}) = 0 \quad , \quad E(d_{ji}) = w_{ji} = \frac{\mu_i t_i^{\{j\}} - q_i^{\{j\}}}{(h_i^{\{1\}})^j} \quad , \quad \frac{\partial w_{ji}}{\partial\boldsymbol{\beta}_r} = \tilde{w}_{ji}\tilde{x}_{ir} \quad , \quad \frac{\partial\tilde{w}_{ji}}{\partial\boldsymbol{\beta}_r} = \tilde{w}_{ji}^*\tilde{x}_{ir},
$$

$$
\frac{\partial d_{0i}}{\partial\boldsymbol{\beta}_r} = y_i\frac{\partial t_i}{\partial\mu_i}\frac{\partial\mu_i}{\partial\eta_i}\frac{\partial\eta_i}{\partial\boldsymbol{\beta}_r} - \frac{\partial q_i}{\partial\mu_i}\frac{\partial\mu_i}{\partial\eta_i}\frac{\partial\eta_i}{\partial\boldsymbol{\beta}_r} = y_i t_i^{\{1\}}\frac{1}{h_i^{\{1\}}}\tilde{x}_{ir} - q_i^{(1)}\frac{1}{h_i^{\{1\}}}\tilde{x}_{ir} = d_{1i}\tilde{x}_{ir},
$$

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

e

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

A.1 Derivadas do logaritmo da função de verossimilhança

Aqui, são calculadas as derivadas da função de log-verossimilhança com relação aos componentes de $\boldsymbol{\beta}$. Assim:

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

Analogamente pode-se obter as derivadas segunda, terceira e quarta, apresentadas abaixo:

$$
\begin{aligned}
U_{rs} &= \frac{\partial^2 \ell(\boldsymbol{\beta}; \boldsymbol{y})}{\partial \boldsymbol{\beta}_r \partial \boldsymbol{\beta}_s} \\
&= \sum_i \Big\{\frac{\partial d_{0i}}{\partial \boldsymbol{\beta}_s}\tilde{x}_{ir} + d_{0i}\frac{\partial \tilde{x}_{ir}}{\partial \boldsymbol{\beta}_s}\Big\} = \sum_i \Big\{d_{1i}\tilde{x}_{is}\tilde{x}_{ir} + d_{0i}\tilde{x}_{irs}\Big\},
\end{aligned}
$$

$$
\begin{aligned}
U_{rst} &= \frac{\partial^3 \ell(\boldsymbol{\beta};\boldsymbol{y})}{\partial\boldsymbol{\beta}_r\partial\boldsymbol{\beta}_s\partial\boldsymbol{\beta}_t} \\
&= \sum_i \Big\{ \frac{\partial d_{1i}}{\partial\boldsymbol{\beta}_t}\tilde{x}_{is}\tilde{x}_{ir} + d_{1i}\frac{\partial\tilde{x}_{is}}{\partial\boldsymbol{\beta}_t}\tilde{x}_{ir} + d_{1i}\tilde{x}_{is}\frac{\partial\tilde{x}_{ir}}{\partial\boldsymbol{\beta}_t} + \frac{\partial d_{0i}}{\partial\boldsymbol{\beta}_t}\tilde{x}_{irs} + d_{0i}\frac{\partial\tilde{x}_{irs}}{\partial\boldsymbol{\beta}_t}\Big\} \\
&= \sum_i \Big\{ \Big[d_{2i} - \frac{d_{1i}h_i^{\{2\}}}{(h_i^{\{1\}})^2}\Big]\tilde{x}_{it}\tilde{x}_{is}\tilde{x}_{ir} + d_{1i}\tilde{x}_{ist}\tilde{x}_{ir} + d_{1i}\tilde{x}_{is}\tilde{x}_{irt} + d_{1i}\tilde{x}_{it}\tilde{x}_{irs} + d_{0i}\tilde{x}_{irst}\Big\} \\
&= \sum_i \Big\{ \Big[d_{2i} - \frac{d_{1i}h_i^{\{2\}}}{(h_i^{\{1\}})^2}\Big]\tilde{x}_{it}\tilde{x}_{is}\tilde{x}_{ir} + d_{1i}(\tilde{x}_{ist}\tilde{x}_{ir} + \tilde{x}_{is}\tilde{x}_{irt} + \tilde{x}_{it}\tilde{x}_{irs}) + d_{0i}\tilde{x}_{irst}\Big\}
\end{aligned}
$$

e

$$
\begin{aligned}
U_{rstu} &= \frac{\partial^4 \ell(\boldsymbol{\beta};\boldsymbol{y})}{\partial\boldsymbol{\beta}_r\partial\boldsymbol{\beta}_s\partial\boldsymbol{\beta}_t\partial\boldsymbol{\beta}_u} \\
&= \sum_i \Big\{ \frac{\partial^2 d_{1i}}{\partial\boldsymbol{\beta}_t\partial\boldsymbol{\beta}_u}\tilde{x}_{is}\tilde{x}_{ir} + \Big[d_{2i} - \frac{d_{1i}h_i^{\{2\}}}{(h_i^{\{1\}})^2}\Big]\tilde{x}_{it}(\tilde{x}_{isu}\tilde{x}_{ir} + \tilde{x}_{is}\tilde{x}_{iru}) \\
&\quad + \frac{\partial d_{1i}}{\partial\boldsymbol{\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 + \frac{\partial d_{0i}}{\partial\boldsymbol{\beta}_u}\tilde{x}_{irst} + d_{0i}\tilde{x}_{irstu}\Big\} \\
&= \sum_i \Big\{ \Big[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}\Big]\tilde{x}_{iu}\tilde{x}_{it}\tilde{x}_{is}\tilde{x}_{ir} + \Big[d_{2i} - \frac{d_{1i}h_i^{\{2\}}}{(h_i^{\{1\}})^2}\Big]\tilde{x}_{itu}\tilde{x}_{is}\tilde{x}_{ir} \\
&\quad + \Big[d_{2i} - \frac{d_{1i}h_i^{\{2\}}}{(h_i^{\{1\}})^2}\Big]\tilde{x}_{it}(\tilde{x}_{isu}\tilde{x}_{ir} + \tilde{x}_{is}\tilde{x}_{iru}) + \Big[d_{2i} - \frac{d_{1i}h_i^{\{2\}}}{(h_i^{\{1\}})^2}\Big]\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 + d_{0i}\tilde{x}_{irstu}\Big\} \\
&= \sum_i \Big\{ \Big[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}\Big]\tilde{x}_{iu}\tilde{x}_{it}\tilde{x}_{is}\tilde{x}_{ir} \\
&\quad + \Big[d_{2i} - \frac{d_{1i}h_i^{\{2\}}}{(h_i^{\{1\}})^2}\Big]\big[\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})\big] \\
&\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 + d_{0i}\tilde{x}_{irstu}\Big\}.
\end{aligned}
$$

A.2 Cálculo de cumulantes

Os cumulantes são obtidos tomando as esperanças das expressões supramecionadas e utilizando, em alguns casos, identidades de Bartlett que tem o objetivo de facilitar a obtenção dos $\kappa' s$. Tais identidades podem ser encontradas em Cordeiro (1999, Capítulo 5, Seção 5.2, p.120).

$$
\begin{aligned}
\kappa_{rs} &= \sum_i w_{1i}\tilde{x}_{is}\tilde{x}_{ir}, \\
\kappa_{rst} &= \sum_i \{q_{1i}\tilde{x}_{ir}\tilde{x}_{is}\tilde{x}_{it} + w_{1i}(\tilde{x}_{ir}\tilde{x}_{ist} + \tilde{x}_{is}\tilde{x}_{irt} + \tilde{x}_{it}\tilde{x}_{irs})\}, \\
\kappa_{rstu} &= \sum_i \Big\{q_{7i}\tilde{x}_{ir}\tilde{x}_{is}\tilde{x}_{it}\tilde{x}_{iu} + q_{1i}\big[\tilde{x}_{ir}\tilde{x}_{is}\tilde{x}_{itu} + \tilde{x}_{ir}\tilde{x}_{it}\tilde{x}_{isu} + \tilde{x}_{is}\tilde{x}_{it}\tilde{x}_{iru} \\
&+ \tilde{x}_{ir}(\tilde{x}_{iu}\tilde{x}_{ist} + \tilde{x}_{is}\tilde{x}_{it} + \tilde{x}_{irt}\tilde{x}_{irs})\big] + w_{1i}(\tilde{x}_{ir}\tilde{x}_{istu} + \tilde{x}_{ist}\tilde{x}_{iru} + \tilde{x}_{isu}\tilde{x}_{irt} \\
&+ \tilde{x}_{is}\tilde{x}_{irtu} + \tilde{x}_{itu}\tilde{x}_{irs} + \tilde{x}_{it}\tilde{x}_{irsu} + \tilde{x}_{iu}\tilde{x}_{irst})\Big\}, \\
\kappa_{r,s,t} &= \sum_i \Big\{ (q_{1i} - 3\tilde{w}_{1i})\,\tilde{x}_{ir}\tilde{x}_{is}\tilde{x}_{it} - w_{1i}(\tilde{x}_{ir}\tilde{x}_{ist} + \tilde{x}_{is}\tilde{x}_{irt} + \tilde{x}_{it}\tilde{x}_{irs})\Big\}, \\
\kappa_{r,st} &= \sum_i (-q_{1i} + \tilde{w}_{1i})\tilde{x}_{ir}\tilde{x}_{is}\tilde{x}_{it} - \sum_i w_{1i}\tilde{x}_{ir}\tilde{x}_{ist}, \\
\kappa_{r,s,tu} &= \sum_i \{(q_{7i} - 2q_{8i} + w^*_{1i} - q_{9i})\tilde{x}_{ir}\tilde{x}_{is}\tilde{x}_{it}\tilde{x}_{iu} - (2q_{5i} + \tilde{w}_{1i})\tilde{x}_{ir}\tilde{x}_{is}\tilde{x}_{itu} - (q_{6i} + w_{1i})\tilde{x}_{irs}\tilde{x}_{itu}\} \\
&\text{e} \\
\kappa_{r,s,t,u} &= \sum_i \Big\{(-3q_{7i} + 8q_{8i} - 6w^*_{1i} + 3q_{9i})\tilde{x}_{ir}\tilde{x}_{is}\tilde{x}_{it}\tilde{x}_{iu} + (6q_{5i} + 6q_{1i} - 6\tilde{w}_{1i})\tilde{x}_{ir}\tilde{x}_{is}\tilde{x}_{itu} \\
&+ (3q_{6i} - 3w_{1i})\tilde{x}_{irs}\tilde{x}_{itu}\Big\}
\end{aligned}
$$

A.2.1 Derivadas dos cumulantes

Calculando as derivadas das expressões da Seção anterior com relação aos componentes de β, obtemos:

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

e

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

A.2.2 Cálculo dos A's

Na Seção 3.2.1 dessa dissertação encontram-se os termos dos A's de maneira separada. Esse apêndice tem por objetivo mostrar, os cálulos realizados para a obtenção das expressões dos A's. Para tanto, considere $\tilde{x}_i^\top$ a i-ésima linha da matriz $\tilde{X}$, $\tilde{\tilde{X}}_i$ uma matriz $p \times p$ cujo elemento (r, s) é $\tilde{x}_{irs}$, $i = 1, \ldots, n$, e K_β^{-1} a inversa da matriz de informação de Fisher K_β dada anteriormente. Seja

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

uma matriz $n \times n$, positiva semi-definida de posto p com elementos $z_{ij} = \tilde{x}_i^\top K_\beta^{-1}\tilde{x}_j$. As matrizes B e C, também de dimensão $n \times n$, tem elementos dados por $b_{ij} = tr(K_\beta^{-1}\tilde{\tilde{X}}_i K_\beta^{-1}\tilde{\tilde{X}}_j)$ e $c_{ij} = \tilde{x}_i^\top K_\beta^{-1}\tilde{\tilde{X}}_j K_\beta^{-1}\tilde{x}_i$, respectivamente. Utilizamos aqui a mesma notação utilizada por Silva (2010), em que Z_d, B_d e C_d representam matrizes diagonais formadas pelos correspondentes elementos das diagonais das matrizes Z, B e C, respectivamente. Denotamos, ainda, $Z^{(3)} = Z^{(2)} \odot Z$, $Z^{(2)} = Z \odot Z$, em que $\odot$ é o produto de Hadamard (Rao, 1973, p. 30), ou seja, o elemeto (i,j) de $Z^{(3)}$ é z_{ij}^3. Ainda, definimos as matrizes diagonais Q_1, Q_2, Q_3, Q_4, Q_5, Q_6, Q_7, Q_8 e Q_9 de dimensão $n \times n$ cujos elementos definidos na Seção 3.2. Além disso, as matrizes C_2, E_2 e J_2 possuem elementos dados, respectivamente, por $c_2 = \sum \tilde{x}_{li} a_{ij} \tilde{x}_{ljk} a_{kr} \tilde{x}_{mr}$, $e_2 = \sum \tilde{x}_{lij} a_{ij}$ e $j_2 = \sum \tilde{x}_{ljk} a_{js} \tilde{x}_{mst} a_{kt}$.

Dos resultados encontrados no Apêndice A.2, temos que:

$$\kappa_{ijk} + 2\kappa_{i,jk} = \sum_l \left\{ (2\tilde{w}_{1l} - q_{1l}) \tilde{x}_{li} \tilde{x}_{lj} \tilde{x}_{lk} + w_{1l} (\tilde{x}_{lj} \tilde{x}_{lik} + \tilde{x}_{lk} \tilde{x}_{lij} - \tilde{x}_{li} \tilde{x}_{ljk}) \right\}. \quad (A.1)$$

Logo, teremos:

$$\begin{aligned}
(\kappa_{ijk} + 2\kappa_{i,jk})(\kappa_{rst} + 2\kappa_{rs,t}) &= \sum_{l,m} \Big\{ (4\tilde{w}_{1l}\tilde{w}_{1m} - 2\tilde{w}_{1l} q_{1m} - 2q_{1l}\tilde{w}_{1m} + q_{1l} q_{1m}) \tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} \\
&+ w_{1l} w_{1m} \Big\{ \tilde{x}_{lj}\tilde{x}_{lik}\tilde{x}_{mr}\tilde{x}_{mst} + \tilde{x}_{lj}\tilde{x}_{lik}\tilde{x}_{ms}\tilde{x}_{mrt} \\
&- \tilde{x}_{lj}\tilde{x}_{lik}\tilde{x}_{mt}\tilde{x}_{mrs} + \tilde{x}_{lk}\tilde{x}_{lij}\tilde{x}_{mr}\tilde{x}_{mst} \\
&+ \tilde{x}_{lk}\tilde{x}_{lij}\tilde{x}_{ms}\tilde{x}_{mrt} - \tilde{x}_{lk}\tilde{x}_{lij}\tilde{x}_{mt}\tilde{x}_{mrs} \\
&- \tilde{x}_{li}\tilde{x}_{ljk}\tilde{x}_{mr}\tilde{x}_{mst} + \tilde{x}_{li}\tilde{x}_{ljk}\tilde{x}_{ms}\tilde{x}_{mrt} \\
&+ \tilde{x}_{li}\tilde{x}_{ljk}\tilde{x}_{mt}\tilde{x}_{mrs} \Big\} + 2\tilde{w}_{1l} w_{1m} \Big\{ \tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}\tilde{x}_{mr}\tilde{x}_{mst} \\
&+ \tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}\tilde{x}_{ms}\tilde{x}_{mrt} - \tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}\tilde{x}_{mt}\tilde{x}_{mrs} \Big\} \\
&- q_{1l} w_{1m} \Big\{ \tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}\tilde{x}_{mr}\tilde{x}_{mst} + \tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}\tilde{x}_{ms}\tilde{x}_{mrt} \\
&- \tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}\tilde{x}_{mt}\tilde{x}_{mrs} \Big\} + 2w_{1l}\tilde{w}_{1m} \Big\{ \tilde{x}_{lj}\tilde{x}_{lik}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} \\
&+ \tilde{x}_{lk}\tilde{x}_{lij}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} - \tilde{x}_{li}\tilde{x}_{ljk}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} \Big\} \\
&- w_{1l} q_{1m} \Big\{ \tilde{x}_{lj}\tilde{x}_{lik}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} + \tilde{x}_{lk}\tilde{x}_{lij}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} \\
&- \tilde{x}_{li}\tilde{x}_{ljk}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} \Big\} \Big\}. \qquad (A.2)
\end{aligned}$$

O próximo passo para obter a expressão de A_{11} é multiplicar $(A.1)$ por $a_{ij} a_{st} m_{kr}$ e aplicar o somatório sobre todos os componentes de β. Temos, em termos matriciais, que:

$$\begin{aligned}
&\sum \left(\kappa_{ijk} + 2\kappa_{i,jk}\right)\left(\kappa_{rst} + 2\kappa_{rs,t}\right) a_{ij} a_{st} m_{kr} \\
&= 4\iota^T \tilde{W}_1 Z_{2d}(Z - Z_2) Z_{2d} \tilde{W}_1 \iota - 2\iota^T \tilde{W}_1 Z_{2d}(Z - Z_2) Z_{2d} Q_1 \iota \\
&- 2\iota^T Q_1 Z_{2d}(Z - Z_2) Z_{2d} \tilde{W}_1 \iota + \iota^T Q_1 Z_{2d}(Z - Z_2) Z_{2d} Q_1 \iota \\
&+ \iota^T W_1 E_2 (Z - Z_2) E_2 W_1 \iota + 2\iota^T \tilde{W}_1 Z_{2d}(Z - Z_2) E_2 W_1 \iota \\
&- \iota^T Q_1 Z_{2d}(Z - Z_2) E_2 W_1 \iota + 2\iota^T W_1 E_2 (Z - Z_2) Z_{2d} \tilde{W}_1 \iota \\
&- \iota^T W_1 E_2 (Z - Z_2) Z_{2d} Q_1 \iota. \qquad (A.3)
\end{aligned}$$

Diante disso, usando o resultado em $(A.3)$, podemos reescrever a expressão de A_{11} dada por:

$$\begin{aligned}
A_{11} &= 3\sum \left(\kappa_{ijk} + 2\kappa_{i,jk}\right)\left(\kappa_{rst} + 2\kappa_{rs,t}\right) a_{ij} a_{st} m_{kr} \\
&= 6\iota^T (2\tilde{W}_1 - Q_1) Z_{2d}(Z - Z_2) Z_{2d} \tilde{W}_1 \iota + 3\iota^T (Q_1 - 2\tilde{W}_1) Z_{2d}(Z - Z_2) Z_{2d} Q_1 \iota \\
&+ 3\iota^T W_1 E_2 (Z - Z_2) E_2 W_1 \iota + 3\iota^T (2\tilde{W}_1 - Q_1) Z_{2d}(Z - Z_2) E_2 W_1 \iota \\
&+ 3\iota^T W_1 E_2 Z_{2d}(Z - Z_2)(2\tilde{W}_1 - Q_1)\iota
\end{aligned} \qquad (A.4)$$

Para obter a expressão de A_{12}, iremos fazer primeiramente o produto de $(A.1)$ por $\kappa_{r,s,t}$, dado a seguir:

$$\begin{aligned}
\left(\kappa_{ijk} + 2\kappa_{i,jk}\right)\kappa_{r,s,t} &= \sum_{i,j} \Big\{(-q_{1l} + 2\tilde{w}_{1l})(2q_{1m} - 3\tilde{w}_{1m})\tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} \\
&+ w_{1l}(2q_{1m} - 3\tilde{w}_{1m})\tilde{x}_{lj}\tilde{x}_{lik}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} \\
&- 4w_{1l}w_{1m}\tilde{x}_{lj}\tilde{x}_{lik}\tilde{x}_{mr}\tilde{x}_{mst}\Big\}. \qquad (A.5)
\end{aligned}$$

Agora, basta multiplicar a expressão $(A.5)$ por $a_{ij}a_{kr}m_{st}$. Logo, obtemos, na forma matricial:

$$\begin{aligned}
\sum \left(\kappa_{ijk} + 2\kappa_{i,jk}\right)\kappa_{r,s,t} a_{ij} a_{kr} m_{st} &= \iota^T (-Q_1 + 2\tilde{W}_1) Z_{2d}(Z - Z_2)_d Z_2 (2Q_1 - 3\tilde{W}_1)\iota \\
&+ \iota^T W_1 C_2 (Z - Z_2)_d (2Q_1 - 3\tilde{W}_1)\iota. \qquad (A.6)
\end{aligned}$$

Assim, usando o resultado em $(A.6)$, temos que a expressão de A_{12} fica dada por:

$$\begin{aligned}
A_{12} &= -6\sum(\kappa_{ijk}+2\kappa_{i,jk})\kappa_{r,s,t}a_{ij}a_{kr}m_{st} \\
&= -6\iota^T(-Q_1+2\tilde{W}_1)Z_{2d}(Z-Z_2)_d Z_2(2Q_1-3\tilde{W}_1)\iota \\
&- 6\iota^T W_1 C_2(Z-Z_2)_d(2Q_1-3\tilde{W}_1)\iota \qquad (A.7)
\end{aligned}$$

Para obter a expressão de A_{13}, primeiramente calcularemos a seguinte quantidade:

$$\kappa_{i,jk}-\kappa_{i,j,k} = \sum_l \Big\{(-3q_{1l}+4\tilde{w}_{1l})\tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}+3w_{1l}\tilde{x}_{li}\tilde{x}_{ljk}+w_{1l}(\tilde{x}_{lj}\tilde{x}_{lik}+\tilde{x}_{lk}\tilde{x}_{lij})\Big\} \qquad (A.8)$$

O próximo passo é realizar o produto de $(A.1)$ por $(A.8)$, dado a seguir:

$$\begin{aligned}
(\kappa_{rst}+2\kappa_{rs,t})(\kappa_{i,jk}-\kappa_{i,j,k}) &= \sum_{l,m}\Big\{(-3q_{1l}+4\tilde{w}_{1l})(-q_{1m}+2\tilde{w}_{1m})\tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} \\
&+ (-3q_{1l}+4\tilde{w}_{1l})w_{1m}(\tilde{x}_{mr}\tilde{x}_{mst}+\tilde{x}_{ms}\tilde{x}_{mrt}-\tilde{x}_{mt}\tilde{x}_{mrs})\tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk} \\
&+ 3w_{1l}(-q_{1m}+2\tilde{w}_{1m})\tilde{x}_{li}\tilde{x}_{ljk}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} \\
&+ 3w_{1l}w_{1m}\tilde{x}_{li}\tilde{x}_{ljk}(\tilde{x}_{mr}\tilde{x}_{mst}+\tilde{x}_{ms}\tilde{x}_{mrt}-\tilde{x}_{mt}\tilde{x}_{mrs}) \\
&+ w_{1l}(-q_{1m}+2\tilde{w}_{1m})((\tilde{x}_{lj}\tilde{x}_{lik})+\tilde{x}_{lk}\tilde{x}_{lij})\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt} \\
&+ w_{1l}w_{1m}((\tilde{x}_{lj}\tilde{x}_{lik})+\tilde{x}_{lk}\tilde{x}_{lij})(\tilde{x}_{mr}\tilde{x}_{mst}+\tilde{x}_{ms}\tilde{x}_{mrt} \\
&- \tilde{x}_{mt}\tilde{x}_{mrs})\Big\}. \qquad (A.9)
\end{aligned}$$

Basta agora realizar o produto de $(A.9)$ por $a_{js}a_{kt}m_{ir}$, que é dado matricialmente da seguinte forma:

$$\begin{aligned}
\sum(\kappa_{rst}+2\kappa_{rs,t})(\kappa_{i,jk}-\kappa_{i,j,k})a_{js}a_{kt}m_{ir} &= \iota^T(-3Q_1+4\tilde{W}_1)Z_2(Z-Z_2)Z_2(-Q_1+2\tilde{W}_1)\iota \\
&+ \iota^T(-3Q_1+4\tilde{W}_1)C_2(Z-Z_2)W_1\iota \\
&+ 5\iota^T W_1 C_2(Z-Z_2)(-Q_1+2\tilde{W}_1)\iota \\
&+ 5\iota^T W_1 J_2(Z-Z_2)W_1\iota \qquad (A.10)
\end{aligned}$$

Dessa forma, usando o resultado obtido em $(A.10)$, a expressão de A_{13} fica dada por:

$$\begin{aligned}
A_{13} &= 6\sum(\kappa_{rst}+2\kappa_{rs,t})(\kappa_{i,jk}-\kappa_{i,j,k})a_{js}a_{kt}m_{ir} \\
&= 6\iota^T(-3Q_1+4\tilde{W}_1)Z_2(Z-Z_2)Z_2(-Q_1+2\tilde{W}_1)\iota+6\iota^T(-3Q_1+4\tilde{W}_1)C_2(Z-Z_2)W_1\iota \\
&+ 30\iota^T W_1 C_2(Z-Z_2)(-Q_1+2\tilde{W}_1)\iota+30\iota^T W_1 J_2(Z-Z_2)W_1\iota \qquad (A.11)
\end{aligned}$$

Finalmente, para obter a expressão de A_{14}, precisamos primeiramente calcular a seguinte soma:

$$\begin{aligned}\kappa_{i,j,k,r} + \kappa_{i,j,kr} &= \sum_{i,j} \Big\{ (-2q_{7l} + 6q_{8l} - 5\tilde{w}^{*}_{1l} + 2q_{9l})\tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}\tilde{x}_{lr} \\ &+ (6q_{1l} + 4q_{5l} - 7\tilde{w}_{1l})\tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lkr} \Big\}. \qquad (A.12)\end{aligned}$$

Agora iremos multiplicar o resultado obtido em $(A.12)$ por $a_{kr}m_{ij}$ que, em notação matricial fica dado por:

$$\begin{aligned}(\kappa_{i,j,k,r} + \kappa_{i,j,kr})\, a_{kr}m_{ij} &= -\iota^T(-2Q_9 - 6Q_8 + 2Q_7 + 5W^*_1)Z_{2d}(Z - Z_2)_d\iota \\ &- \iota^T(7\tilde{W}_1 - 6Q_1 - 4Q_5)(Z - Z_2)_d E_2\iota. \qquad (A.13)\end{aligned}$$

Portanto, usando o resultado em $(A.13)$, obtemos a seguinte expressão para A_{14}:

$$\begin{aligned}A_{14} &= -6\sum (\kappa_{i,j,k,r} + \kappa_{i,j,kr})\, a_{kr}a_{kr} \\ &= 6\iota^T(-2Q_9 - 6Q_8 + 2Q_7 + 5W^*_1)Z_{2d}(Z - Z_2)_d\iota \\ &+ 6\iota^T(7\tilde{W}_1 - 6Q_1 - 4Q_5)(Z - Z_2)_d E_2\iota. \qquad (A.14)\end{aligned}$$

Agora faremos os cálculos para a expressão de A_{21}. Precisamos num primeiro momento calcular o seguinte produto:

$$\kappa_{i,j,k}.\kappa_{r,s,t} = \sum_{l,m} \{(2q_{1l} - 3\tilde{w}_{1l})(2q_{1m} - 3\tilde{w}_{1m})\tilde{x}_{li}\tilde{x}_{lj}\tilde{x}_{lk}\tilde{x}_{mr}\tilde{x}_{ms}\tilde{x}_{mt}\}. \qquad (A.15)$$

Assim, multiplicaremos o resultado em $(A.15)$ por $a_{kr}m_{ij}m_{st}$, obtendo a seguinte expressão matricial:

$$\sum \kappa_{i,j,k}.\kappa_{r,s,t}a_{kr}m_{ij}m_{st} = \iota^T(2Q_1 - 3\tilde{W}_1)(Z - Z_2)_d Z_2 (Z - Z_2)_d(2Q_1 - 3\tilde{W}_1)\iota \qquad (A.16)$$

Dessa forma, usando o resultado obtido em $(A.16)$, obtemos a seguinte expressão para A_{21}:

$$\begin{aligned}A_{21} &= -3\sum \kappa_{i,j,k}.\kappa_{r,s,t}a_{kr}m_{ij}m_{st} \\ &= -3\iota^T(2Q_1 - 3\tilde{W}_1)(Z - Z_2)_d Z_2 (Z - Z_2)_d(2Q_1 - 3\tilde{W}_1)\iota. \qquad (A.17)\end{aligned}$$

Para o cálculo de A_{22}, precisaremos da multipicação de $(A.5)$ por $a_{ij}m_{kr}m_{st}$, uma vez que a expressão de A_{22} é bastante parecida com a de A_{12}. Então:

$$\begin{aligned} \sum \left(\kappa_{ijk} + 2\kappa_{i,jk}\right)\kappa_{r,s,t}a_{ij}m_{kr}m_{st} &= \iota^T(-Q_1 + 2\tilde{W}_1)(Z - Z_2)_d Z_{2d}(Z - Z_2)(2Q_1 - 3\tilde{W}_1)\iota \\ &+ \iota^T W_1 E_2(Z - Z_2)(Z - Z_2)_d(2Q_1 - 3\tilde{W}_1)\iota. \qquad (A.18) \end{aligned}$$

Logo, usando o resultado em $(A.18)$, obtemos a seguinte expressão para A_{22}:

$$\begin{aligned} A_{22} &= 6\sum \left(\kappa_{ijk} + 2\kappa_{i,jk}\right)\kappa_{r,s,t}a_{ij}m_{kr}m_{st} \\ &= 6\iota^T(-Q_1 + 2\tilde{W}_1)(Z - Z_2)_d Z_{2d}(Z - Z_2)(2Q_1 - 3\tilde{W}_1)\iota \\ &+ 6\iota^T W_1 E_2(Z - Z_2)(Z - Z_2)_d(2Q_1 - 3\tilde{W}_1)\iota \qquad (A.19) \end{aligned}$$

O primeiro passo para obter a expressão de A_{23} é multiplicar $(A.15)$ por $a_{kt}m_{ir}m_{js}$, dada a seguir:

$$\sum \kappa_{i,j,k}.\kappa_{r,s,t}a_{kt}m_{ir}m_{js} = \iota^T(2Q_1 - 3\tilde{W}_1)(Z - Z_2)Z_2(Z - Z_2)(2Q_1 - 3\tilde{W}_1)\iota. \quad (A.20)$$

Assim, usando o resultado em $(A.20)$, temos que a expressão de A_{23} fica dada por:

$$\begin{aligned} A_{23} &= -6\sum \kappa_{i,j,k}.\kappa_{r,s,t}a_{kt}m_{ir}m_{js} \\ &= -6\iota^T(2Q_1 - 3\tilde{W}_1)(Z - Z_2)Z_2(Z - Z_2)(2Q_1 - 3\tilde{W}_1)\iota. \qquad (A.21) \end{aligned}$$

Para obter a expressão de A_{24}, basta multiplicar $\kappa_{i,j,k,r}$ por $m_{ij}m_{kr}$. Logo:

$$\sum \kappa_{i,j,k,r}m_{ij}m_{kr} = \iota^T(8Q_8 - 3Q_7 + 3Q_9 - 6\tilde{W}_1^*)(Z - Z_2)_d(Z - Z_2)_d\iota. \qquad (A.22)$$

Portanto, temos que a expressão de A_{24}, usando resultado em $(A.22)$, é dada por:

$$\begin{aligned} A_{24} &= 3\sum \kappa_{i,j,k,r}m_{ij}m_{kr} \\ &= 3\iota^T(8Q_8 - 3Q_7 + 3Q_9 - 6\tilde{W}_1^*)(Z - Z_2)_d(Z - Z_2)_d\iota. \qquad (A.23) \end{aligned}$$

Por fim, as expressões de A_{31} e A_{32} são simples de obter usando em ambos os casos o resultado em $(A.14)$. Para A_{31}, basta multiplicar $(A.16)$ por $m_{ij}m_{kr}m_{st}$, como descrito a seguir:

$$\begin{aligned} \sum \kappa_{i,j,k}.\kappa_{r,s,t}m_{ij}m_{kr}m_{st} &= 2\iota^T Q_1(Z - Z_2)_d(Z - Z_2)(Z - Z_2)_d(2Q_1 - 3W_1)\iota \\ &+ 3\iota^T W_1(Z - Z_2)_d(Z - Z_2)(Z - Z_2)_d(3W_1 - 2Q_1)\iota. \quad (A.24) \end{aligned}$$

Usando o resultado em $(A.24)$, temos que a seguinte expressão para A_{31}:

$$\begin{aligned} A_{31} &= 3\sum \kappa_{i,j,k}.\kappa_{r,s,t}m_{ij}m_{kr}m_{st} \\ &= 6\iota^T Q_1(Z-Z_2)_d(Z-Z_2)(Z-Z_2)_d(2Q_1-3W_1)\iota \\ &+ 9\iota^T W_1(Z-Z_2)_d(Z-Z_2)(Z-Z_2)_d(3W_1-2Q_1)\iota. \end{aligned} \tag{A.25}$$

E, para a expressão de A_{32}, basta multiplicar $(A.15)$ por $m_{ir}m_{js}m_{kt}$, como a seguir:

$$\begin{aligned} \sum \kappa_{i,j,k}.\kappa_{r,s,t}m_{ir}m_{js}m_{kt} &= 2\iota^T Q_1(Z-Z_2)(Z-Z_2)(Z-Z_2)(2Q_1-3\tilde{W}_1)\iota \\ &- 3\iota^T \tilde{W}_1(Z-Z_2)(Z-Z_2)(Z-Z_2)(2Q_1-3\tilde{W}_1)\iota. \end{aligned} \tag{A.26}$$

Daí, usando o resultado em $(A.26)$, obtemos a seguinte expressão para A_{32}:

$$\begin{aligned} A_{32} &= 2\sum \kappa_{i,j,k}.\kappa_{r,s,t}m_{ir}m_{js}m_{kt} \\ &= 4\iota^T Q_1(Z-Z_2)(Z-Z_2)(Z-Z_2)(2Q_1-3\tilde{W}_1)\iota \\ &- 6\iota^T \tilde{W}_1(Z-Z_2)(Z-Z_2)(Z-Z_2)(2Q_1-3\tilde{W}_1)\iota. \end{aligned} \tag{A.27}$$

Tendo em vista os resultados acima e a Seção 3.2.1, podemos obter as seguintes expressões para os $A's$:

$$\begin{aligned} A_1 &= A_{11}+A_{12}+A_{13}+A_{14} \\ &= (A.4)+(A.7)+(A.11)+(A.14) \\ &= 6\iota^T(2\tilde{W}_1-Q_1)Z_{2d}(Z-Z_2)Z_{2d}\tilde{W}_1\iota + 3\iota^T(Q_1-2\tilde{W}_1)Z_{2d}(Z-Z_2)Z_{2d}Q_1\iota \\ &+ 3\iota^T W_1E_2(Z-Z_2)E_2W_1\iota + 3\iota^T(2\tilde{W}_1-Q_1)Z_{2d}(Z-Z_2)E_2W_1\iota \\ &+ 3\iota^T W_1E_2Z_{2d}(Z-Z_2)(2\tilde{W}_1-Q_1)\iota \\ &- 6\iota^T(-Q_1+2\tilde{W}_1)Z_{2d}(Z-Z_2)_dZ_2(2Q_1-3\tilde{W}_1)\iota \\ &+ 36\iota^T(-Q_1+2\tilde{W}_1)Z_{2d}Z_2D_2W_1\iota - 6\iota^T W_1C_2(Z-Z_2)_d(2Q_1-3\tilde{W}_1)\iota \\ &+ 6\iota^T(-3Q_1+4\tilde{W}_1)Z_2(Z-Z_2)Z_2(-Q_1+2\tilde{W}_1)\iota \\ &+ 6\iota^T(-3Q_1+4\tilde{W}_1)C_2(Z-Z_2)W_1\iota + 30\iota^T W_1C_2(Z-Z_2)(-Q_1+2\tilde{W}_1)\iota \\ &+ 30\iota^T W_1J_2(Z-Z_2)W_1\iota + 6\iota^T(-2Q_9-6Q_8+2Q_7+5W_1^*)Z_{2d}(Z-Z_2)_d\iota \\ &+ 6\iota^T(7\tilde{W}_1-6Q_1-4Q_5)(Z-Z_2)_dE_2\iota, \end{aligned}$$

$$
\begin{aligned}
A_2 &= A_{21} + A_{22} + A_{23} + A_{24} \\
&= (A.17) + (A.19) + (A.21) + (A.23) \\
&= 6\iota^T W_1 E_2 (Z - Z_2)(Z - Z_2)_d (2Q_1 - 3\tilde{W}_1).\iota \\
&- 3\iota^T (2Q_1 - 3\tilde{W}_1)(Z - Z_2)_d Z_2 (Z - Z_2)_d (2Q_1 - 3\tilde{W}_1)\iota \\
&- 6\iota^T (2Q_1 - 3\tilde{W}_1)(Z - Z_2) Z_2 (Z - Z_2)(2Q_1 - 3\tilde{W}_1)\iota \\
&+ 3\iota^T (8Q_8 - 3Q_7 + 3Q_9 - 6\tilde{W}_1^*)(Z - Z_2)_d (Z - Z_2)_d \iota \\
&+ 6\iota^T (-Q_1 + 2\tilde{W}_1)(Z - Z_2)_d Z_{2d} (Z - Z_2)(2Q_1 - 3\tilde{W}_1)\iota \qquad \text{e}
\end{aligned}
$$

$$
\begin{aligned}
A_3 &= A_{31} + A_{32} \\
&= (A.25) + (A.27) \\
&= 3\iota^T (2Q_1 - 3\tilde{W}_1)(Z - Z_2)_d (Z - Z_2)(Z - Z_2)_d (2Q_1 - 3\tilde{W}_1)\iota \\
&+ 2\iota^T (2Q_1 - 3\tilde{W}_1)(Z - Z_2)^{(3)} (2Q_1 - 3\tilde{W}_1)\iota.
\end{aligned}
$$

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

Referências Bibliográficas

[1] Aho, J. (1978). Freshwater Snail Populations and Equilibrium-Theory of Island Biogeography .1. Case-Study in Southern Finland. Annales Zoologici Fennici, v. 15, n. 2, pp. 146-154.

[2] Barbour, C. D. and Brown, J. H. (1974) .Fish Species-Diversity in Lakes. American Naturalist, v. 108, n. 962, pp. 473-489.

[3] Bartlett, M. S.(1937). Properties of sufficiency and statistical tests. *Procedings of the Royal Society A*, **160**, 268-282.

[4] Braga, K. S. P. R. (2007). Aperfeiçoamento de testes de hipóteses para modelos não-lineares simétricos. Dissertação de Mestrado em Estatística. Universidade Federal de Pernambuco.

[5] Brito, C. C. R. (2009). Correção de Bartlett em modelos nãoŨlineares simétricos heteroscedásticos. Dissertação de Mestrado em Estatística. Universidade Federal de Pernambuco.

[6] Browne, R. A. (1981). Lakes as Islands - Biogeographic Distribution, Turnover Rates, and Species Composition in the Lakes of Central New-York. Journal of Biogeography, v. 8, n. 1, pp. 75-83.

[7] Cavalcanti, A. B. (2009). Aperfeiçoamento de métodos estatísticos em modelos de regressão da família exponencial. Tese de doutorado em Estatística. Instituto de Matemática e Estatística da USP.

[8] Cordeiro, G. M. (1982). *Improved Likelihood Ratio Statistics for Generalized Linear Models*. Ph.D. thesis, Imperial College of Science and Technology, University of London.

[9] Cordeiro, G. M. (1983). Improved Likelihood Ratio Statistics for Generalized Linear Models. *Journal of the Royal Statistical Society B*, **45**, 404-413.

[10] Cordeiro, G. M.; Ferrari, S. L. P. (1991). A modified score test statistic having chi-squared distribution to order n^{-1}. *Biometrika* **78**, 573-582.

[11] Cordeiro, G. M. (1987). On the corrections to the likelihood ratio statistic. *Biometrika*, **74** ,265-274.

[12] Cordeiro, G. M., Ferrari, S. L. P.; Paula, G. A. (1993). Improved score tests for generalized linear models. *Journal of the Royal Statistical Society B* **55**, 661-674.

[13] Cordeiro, G. M.; Ferrari, S.L.P.; Cysneiros, A.H.M.A. (1998). A formula to improve score test statistics. *Journal of Statistical Computatio*, **62**, 123-136.

[14] Cordeiro, G. M. (1999). Introdução à Teoria Assintótica. ***IMPA*** Instituto de Matemática Pura e Aplicada, Rio de Janeiro.

[15] Cordeiro, G. M.; Cysneiros, H. M. A.; Cysneiros, F. J. A. (2006), Bartlett Adjustments for Overdispersed Generalized Linear Models. *Communications in Statistics- Theory and Methods*, **35**, 937-952.

[16] Cordeiro, G. M.; Andrade, M. G.; de Castro, M. (2009). Power series generalized nonlinear models.*Computational Statistics and Data Analysis* **53**,1155-1166.

[17] Cox, D. R.; Snell, E. (1968). A general definition of residuals. *Journal of the Royal Statistical Society B*, **30**, 248-275.

[18] Cox, D. R.; Hinkley, D. V. (1974). Theoretical Statistics. Chapman & Hall, London.

[19] Cribari-Neto, F.; Cordeiro, G. M. (1997). On Bartlett and Bartlett-type corrections. *Econometric Reviews* **15**, 339-367.

[20] Cysneiros, A. H. M. A. (1997). Correções de Bartlett e tipo-Bartlett em modelos lineares generalzados. Dissertação de Mestrado em Estatística. Universidade de São Paulo.

[21] 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**, n. 3, p. 255-265.

[22] Doornik, J. A. (2006). Ox: *An Object-Oriented Matrix Language.* Timberlake Consultants Press, London; Oxford, http://www.doornik.com. 5th ed.

[23] Ferrari, S.L.P. ; Uribe-Opazo, M. A. (2001). Corrected likelihood ratio tests in a class of symmetric linear regression models. *Brazilian Journal of Probability and Statistics*, São Paulo - Brasil, **15**, 49-67.

[24] Frydenberg, M.; Jensen, J.L. (1989). Is the “improved likelihood statistic"really improved in discrete case? *Biometrika* **76**, 655-661.

[25] Harris, P. (1985). An asymptotic expansion for the null distribution of the eficient score statistic. *Biometrika* **72**, 653-659.

[26] Hayakawa, T. (1977). The likelihood ratio criterion and the asymptotic expansion of its distribution. *Annals of the Institute of Statistical Mathematics A* **29**, 359-378.

[27] Hayakawa, T.; Puri, M. L. (1985). Asymptotic expansions of the distributions of some test statistics. *Annals of the Institute of Statistical Mathematics A*, **62**, 451460.

[28] Kakizawa, Y. (1996), Higher order monotone Bartlett-type adjustment for some multivariate test statistic, *Biometrika* **83**, 923-659.

[29] Lassen, H. H. (1975) .Diversity of Freshwater Snails in View of Equilibrium Theory of Island Biogeography. Oecologia, v. 19, n. 1, pp. 1-8.

[30] Lemonte, A. J. (2009). Estatística gradiente e refinamento de métodos assintóticos no modelo de regressão Birnbaum-Saunders. Tese de Doutorado, Programa de pós-graduação em Estatística, Universidade de São Paulo.

[31] Lemonte, A. J.; Ferrari, S.L.P. (2009). Small-Sample Corrections for Score Tests in Birnbaum-Saunders Regressions. *Communications in Statistics.* Theory and Methods, **40**, 232-243

[32] Nascimento, K. P. (2010).Correção tipo-Bartlett em modelos não-lineares simétricos heteroscedásticos. Dissertação de Mestrado em Estatística. Universidade Federal de Pernambuco.

[33] Nocedal, J; Wright, S. T. (1999). *Numerical Optimization.* New York: Springer.

[34] Rao, C. R. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society* **44** ,40-57.

[35] Rao, C. R. (1973). *Linear Statistical Inference and its Applications.* 2nd ed. New York, John Wiley. 625p. (Wiley Series in Probability and Mathematical Statistics).

[36] Sepkoski, J.J. Jr. Rex, M. A. (1974). Distribution of freshwater mussels: coastal rivers as biogeographic islands. Syst. Zool. 23: 165-168.

[37] Silva, P. G. (2010). Correção de viés e de Bartlett em modelos em série de potência não-lineares generalizados. Dissertação de Mestrado em Estatística. Universidade Federal de Pernambuco.

[38] Terrell, G. R. (2002). The Gradient Statistic. *Computing Science and Statistics*, **34**, 206-215.

[39] Uribe-Opazo, M. A. (1997). Aperfeiçoamento de Testes Estatísticos em várias Famílias de distribuição . Tese de Doutorado, Programa de pós-graduação em Estatística, Universidade de São Paulo.

[40] Uribe-Opazo, M. A; Ferrari, S. L. P.; Cordeiro, G. M. (2008). Improved score tests in symmetric linear regression models. *Communications in Statistics.* Theory and Methods, **37**, 261-276.