



Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza
Departamento de Estatística
Pós-graduação em Estatística

**INFERÊNCIA E DIAGNÓSTICO EM
MODELOS NÃO LINEARES
LOG-GAMA GENERALIZADOS**

Priscila Gonçalves da Silva

Tese de Doutorado

Recife
2016

PRISCILA GONÇALVES DA SILVA

**INFERÊNCIA E DIAGNÓSTICO EM MODELOS NÃO LINEARES LOG-GAMA
GENERALIZADOS**

Tese apresentada ao Programa de Pós-graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Doutor em Estatística.

Orientadora:

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

Recife
2016

Catálogo na fonte
Bibliotecário Jefferson Luiz Alves Nazareno CRB 4-1758

S586i Silva, Priscila Gonçalves da.
Inferência e diagnóstico em modelos não lineares Log-Gama generalizados / Priscila Gonçalves da Silva. – 2016.
145f.: fig., tab.

Orientadora: Audrey Helen Mariz de Aquino Cysneiros.
Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN. Estatística, Recife, 2016.
Inclui referências e apêndice.

1. Estatística matemática. 2. Teoria assintótica. 3. Análises de diagnóstico. I. Cysneiros, Audrey Helen Mariz de Aquino. (Orientadora). II. Título.

310 CDD (22. ed.) UFPE-MEI 2016-163

PRISCILA GONÇALVES DA SILVA

**INFERÊNCIA E DIAGNÓSTICO EM MODELOS NÃO LINEARES LOG-GAMA
GENERALIZADOS**

Tese apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Estatística.

Aprovada em: 04 de novembro de 2016.

BANCA EXAMINADORA

Prof.^a Audrey Helen Mariz de Aquino Cysneiros
UFPE

Prof. Francisco José de Azevedo Cysneiros
UFPE

Prof. Getúlio José Amorim do Amaral
UFPE

Prof. Vinicius Quintas Souto Maior
UFPE

Prof.^a Lourdes Coral Contreras Montenegro
UFMG

A Deus, o Autor da vida, e a minha família, o meu alicerce.

Agradecimentos

- Primeiramente ao meu Deus, por ter me proporcionado mais esta conquista.
- À professora Audrey Helen Mariz de Aquino Cysneiros, pela confiança e por sua sublime orientação com que direcionou esse trabalho.
- Aos meus pais, Josabete e Natanael, pela educação que a mim foi dada, pelo apoio e compreensão.
- A toda minha família pelo apoio e incentivo, em especial a Patricia, Thiago, Jenilson, Maciana, João, Clarice e Vó Amara.
- Aos meus colegas de trabalho pelo apoio dado durante o curso, em especial a Jansen Dodô, Mayara Leon, Normando Vieira e Tânia Maria.
- Aos meus colegas de Doutorado por compartilharmos momentos de dificuldades e superação.
- A todos os professores e funcionários do Departamento de Estatística da UFPE por seus trabalhos realizados, em especial à professora Cristina Raposo pelo seu incentivo.
- Ao professor Francisco Cysneiros pelas sugestões.

Resumo

Young e Bakir (1987) propôs a classe de Modelos Lineares Log-Gama Generalizados (MLLGG) para analisar dados de sobrevivência. No nosso trabalho, estendemos a classe de modelos proposta por Young e Bakir (1987) permitindo uma estrutura não linear para os parâmetros de regressão. A nova classe de modelos é denominada como Modelos Não Lineares Log-Gama Generalizados (MNLLGG). Com o objetivo de obter a correção de viés de segunda ordem dos estimadores de máxima verossimilhança (EMV) na classe dos MNLLGG, desenvolvemos uma expressão matricial fechada para o estimador de viés de Cox e Snell (1968). Analisamos, via simulação de Monte Carlo, os desempenhos dos EMV e suas versões corrigidas via Cox e Snell (1968) e através da metodologia bootstrap (Efron, 1979). Propomos também resíduos e técnicas de diagnóstico para os MNLLGG, tais como: alavancagem generalizada, influência local e influência global. Obtivemos, em forma matricial, uma expressão para o fator de correção de Bartlett à estatística da razão de verossimilhanças nesta classe de modelos e desenvolvemos estudos de simulação para avaliar e comparar numericamente o desempenho dos testes da razão de verossimilhanças e suas versões corrigidas em relação ao tamanho e poder em amostras finitas. Além disso, derivamos expressões matriciais para os fatores de correção tipo-Bartlett às estatísticas escore e gradiente. Estudos de simulação foram feitos para avaliar o desempenho dos testes escore, gradiente e suas versões corrigidas no que tange ao tamanho e poder em amostras finitas.

Palavras-chave: Bootstrap. Correção de Bartlett. Correção de viés. Correção tipo-Bartlett. Distribuição log-gama generalizada. Modelo não linear. Resíduos. Técnicas de diagnóstico. Teste da razão de verossimilhanças. Teste escore. Teste gradiente.

Abstract

Young e Bakir (1987) proposed the class of generalized log-gamma linear regression models (GLGLM) to analyze survival data. In our work, we extended the class of models proposed by Young e Bakir (1987) considering a nonlinear structure for the regression parameters. The new class of models is called generalized log-gamma nonlinear regression models (GLGNLM). We also propose matrix formula for the second-order bias of the maximum likelihood estimate of the regression parameter vector in the GLGNLM class. We use the results by Cox and Snell (1968) and bootstrap technique [Efron (1979)] to obtain the bias-corrected maximum likelihood estimate. Residuals and diagnostic techniques were proposed for the GLGNLM, such as generalized leverage, local and global influence. An general matrix notation was obtained for the Bartlett correction factor to the likelihood ratio statistic in this class of models. Simulation studies were developed to evaluate and compare numerically the performance of likelihood ratio tests and their corrected versions regarding size and power in finite samples. Furthermore, general matrix expressions were obtained for the Bartlett-type correction factor for the score and gradient statistics. Simulation studies were conducted to evaluate the performance of the score and gradient tests with their corrected versions regarding to the size and power in finite samples.

Keywords: Bartlett correction. Bartlett-type correction. Bias correction. Bootstrap. Diagnostic techniques. Generalized log-gamma distribution. Gradient test. Likelihood ratio test. Nonlinear model. Residual. Score test.

Lista de Figuras

3.1	Estimação intevarlar dos parâmetros do MNLLGG quando $\beta_0 = \beta_1 = \beta_2 = 1$, $\theta = 1$ e $\lambda = 1$ para $n = 50$.	39
3.2	Valores observados versus valores ajustados a partir dos Modelos I, II e III para os dados em estudo.	42
4.1	Curva de poder dos testes para o MLLGG com $\theta = 1$, $\lambda = 1$, $p = 4$, $n = 20$.	69
4.2	Discrepância relativa de quantis para o MLLGG com $\theta = 1$, $\lambda = 0,5$, $p = 4$, $n = 20$.	70
5.1	Histogramas das estimativas do resíduo de Pearson.	86
5.2	Histogramas das estimativas do resíduo padronizado.	86
5.3	Histogramas das estimativas do resíduo componente do desvio padronizado.	87
5.4	Gráfico dos resíduos versus o índice de observações e envelope para o Modelo III.	99
5.5	Gráfico de alavancagem generalizada para o Modelo III.	100
5.6	Gráfico de C_i contra o índice de observações segundo os esquemas de perturbação (a) aditiva na resposta (b) casos ponderados e (c) aditiva no preditor para o Modelo III.	101
5.7	Gráfico de Influência Global para o Modelo III.	102

Lista de Tabelas

3.1	Comparação dos três estimadores para o MNLLGG quando $\beta_0 = \beta_1 = 1$, $\beta_2 = 1,25$, $\theta = 1$ e $\lambda = 1$ para $n = 20, 50$ e 100 .	35
3.2	Comparação dos três estimadores para o MNLLGG quando $\beta_0 = \beta_1 = 1$, $\beta_2 = 1; 1,5; 2$, $\theta = 1$, $\lambda = 1$ e $n = 30$.	36
3.3	Estimação intevarlar dos parâmetros do MNLLGG quando $\beta_0 = \beta_1 = 1$, $\beta_2 = 1,25$, $\theta = 1$, $\lambda = 1$, $n = 20$ e $(1 - \alpha)\% = 90\%, 95\%, 99\%$.	37
3.4	Estimação intevarlar dos parâmetros do MNLLGG, ao nível de confiança de 95% quando $\beta_0 = \beta_1 = 1$, $\beta_2 = 1,25$, $\theta = 1$ e $\lambda = 1$ para $n = 20, 50$ e 100 .	38
3.5	Resultado da estimação pontual e intervalar dos parâmetros dos Modelos III para os dados em estudo.	41
4.1	Taxas de rejeição dos testes $LR, LR^*, LR_1^*, LR_{boot}, LR_{boot}^*, S_r, S_r^*, K(S_r), \tilde{S}_r, S_g$ e S_g^* em MLLGG com $\theta = 1$, $n = 30$, $p = 8$ e diversos valores de λ .	65
4.2	Taxas de rejeição dos testes $LR, LR^*, LR_1^*, LR_{boot}, LR_{boot}^*, S_r, S_r^*, K(S_r), \tilde{S}_r, S_g$ e S_g^* em MLLGG com $\theta = 1$, $p = 8$, $\lambda = 1$ e diversos valores de n .	66
4.3	Taxas de rejeição dos testes $LR, LR^*, LR_1^*, LR_{boot}, LR_{boot}^*, S_r, S_r^*, K(S_r), \tilde{S}_r, S_g$ e S_g^* em MLLGG com $\theta = 1$, $n = 30$, $\lambda = 1$ e diversos valores de p .	67
4.4	Poder dos testes $LR, LR^*, LR_1^*, LR_{boot}, LR_{boot}^*, S_r, S_r^*, K(S_r), \tilde{S}_r, S_g$ e S_g^* , em MLLGG com $\theta = 1$, $\lambda = 1$, $p = 4$, $n = 20$.	68
4.5	Média e variância da distribuição χ_2^2 e das estatísticas $LR, LR^*, LR_1^*, LR_{boot}^*, S_r, S_r^*, K(S_r), \tilde{S}_r, S_g$ e S_g^* em MLLGG com $p = 4$, $n = 20$ e $\lambda = 0,5$.	70
4.6	Quantis da distribuição χ_2^2 e das estatísticas $LR, LR^*, LR_1^*, LR_{boot}^*, S_r, S_r^*, K(S_r), \tilde{S}_r, S_g$ e S_g^* em MLLGG com $p = 4$, $n = 20$ e $\lambda = 0,5$.	70
4.7	Taxas de rejeição dos testes $LR, LR^*, LR_1^*, LR_{boot}^*, LR_{boot}, S_r, S_r^*, K(S_r)$ e \tilde{S}_r em MNLLGG com $\theta = 1$, $p = 5$, $n = 20$ e diversos valores de λ .	72
4.8	Taxas de rejeição dos testes $LR, LR^*, LR_1^*, LR_{boot}^*, LR_{boot}, S_r, S_r^*, K(S_r)$ e \tilde{S}_r em MNLLGG com $\theta = 1$, $p = 5$, $\lambda = 1$ e diversos valores de n .	72
4.9	Taxas de rejeição dos testes $LR, LR^*, LR_1^*, LR_{boot}^*, LR_{boot}, S_r, S_r^*, K(S_r)$ e \tilde{S}_r em MNLLGG com $\lambda = 1$, $\theta = 1$, $n = 20$ e diversos valores de p .	73
4.10	Poder dos testes $LR, LR^*, LR_1^*, LR_{boot}^*, LR_{boot}, S_r, S_r^*, K(S_r)$ e \tilde{S}_r em MNLLGG com $\theta = 1$, $\lambda = 1$, $p = 4$, $n = 20$.	74
4.11	Média e variância da distribuição χ_2^2 e das estatísticas $LR, LR^*, LR_1^*, LR_{boot}^*, LR_{boot}, S_r, S_r^*, K(S_r)$ e \tilde{S}_r em MNLLGG com $p = 4$, $n = 20$ e $\lambda = 0,5$.	75

4.12	Quantis da distribuição χ_2^2 e das estatísticas $LR, LR^*, LR_1^*, LR_{boot}^*, LR_{boot}, S_r, S_r^*, K(S_r)$ e \tilde{S}_r em MNLLGG com $p = 4, n = 20$ e $\lambda = 0,5$.	75
4.13	P-valores dos testes de hipóteses no modelo 34.	77
5.1	Medidas descritivas das estimativas do resíduo de Pearson no MNLLGG.	84
5.2	Medidas descritivas das estimativas do resíduo padronizado no MNLLGG.	85
5.3	Medidas descritivas das estimativas do resíduo componente do desvio padronizado no MNLLGG.	85
5.4	Análise de sensibilidade ao remover as observações influentes no Modelo III.	100

Sumário

1	Introdução	13
2	Modelo Não Linear Log-Gama Generalizado	17
2.1	Definição	19
2.2	Aspectos Inferenciais	20
3	Correção de Viés em Modelo Não Linear Log-Gama Generalizado	24
3.1	Introdução	24
3.2	Correção de Cox e Snell	25
3.3	Correção por Bootstrap	30
3.4	Estimação Intervalar	31
3.5	Resultados Numéricos	33
3.6	Aplicação	40
3.7	Comentários	42
4	Correção de Bartlett e Tipo-Bartlett em Modelo Não Linear Log-Gama Generalizado	44
4.1	Introdução	44
4.2	Testes de Hipóteses	48
4.3	Correção de Bartlett	49
4.3.1	Correção de Bartlett em MNLLGG com θ fixo	51
4.4	Correção de Bartlett Bootstrap	55
4.5	Correção Tipo-Bartlett	56
4.5.1	Correção Tipo-Bartlett à Estatística Escore	56
4.5.2	Correção Tipo-Bartlett à Estatística Gradiente	59
4.5.3	Correção Tipo-Bartlett em MNLLGG com θ fixo	61
4.6	Resultados Numéricos	63
4.6.1	Modelos Lineares	64
4.6.2	Modelos Não-Lineares	71
4.7	Aplicações	75
4.7.1	Exemplo 1	75
4.7.2	Exemplo 2	76
4.8	Comentários	78

5	Técnicas de Diagnóstico no Modelo Não Linear Log-Gama Generalizado	80
5.1	Introdução	80
5.2	Análise de Resíduos	81
5.3	Análise de Diagnóstico	88
5.3.1	Alavancagem Generalizada	88
5.3.2	Influência Local	89
	Perturbação Aditiva na Resposta	91
	Perturbação de Casos Ponderados	93
	Perturbação Aditiva no Preditor	94
5.3.3	Influência Global	95
5.4	Aplicação	97
5.5	Comentários	102
6	Considerações Finais	104
	Referências Bibliográficas	106
	Apêndice Apêndice	113
	Apêndice A Cálculo dos Momentos	113
A.1	Derivadas da Função de Log-Verossimilhança	113
A.2	Cálculo dos Cumulantes	114
A.3	Derivadas dos Cumulantes	116
	Apêndice B Cálculo dos Termos de $B(\hat{\xi})$	118
B.1	Cálculo dos Termos de $B(\hat{\beta}_b)$	118
	Apêndice C Cálculo do Fator de Correção de Bartlett	121
C.1	Cálculo da Primeira Parcela $\sum \lambda_{rstu}$	122
C.2	Cálculo da Segunda Parcela $\sum \lambda_{rstuvw}$	124
	Apêndice D Cálculo do Fator de Correção Tipo-Bartlett	130
D.1	Correção Tipo-Bartlett à Estatística Escore	130
D.2	Correção Tipo-Bartlett à Estatística Gradiente	135
	Apêndice E Influência Global	139
	Apêndice F Material Suplementar	141
F.1	Fator de Correção de Bartlett em MNLLGG com θ não fixo	141
F.2	Fator de Correção tipo-Bartlett à Estatística Gradiente em MNLLGG	143

1

Introdução

A eficácia de um modelo de regressão paramétrico depende da escolha correta da distribuição da variável aleatória dependente. Diferentes modelos de regressão foram propostos na literatura para análise de dados de sobrevivência, tais como aqueles baseados nas distribuições gama, log-normal, Weibull e valores extremos. Exceto para a distribuição valor extremo, as outras distribuições normalmente fornecem ajustes satisfatórios na parte central dos dados, mas muitas vezes não conseguem ter um bom ajuste nas caudas. Isto porque a quantidade de dados relevantes para inferência na cauda da distribuição pode ser limitada, mesmo quando grandes quantidades de dados estão disponíveis para inferência na parte central da distribuição. A classe dos modelos de regressão baseados na distribuição log-gama generalizada (Lawless, 1980) têm sido amplamente aplicada nas áreas de análise de confiabilidade e sobrevivência, inclui como casos especiais o modelo normal, já bastante conhecido, e o modelo de regressão valor extremo tipo I, que é comumente utilizado para modelagem de dados referentes à terremoto, inundação, entre outros desastres naturais. Neste trabalho, propomos a classe de modelos de regressão não lineares log-gama generalizados (MNLLGG) que abrange o modelo de regressão linear log-gama generalizado (MLLGG) apresentado por Young e Bakir (1987).

A estimação dos parâmetros nos MNLLGG é feita pelo método da máxima verossimilhança, que por sua vez fornece, em geral, estimadores viesados. Em alguns casos, o viés

é considerado insignificante quando comparado ao erro padrão dos estimadores de máxima verossimilhança (EMV), visto que ele é de ordem n^{-1} , enquanto o desvio padrão da estimativa é de ordem $n^{-1/2}$. Porém, no caso de modelos não lineares quando o tamanho da amostra é pequeno ou a informação de Fisher é reduzida, o viés passa a ter uma magnitude comparável ao erro padrão do EMV (Cordeiro, 1999). Deste modo, é de suma importância o cálculo do viés de segunda ordem dos EMV a fim de obtermos estimadores mais precisos.

Para os testes de hipóteses sobre os parâmetros dos MNLLGG, podemos utilizar as estatísticas da razão de verossimilhanças (LR), escore (S_r) e gradiente (S_g). As distribuições exatas dessas estatísticas são geralmente desconhecidas. Mas, em problemas regulares e em grandes amostras, as três estatísticas possuem distribuições que podem ser aproximadas, sob a hipótese nula (H_0), pela distribuição qui-quadrado (χ^2). Para a estatística LR o erro dessa aproximação é de ordem n^{-1} , enquanto para as estatísticas S_r e S_g esse erro é de ordem $n^{-1/2}$. Encontramos vários estudos na literatura cujo objetivo é obter ajustes para as estatísticas LR , S_r e S_g que reduzam esses erros de aproximação. A idéia é modificar essas estatísticas por um fator de correção. No caso da estatística LR , Bartlett (1937) propôs um fator de correção produzindo uma estatística modificada com o primeiro momento igual ao da distribuição χ^2 de referência, sob H_0 . Por outro lado, Cordeiro e Ferrari (1991) e Vargas *et al.* (2013) propuseram fatores de correções tipo-Bartlett para as estatísticas S_r e S_g , respectivamente, produzindo novas estatísticas com distribuição χ^2 até erro de ordem n^{-1} , sob H_0 .

A precisão das inferências realizadas, no entanto, dependerá da qualidade do modelo. Por isso, é necessário verificar se o modelo ajustado se adequa bem aos dados e se existem observações discrepantes com alguma influência desproporcional ou inferencial nos resultados do ajuste. Esta análise é feita através dos resíduos produzidos pelo modelo, assumindo o modelo como correto e investigando se as conclusões feitas são coerentes. Além disso, é importante verificar a existência de observações influentes, isto é, pontos que exercem uma influência desproporcional nas estimativas dos parâmetros do modelo ao serem omitidos no ajuste do mesmo ou submetidos a uma pequena perturbação.

Diante do exposto, neste trabalho destacamos dois aspectos inferenciais no MNLLGG usando a teoria assintótica: o primeiro corresponde à obtenção da expressão do viés de segunda ordem dos EMV dos parâmetros do modelo e o segundo visa à obtenção de ajustes para as estatísticas LR , S_r e S_g . Sobre estes temas, encontramos na literatura, os mais recentes trabalhos de Cordeiro e Cribari-Neto (2014) e de Lemonte (2016). Adicionalmente, desenvolvemos técnicas de análise de resíduo e diagnóstico para o MNLLGG. Nossos resultados não se aplicam para observações censuradas.

O trabalho é desenvolvido como segue. No Capítulo 2, introduzimos o modelo não linear log-gama generalizado, bem como seus aspectos inferenciais, considerando o parâmetro de forma λ^2 fixo. No Capítulo 3, desenvolvemos uma expressão, em forma fechada, para o viés de segunda ordem de Cox e Snell (1968) dos EMV dos parâmetros do MNLLGG. Além disso, discorremos sobre a obtenção da correção de viés via a metodologia bootstrap (Efron, 1979). Resultados numéricos sobre o desempenho dos EMV, bem como das suas versões corrigidas, em amostras de tamanho finito e aplicações a dados reais também são apresentados neste capítulo. No Capítulo 4 apresentamos a obtenção, em notação matricial, do fator de correção de Bartlett para a estatística LR e do fator de correção tipo-Bartlett para as estatísticas S_r e S_g na classe dos MNLLGG. Estudos de simulação via Monte Carlo são apresentados com o objetivo de comparar o desempenho dos testes baseados nas estatísticas LR , S_r e S_g com suas respectivas versões corrigidas em relação ao tamanho e poder em amostras finitas. No Capítulo 5 são desenvolvidas técnicas de diagnóstico para os MNLLGG baseadas nas medidas de alavancagem generalizada, influência local e influência global. Finalmente, no Capítulo 6, encontram-se as considerações finais.

Os resultados numéricos foram obtidos utilizando a versão 4.10 da linguagem matricial de programação $\text{O}\times$ para sistema operacional Windows. Esta linguagem foi criada por Jurgen Doornik, em 1994, na Universidade de Oxford (Inglaterra). Ela é muito flexível com sintaxe similar às sintaxes das linguagens de programação C e C++. Mais detalhes sobre esta linguagem de programação podem ser encontrados em Doornik (2007) e em Cribari-Neto e Zarkos

(2003). As apresentações gráficas foram produzidas no ambiente de programação R, tendo sido utilizada a versão 3.1.1 para a plataforma Windows. O R é um ambiente integrado que possui grandes facilidades para manipulação de dados, geração de gráficos e modelagem estatística em geral (vide Cribari-Neto e Zarkos, 1999; Ihaka e Gentleman 1996; Venables e Ripley, 2002; Rizzo, 2008; Fox e Weisberg, 2011; Mello e Peternelli, 2013).

2

Modelo Não Linear Log-Gama Generalizado

Dados de sobrevivência são utilizados em análises em que o interesse está no tempo de ocorrência de um determinado evento. Nas indústrias, o interesse pode estar no tempo médio de vida de um produto ou no tempo de falha associados aos materiais expostos à fadiga, um dano estrutural que ocorre quando um material é exposto a flutuações de estresse e tensão. Na área médica o tempo entre o diagnóstico de uma doença e o falecimento ou a cura do paciente, assim como ao tempo entre a remissão (paciente livre de sintomas) e a recidiva da doença são de suma importância para o tratamento da mesma. Já na área ambiental, o tempo de ocorrências de alguns fenômenos, tais como temperatura mínima, chuvas em período de seca, entre outros, é de suma importância na prevenção de catástrofes ambientais.

Na literatura, encontramos várias distribuições utilizadas para modelar a taxa de ocorrência do evento, comumente chamada de taxa de falha. A distribuição gama generalizada (GG), proposta por Stacy (1962) abrange algumas dessas distribuições tais como a exponencial, Weibull, gama e a log-normal. No entanto, dificuldades em realizar inferência para a distribuição GG, levou Prentice (1974) a propor uma reparametrização de forma a obter a distribuição log-gama generalizada (LGG). Uma nova reparametrização apresentada por Lawless (1980) é definida como segue. Seja W uma variável aleatória gama com parâmetro escalar unitário e parâmetro de forma λ^2 . A variável aleatória transformada $Z = \lambda[\log(W) - 2\log(\lambda)]$ tem distribuição

LGG com função de densidade (Lawless, 2003, p. 44)

$$f(z; \lambda) = \frac{\lambda^{2\lambda^2-1}}{\Gamma(\lambda^2)} \exp[\lambda z + \lambda^2 c(\lambda) - \lambda^2 \exp\{\lambda^{-1}z + c(\lambda)\}], z \in \mathbb{R}, \quad (1)$$

em que $c(\lambda) = \psi(\lambda^2) - \log(\lambda^2)$ com $\psi(p) = d \log \Gamma(p) / dp$ sendo a função digama. A distribuição LGG inclui casos especiais, como a distribuição normal quando $\lambda \rightarrow \infty$ e a distribuição valor extremo tipo I quando $\lambda = 1$. Temos que $E(Z) = 0$ e $Var(Z) = \lambda^2 \psi^{(1)}(\lambda^2)$, em que $\psi^{(1)}(\cdot)$ denota a primeira derivada da função digama. Neste trabalho, assumimos que o parâmetro de forma λ^2 é uma constante positiva fixa.

O modelo de regressão linear que considera a distribuição LGG para a variável resposta (modelo linear log-gama generalizado - MLLGG) surge, então, como membro da família dos modelos de locação e escala para o logaritmo do tempo de falha apresentada por Lawless (1980). Esta família de modelos, também conhecida como modelos de tempo de vida acelerado, analisa o efeito de uma ou mais covariáveis associadas ao tempo de falha e ainda inclui os modelos de regressão normal, log-normal e gama generalizada. O nosso trabalho é baseado na parametrização utilizada por Young e Bakir (1987), os quais propuseram um estimador de máxima verossimilhança (EMV) corrigido para os parâmetros de regressão e o parâmetro escalar do MLLGG. No entanto, vários estudos são encontrados na literatura baseados em outras parametrizações da distribuição LGG, como por exemplo, DiCiccio (1987), o qual obteve estimativas aproximadas para os quantis e o parâmetro de escala da distribuição LGG; Ortega, Bolfarine e Paula (2003), que consideraram um estudo de influência local em modelo log-gama generalizado estendido com observações censuradas; Ortega, Cancho e Paula (2008) analisaram aspectos de inferência estatística e modelagem na presença de dados censurados neste mesmo modelo; Ortega *et al.* (2012) apresentaram uma nova extensão da LGG, a distribuição log-gama generalizada exponenciada; Hashimoto *et al.* (2013) propuseram um modelo de regressão de locação e escala baseada na distribuição LGG para a modelagem de dados com censura intervalar; Agostinelli, Marazzi e Yohai (2014) propuseram estimadores robustos no

modelo log-gama generalizado estendido.

Uma das particularidades da família LGG é a distribuição Valor-Extremo, amplamente utilizada na modelagem de valores extremos de fenômenos da natureza. Barreto-Souza e Vasconcellos (2011) introduziram o modelo de regressão não linear valor-extremo e derivaram uma expressão de viés de segunda-ordem de Cox e Snell (1968) dos EMV dos parâmetros deste modelo. Vale salientar que Barreto-Souza e Vasconcellos (2011) assumiram uma estrutura de regressão não linear tanto para parâmetro de locação quanto para o parâmetro de escala.

2.1 Definição

Considere que Y_1, \dots, Y_n sejam variáveis independentes. Introduzimos o modelo não linear log-gama generalizado (MNLLGG) definido por

$$Y_i = \eta(x_i, \beta) + \theta Z_i, \quad i = 1, \dots, n, \quad (2)$$

em que os erros Z_1, \dots, Z_n são variáveis aleatórias independentes e identicamente distribuídas com função densidade da LGG dada em (1), $\beta = (\beta_1, \dots, \beta_p)^\top$ é um vetor de p ($p < n$) parâmetros desconhecidos a serem estimados, $x_i = (x_{i1}, \dots, x_{ik})^\top$ representa os valores de k variáveis explicativas. Assumimos que β é definido no subconjunto Ω_β de \mathbb{R}^p e $\eta(\cdot; \cdot)$ é uma função possivelmente não linear no segundo argumento, contínua e diferenciável com respeito aos componentes de β tal que a matriz de derivadas $\tilde{X} = \tilde{X}(\beta) = \partial \eta / \partial \beta^\top$, com $\eta = (\eta_1, \dots, \eta_n)^\top$, tem posto p para todo β . O parâmetro de escala θ , assim como β , é considerado desconhecido. Quando $\eta(x_i, \beta) = x_i^\top \beta$ temos o MLLGG de Young e Bakir (1987).

Consideremos um MNLLGG definido por (1) e (2) e o vetor $(p+1)$ de parâmetros $\xi = (\beta, \theta)^\top$. Quando $\lambda = 1$, o MNLLGG é equivalente ao modelo de regressão não linear valor-extremo proposto por Barreto-Souza e Vasconcellos (2011). Inferência para o modelo de regressão proposto pode ser realizado com base nos EMV do vetor de parâmetros ξ . Dado o

vetor de observações $(y_1, \dots, y_n)^\top$, o logaritmo da função de verossimilhança (função de log-verossimilhança) dos parâmetros para ξ é dado por:

$$l(\xi) = (2\lambda^2 - 1)n \log \lambda - n \log \Gamma(\lambda^2) + \lambda \sum_{i=1}^n z_i + n\lambda^2 c(\lambda) - \lambda^2 \sum_{i=1}^n \exp\left\{\frac{z_i}{\lambda} + c(\lambda)\right\} - n \log \theta. \quad (3)$$

Assumimos que a função de log-verossimilhança satisfaz as condições de regularidades (Cox e Hinkley, 1974, Capítulo 9).

2.2 Aspectos Inferenciais

Definimos a matriz estendida \tilde{X}^* , de dimensão $(n+1) \times (p+1)$, da forma

$$\tilde{X}^* = \begin{pmatrix} \tilde{X} & \mathbf{0}_1 \\ \mathbf{0}_2^\top & 1 \end{pmatrix},$$

em que $\mathbf{0}_1$ e $\mathbf{0}_2$ são vetores de zeros de tamanhos n e p , respectivamente. Dessa forma, a função escore para o vetor de parâmetros ξ , condicionando em λ , é dada por:

$$U_\xi = \frac{\partial l(\xi)}{\partial \xi} = -\frac{\lambda}{\theta} \tilde{X}^{*\top} M \mathbf{1},$$

em que $\mathbf{1}$ é vetor de tamanho n de uns e $M = (M_1, m_2^\top)^\top$ é uma matriz de dimensão $(n+1) \times n$, com $M_1 = \text{diag}\{m_{11}, \dots, m_{1n}\}$ sendo uma matriz diagonal de dimensão $n \times n$ cujo i -ésimo elemento é $m_{1i} = 1 - \exp\{\lambda^{-1}z_i + c(\lambda)\}$ e m_2 um vetor de tamanho n , com $m_{2i} = m_{1i}z_i + 1/\lambda$, $i = 1, \dots, n$.

Seja $-\check{L}_{\xi\xi}$ a matriz de informação observada de Fisher para $\xi = (\beta^\top, \theta)^\top$, temos que

$$\check{L}_{\xi\xi} = \frac{\partial^2 l(\xi)}{\partial \xi \partial \xi^\top} = \begin{pmatrix} \check{L}_{\beta\beta} & \check{L}_{\beta\theta} \\ \check{L}_{\theta\beta} & \check{L}_{\theta\theta} \end{pmatrix} \quad (4)$$

sendo

$$\begin{aligned}\ddot{L}_{\beta\beta} &= -\frac{1}{\theta^2}\tilde{X}^\top(I-M_1)\tilde{X} - \frac{\lambda}{\theta^2}[M_1][X^{**}], \\ \ddot{L}_{\beta\theta} &= -\frac{1}{\theta^2}\tilde{X}^\top(I-M_1)z + \frac{\lambda}{\theta^2}\tilde{X}^\top M_1 \mathbf{1}, \\ \ddot{L}_{\theta\beta} &= \ddot{L}_{\beta\theta}^\top e \\ \ddot{L}_{\theta\theta} &= -\frac{1}{\theta^2}z^\top(I-M_1)z + 2\frac{\lambda}{\theta^2}z^\top M_1 \mathbf{1} + \frac{1}{\theta^2}\mathbf{1}^\top \mathbf{1},\end{aligned}$$

em que $z = (z_1, \dots, z_n)^\top$ é um vetor de dimensão $n \times 1$, $X^{**} = X^{**}(\beta) = \partial^2 \mu / \partial \beta \partial \beta^\top$ é um “array” de dimensão $n \times p \times p$ e $[\cdot][\cdot]$ representa multiplicação de uma matriz por um “array” chamado de produto cochete e definido em Wei (1998, p.188), Bates e Watts (1988) e Seber e Wild (1989, p.691).

A matriz de informação esperada para o vetor de parâmetros ξ , condicionando em λ , é dada por:

$$K_\xi = E\left(-\ddot{L}_{\xi\xi}\right) = \begin{pmatrix} K_{\beta\beta} & K_{\beta\theta} \\ K_{\theta\beta} & K_{\theta\theta} \end{pmatrix} = \tilde{X}^{*\top} W \tilde{X}^*,$$

sendo W uma matriz $(n+1) \times (n+1)$ dada por:

$$W = \frac{1}{\lambda\theta^2} \begin{pmatrix} \lambda I & \mathbf{1} \\ \mathbf{1}^\top & q \end{pmatrix},$$

com I sendo a matriz identidade de dimensão $n \times n$ e $q = n\lambda \left\{1 + \lambda^2 \psi^{(1)}(\lambda^2)\right\}$. Note que W depende do parâmetro θ desconhecido. É importante salientar que os parâmetros não são ortogonais pois $K_{\beta\theta} \neq 0$. Consequentemente, temos que a inversa da matriz de informação é dada por:

$$K_\xi^{-1} = \begin{pmatrix} K^{\beta\beta} & K^{\beta\theta} \\ K^{\theta\beta} & K^{\theta\theta} \end{pmatrix} = (\tilde{X}^{*\top} W \tilde{X}^*)^{-1}. \quad (5)$$

Método de otimização não linear como o algoritmo escore de Fisher ou o algoritmo de Newton pode ser usado para obter os EMV do vetor de parâmetros ξ . Usamos o processo iterativo escore de Fisher, o qual é definido como $\xi^{(k+1)} = \xi^{(k)} + K_{\xi^{(k)}}^{-1} U_{\xi^{(k)}}$, $k = 0, 1, \dots$. Esse processo iterativo pode ser reescrito como um processo de mínimos quadrados ponderados, como se segue:

$$\begin{aligned} \xi^{(k+1)} &= \xi^{(k)} - \frac{\lambda}{\theta^{(k)}} (\tilde{X}^{*(k)\top} W^{(k)} \tilde{X}^{*(k)})^{-1} \tilde{X}^{*(k)\top} M^{(k)} \mathbf{1} \\ &= (\tilde{X}^{*(k)\top} W^{(k)} \tilde{X}^{*(k)})^{-1} \tilde{X}^{*(k)\top} W^{(k)} \delta^{(k)}, \quad k = 0, 1, \dots, \end{aligned} \quad (6)$$

em que $\delta^{(k)} = \tilde{X}^{*(k)} \xi^{(k)} - \frac{\lambda}{\theta^{(k)}} (W^{(k)})^{-1} M^{(k)} \mathbf{1}$ desempenha o papel de uma variável dependente modificada, enquanto $W^{(k)}$ é uma matriz de pesos que muda a cada passo do processo iterativo. Estas equações podem ser implementadas em qualquer software com uma rotina de regressão linear ponderada para calcular o EMV de ξ iterativamente. Aproximações iniciais $\xi^{(0)}$ para o algoritmo iterativo são tomadas para avaliar $M^{(0)}$, $W^{(0)}$, $\tilde{X}^{*(0)}$ e $\delta^{(0)}$ a partir do qual a equação (6) pode ser usado para obter a próxima estimativa $\xi^{(1)}$. Estes novos valores podem atualizar M , W , \tilde{X}^* e δ e assim continuar as iterações até atingir a convergência.

O estimador restrito $\hat{\xi}$, dado λ , tem uma distribuição assintoticamente normal com média ξ e matriz de covariâncias $(K_{\xi})^{-1}$ consistentemente estimada por $(\widehat{X}^{*\top} \widehat{W} \widehat{X}^*)^{-1}$.

Em particular, quando consideramos os parâmetros de escala θ e de forma λ^2 fixos, a função escore para o parâmetro β é dada por:

$$U_{\beta} = -\frac{\lambda}{\theta} \tilde{X}^{\top} M_1 \mathbf{1}.$$

Já a matriz de informação observada de Fisher para β reduz à $\ddot{L}_{\beta\beta}$, dada em (4). Consequentemente, a matriz de informação de Fisher para o parâmetro β , condicionando em θ e em λ , é

dado por

$$K_{\beta} = \frac{1}{\theta^2} \tilde{X}^{\top} \tilde{X}. \quad (7)$$

Para a obtenção do EMV de β , o processo iterativo escore de Fisher (6) pode ser reescrito como:

$$\begin{aligned} \beta^{(k+1)} &= \beta^{(k)} + (\theta^{-2} \tilde{X}^{(k)\top} \tilde{X}^{(k)})^{-1} \left(-\lambda \theta^{-1} \tilde{X}^{(k)\top} M_1^{(k)} \mathbf{1} \right) \\ &= (\tilde{X}^{(k)\top} \tilde{X}^{(k)})^{-1} \tilde{X}^{(k)\top} \boldsymbol{\delta}^{(k)}, \quad k = 0, 1, \dots, \end{aligned} \quad (8)$$

em que $\boldsymbol{\delta}^{(k)} = \tilde{X}^{(k)} \beta^{(k)} - \lambda \theta M_1^{(k)} \mathbf{1}$.

3

Correção de Viés em Modelo Não Linear Log-Gama Generalizado

3.1 Introdução

O principal objetivo da teoria assintótica é calcular o viés de segunda-ordem do estimador de máxima verossimilhança (EMV). É bem conhecido que, sob condições gerais de regularidade, os EMV são consistentes, assintoticamente eficientes e com distribuição assintótica normal. Estes estimadores tipicamente apresentam vieses de ordem $O(n^{-1})$ ¹ para tamanho da amostra (n) grande, o que na prática são comumente ignorados com a justificativa de que eles são pequenos quando comparados com os erros padrão dos estimadores que são de ordem $O(n^{-1/2})$. Para tamanhos de amostras pequenos, no entanto, esses vieses podem ser apreciáveis e terem a mesma magnitude dos erros padrão correspondentes. Em tais casos, os vieses não podem ser desprezados, portanto as fórmulas correspondentes para o seu cálculo precisam ser estabelecidos para uma ampla gama de distribuições e de modelos de regressão. Alguns métodos baseados na avaliação do viés de segunda-ordem para melhorar o EMV podem ser

¹Se $\{a_n\}_{n \geq 1}$ e $\{b_n\}_{n \geq 1}$ são duas seqüências de números reais, dizemos que a_n é de ordem menor que b_n e escrevemos $a_n = o(b_n)$ se $\lim_{n \rightarrow \infty} a_n/b_n = 0$. Dizemos que a_n é de ordem no máximo igual a b_n , denotado por $a_n = O(b_n)$, se existe um número real $M > 0$ tal que $|a_n/b_n| \leq M$, isto é, a razão $|a_n/b_n|$ é limitada. Observamos que se $a_n = o(b_n)$, então $a_n = O(b_n)$ e que quando $b_n \rightarrow 0$ a ordem fornece noção da taxa de convergência de a_n para zero.

muito eficazes. A aproximação normal usual pode frequentemente ser melhorada fazendo um ajuste simples ao viés do EMV. No entanto, a correção de viés pode aumentar ou diminuir o erro quadrado médio (EQM), o qual mede a acurácia do estimador. Para alguns casos especiais, aproximações para o viés dos EMV em modelos simples podem ser obtidas analiticamente.

A metodologia da correção de viés foi aplicado a vários modelos de regressão. Citamos os seguintes modelos: modelo linear log-gama generalizado (Young e Bakir, 1987), modelo normal não linear (Cook, Tsai e Wei, 1986), modelo linear generalizado (Cordeiro e McCullagh, 1991), modelo de regressão não linear multivariado (Cordeiro e Vasconcellos, 1997), modelo linear generalizado com covariáveis modelando o parâmetro de dispersão (Botter e Cordeiro, 1998), modelo de regressão não linear simétrico (Cordeiro *et al.*, 2000), modelo de regressão beta (Ospina, Cribari-Neto e Vasconcellos, 2006) e modelo de regressão para valores extremos (Barreto-Souza e Vasconcellos, 2011). Para quase todos estes modelos, é um resultado notável que o termo principal de viés assintótico pode ser calculado por meio de regressão linear reponderado. O viés aproximado pode ser usado para produzir estimadores corrigidos subtraindo os vieses de ordem $O(n^{-1})$ do EMV.

3.2 Correção de Cox e Snell

O logaritmo da função de verossimilhança (função de log-verossimilhança), $l = l(\xi)$, apresentado em (3) é função do vetor $\xi = (\beta_1, \dots, \beta_p, \theta)^\top$ de parâmetros desconhecidos e de dimensão $(p + 1) \times 1$. Assumimos que $l = l(\xi)$ é regular (Cox e Hinkley, 1974) com respeito à todas as derivadas em relação à β_r , $r = 1, \dots, p$, e θ até a terceira ordem. Definimos os subscritos a, b, c, d como indexadores do espaço paramétrico correspondentes às componentes do vetor ξ , enquanto os subscritos r, s, t, \dots como indexadores do espaço paramétrico correspondente ao vetor β . Introduzimos as seguintes derivadas da função de log-verossimilhança: $U_r = \partial l / \partial \beta_r$, $U_\theta = \partial l / \partial \theta$, $U_{rs} = \partial^2 l / \partial \beta_r \partial \beta_s$, $U_{r\theta} = \partial^2 l / \partial \beta_r \partial \theta$, $U_{rst} = \partial^3 l / \partial \beta_r \partial \beta_s \partial \beta_t$ e assim por diante. A notação padrão é adotada para os cumulantes conjuntos das derivadas

da função de log-verossimilhança (Cordeiro e McCullagh, 1991): $\kappa_{rs} = E(U_{rs})$, $\kappa_{r\theta} = E(U_{r\theta})$, $\kappa_{r,s} = E(U_r U_s)$, $\kappa_{r,\theta} = E(U_r U_\theta)$, $\kappa_{rst} = E(U_{rst})$, $\kappa_{r\theta,t} = E(U_{r\theta} U_t)$, etc. Todos os κ 's referem-se a um total sobre a amostra e são, em geral, de ordem n . Os elementos da matriz de informação de Fisher K_ξ para ξ são $\kappa_{a,b} = -\kappa_{ab}$ e $\kappa^{a,b} = -\kappa^{ab}$ representa o elemento correspondente de K_ξ^{-1} .

Segundo Cox e Snell (1968), para observações independentes, mas não necessariamente identicamente distribuídas, o viés de ordem n^{-1} do EMV $\hat{\xi}_b$ de ξ_b , para $b = 1, 2, \dots, p+1$, pode ser expresso da seguinte forma:

$$B(\hat{\xi}_b) = \sum_{a,c,d} \kappa^{ba} \kappa^{cd} \left(\kappa_{ac}^{(d)} - \frac{1}{2} \kappa_{acd} \right). \quad (9)$$

Definimos um EMV corrigido $\tilde{\xi}_b$ para o parâmetro ξ_b por $\tilde{\xi}_b = \hat{\xi}_b - \hat{B}(\hat{\xi}_b)$, sendo $\hat{B}(\hat{\xi}_b)$ o EMV do viés (9). O viés de $\tilde{\xi}_b$ é de ordem $O(n^{-2})$, pois $E(\tilde{\xi}_b) = \xi_b + O(n^{-2})$. Além disso, podemos esperar que o estimador corrigido $\tilde{\xi}_b$ tenha melhores propriedades do que $\hat{\xi}_b$, cujo viés é de ordem $O(n^{-1})$.

Assim a expressão (9) para calcular o viés de segunda ordem do EMV da b -ésima componente de β , para $b = 1, 2, \dots, p$, é dada por

$$\begin{aligned} B(\hat{\beta}_b) &= \sum_{r,s,t} \kappa^{br} \kappa^{st} \left(\kappa_{rs}^{(t)} - \frac{1}{2} \kappa_{rst} \right) + \sum_{r,s} \kappa^{br} \kappa^{s\theta} \left(\kappa_{rs}^{(\theta)} - \frac{1}{2} \kappa_{rs\theta} \right) \\ &+ \sum_{r,t} \kappa^{br} \kappa^{\theta t} \left(\kappa_{r\theta}^{(t)} - \frac{1}{2} \kappa_{r\theta t} \right) + \sum_{s,t} \kappa^{b\theta} \kappa^{st} \left(\kappa_{\theta s}^{(t)} - \frac{1}{2} \kappa_{\theta s t} \right) \\ &+ \sum_r \kappa^{br} \kappa^{\theta\theta} \left(\kappa_{r\theta}^{(\theta)} - \frac{1}{2} \kappa_{r\theta\theta} \right) + \sum_s \kappa^{b\theta} \kappa^{s\theta} \left(\kappa_{\theta s}^{(\theta)} - \frac{1}{2} \kappa_{\theta s\theta} \right) \\ &+ \sum_t \kappa^{b\theta} \kappa^{\theta t} \left(\kappa_{\theta\theta}^{(t)} - \frac{1}{2} \kappa_{\theta\theta t} \right) + \kappa^{b\theta} \kappa^{\theta\theta} \left(\kappa_{\theta\theta}^{(\theta)} - \frac{1}{2} \kappa_{\theta\theta\theta} \right). \end{aligned} \quad (10)$$

Da mesma forma, temos que a fórmula para calcular o viés de ordem $O(n^{-1})$ do EMV de

θ é da forma

$$\begin{aligned}
B(\hat{\theta}) &= \sum_{r,s,t} \kappa^{\theta r} \kappa^{st} \left(\kappa_{rs}^{(t)} - \frac{1}{2} \kappa_{rst} \right) + \sum_{r,s} \kappa^{\theta r} \kappa^{s\theta} \left(\kappa_{rs}^{(\theta)} - \frac{1}{2} \kappa_{rs\theta} \right) \\
&+ \sum_{r,t} \kappa^{\theta r} \kappa^{\theta t} \left(\kappa_{r\theta}^{(t)} - \frac{1}{2} \kappa_{r\theta t} \right) + \sum_{s,t} \kappa^{\theta\theta} \kappa^{st} \left(\kappa_{\theta s}^{(t)} - \frac{1}{2} \kappa_{\theta st} \right) \\
&+ \sum_r \kappa^{\theta r} \kappa^{\theta\theta} \left(\kappa_{r\theta}^{(\theta)} - \frac{1}{2} \kappa_{r\theta\theta} \right) + \sum_s \kappa^{\theta\theta} \kappa^{s\theta} \left(\kappa_{\theta s}^{(\theta)} - \frac{1}{2} \kappa_{\theta s\theta} \right) \\
&+ \sum_t \kappa^{\theta\theta} \kappa^{\theta t} \left(\kappa_{\theta\theta}^{(t)} - \frac{1}{2} \kappa_{\theta\theta t} \right) + \kappa^{\theta\theta} \kappa^{\theta\theta} \left(\kappa_{\theta\theta}^{(\theta)} - \frac{1}{2} \kappa_{\theta\theta\theta} \right). \quad (11)
\end{aligned}$$

Devido ao fato dos parâmetros não serem ortogonais, as entradas da matriz de informação esperada não são todas iguais as zeros, conseqüentemente todos os termos dados em (10) e (11) devem ser calculados. Os cálculos dos cumulantes e as derivadas dos cumulantes estão apresentados no Apêndice A e os cálculos dos termos de (10) e (11) estão no Apêndice B. Baseado na equação (10) obtemos o seguinte resultado:

$$\begin{aligned}
B(\hat{\beta}_b) &= -\frac{1}{2\theta^2} \sum_i \left(\sum_r \kappa^{br} \tilde{x}_{ir} \right) \left(\sum_{s,t} \kappa^{st} \tilde{x}_{ist} \right) - \frac{1}{2\lambda\theta^3} \sum_i \left(\sum_r \kappa^{br} \tilde{x}_{ir} \right) \left(\sum_{s,t} \tilde{x}_{is} \kappa^{st} \tilde{x}_{it} \right) \\
&- \frac{1}{\theta^3\lambda^2} \sum_i \left(\sum_r \kappa^{br} \tilde{x}_{ir} \right) \left(\sum_s \kappa^{s\theta} \tilde{x}_{is} \right) - \frac{1}{\theta^3} \left(1 + \frac{1}{2\lambda^2} \right) \kappa^{b\theta} \sum_i \left(\sum_{s,t} \tilde{x}_{is} \kappa^{st} \tilde{x}_{it} \right) \\
&- \frac{1}{2\lambda\theta^2} \kappa^{b\theta} \sum_i \left(\sum_{s,t} \kappa^{st} \tilde{x}_{ist} \right) - \frac{\lambda\psi^{(1)}(\lambda^2)}{2\theta^3} \kappa^{\theta\theta} \sum_i \left(\sum_r \kappa^{br} \tilde{x}_{ir} \right) \\
&- \frac{1}{2\theta^3} \left(\frac{4}{\lambda} + 2\lambda\psi^{(1)}(\lambda^2) \right) \kappa^{b\theta} \sum_i \left(\sum_t \kappa^{\theta t} \tilde{x}_{it} \right) \\
&- \kappa^{b\theta} \kappa^{\theta\theta} \frac{n}{2\theta^3} \left\{ 2\lambda^2\psi^{(1)}(\lambda^2) + 3\psi^{(1)}(\lambda^2) - \lambda^2\psi^{(2)}(\lambda^2) \right\},
\end{aligned}$$

em que $\tilde{x}_{ir} = \partial\eta_i/\partial\xi_r$, $\tilde{x}_{irs} = \partial^2\eta_i/\partial\xi_r\partial\xi_s$, $\tilde{x}_{irst} = \partial^3\eta_i/\partial\xi_r\partial\xi_s\partial\xi_t$, etc, segundo a notação proposta por Cordeiro e Paula (1989), e \sum_i é o somatório sobre todas as observações.

Agora definiremos algumas quantidades. Sejam \tilde{x}_i^\top a i -ésima linha da matriz \tilde{X} , $\tilde{\tilde{X}}_i$ uma matriz $p \times p$ cujo elemento (r,s) é \tilde{x}_{irs} , $i = 1, \dots, n$, e $K^{\beta\beta}$ a submatriz $(p \times p)$ da inversa da matriz de informação de Fisher K_ξ dada pela equação (5). Definimos a matriz $Z = \tilde{X}^\top K^{\beta\beta} \tilde{X}$ de

dimensão $n \times n$ com elementos $z_{ij} = \sum \tilde{x}_{is} \mathbf{k}^{st} \tilde{x}_{js}$ e as matrizes diagonais $D = \text{diag}\{d_{11}, \dots, d_{1n}\}$, com $d_{1i} = \text{tr}(K^{\beta\beta} \tilde{X}_i)$, e Z_d formada pelos elementos da diagonal da matriz Z . Adicionalmente, definimos: $q_1 = (2\theta^2)^{-1}$, $q_2 = (2\lambda\theta^3)^{-1}$, $q_3 = (\theta^3\lambda^2)^{-1}$, $q_4 = \theta^{-3} (1 + (2\lambda^2)^{-1})$, $q_5 = (2\lambda\theta^2)^{-1}$, $q_6 = (2\theta^3)^{-1} \lambda \psi^{(1)}(\lambda^2)$, $q_7 = (2\theta^3)^{-1} (4\lambda^{-1} + 2\lambda \psi^{(1)}(\lambda^2))$ e $q_8 = n(2\theta^3)^{-1} \times \{2\lambda^2 \psi^{(1)}(\lambda^2) + 3\psi^{(1)}(\lambda^2) - \lambda^2 \psi^{(2)}(\lambda^2)\}$. Assim, podemos escrever $B(\hat{\beta}_b)$ em notação matricial:

$$\begin{aligned} B(\hat{\beta}_b) &= q_1 \rho_b^\top K^{\beta\beta} \tilde{X}^\top D \mathbf{1} + q_2 \rho_b^\top K^{\beta\beta} \tilde{X}^\top Z_d \mathbf{1} + q_3 \rho_b^\top K^{\beta\beta} \tilde{X}^\top \tilde{X} K^{\beta\theta} + q_4 \rho_b^\top K^{\beta\theta} \mathbf{1}^\top Z_d \mathbf{1} \\ &+ q_5 \rho_b^\top K^{\beta\theta} \mathbf{1}^\top D \mathbf{1} + q_6 \rho_b^\top K^{\beta\beta} \tilde{X}^\top \mathbf{1} K^{\theta\theta} + q_7 \rho_b^\top K^{\beta\theta} \mathbf{1}^\top \tilde{X} K^{\beta\theta} + q_8 \rho_b^\top K^{\beta\theta} K^{\theta\theta}. \end{aligned}$$

Definimos o bloco superior de K_ξ^{-1} , de dimensão $p \times (p+1)$, como $K^{\beta*} = (K^{\beta\beta} K^{\beta\theta})$ e os vetores de tamanhos $n+1$: $\delta_1 = (q_1 \mathbf{1}^\top D, q_5 \mathbf{1}^\top D \mathbf{1})^\top$ e δ_2 ,

$$\delta_2 = \begin{pmatrix} q_2 Z_d \mathbf{1} + q_3 \tilde{X} K^{\beta\theta} + q_6 \mathbf{1} K^{\theta\theta} \\ q_4 \mathbf{1}^\top Z_d \mathbf{1} + q_7 \mathbf{1}^\top \tilde{X} K^{\beta\theta} + q_8 K^{\theta\theta} \end{pmatrix}.$$

Desta forma, o viés de ordem $O(n^{-1})$ de $\hat{\beta}$ pode ser expresso por:

$$B(\hat{\beta}) = K^{\beta*} \tilde{X}^{*\top} (\delta_1 + \delta_2). \quad (12)$$

Analogamente, $B(\hat{\theta})$ é dado por:

$$\begin{aligned}
B(\hat{\theta}) &= -\frac{1}{2\theta^2} \sum_i \left(\sum_r \kappa^{\theta r} \tilde{x}_{ir} \right) \left(\sum_{s,t} \kappa^{st} \tilde{x}_{ist} \right) - \frac{1}{2\lambda\theta^3} \sum_i \left(\sum_r \kappa^{\theta r} \tilde{x}_{ir} \right) \left(\sum_{s,t} \tilde{x}_{is} \kappa^{st} \tilde{x}_{it} \right) \\
&- \frac{1}{\theta^3 \lambda^2} \sum_i \left(\sum_r \kappa^{\theta r} \tilde{x}_{ir} \right) \left(\sum_s \kappa^{s\theta} \tilde{x}_{is} \right) - \frac{1}{\theta^3} \left(1 + \frac{1}{2\lambda^2} \right) \kappa^{\theta\theta} \sum_i \left(\sum_{s,t} \tilde{x}_{is} \kappa^{st} \tilde{x}_{it} \right) \\
&- \frac{1}{2\lambda\theta^2} \kappa^{\theta\theta} \sum_i \left(\sum_{s,t} \kappa^{st} \tilde{x}_{ist} \right) - \frac{\lambda \psi^{(1)}(\lambda^2)}{2\theta^3} \kappa^{\theta\theta} \sum_i \left(\sum_r \kappa^{\theta r} \tilde{x}_{ir} \right) \\
&- \frac{1}{2\theta^3} \left(\frac{4}{\lambda} + 2\lambda \psi^{(1)}(\lambda^2) \right) \kappa^{\theta\theta} \sum_i \left(\sum_t \kappa^{\theta t} \tilde{x}_{it} \right) \\
&- \kappa^{\theta\theta} \kappa^{\theta\theta} \frac{n}{2\theta^3} \left\{ 2\lambda^2 \psi^{(1)}(\lambda^2) + 3\psi^{(1)}(\lambda^2) - \lambda^2 \psi^{(2)}(\lambda^2) \right\}.
\end{aligned}$$

Podemos escrever em notação matricial como

$$\begin{aligned}
B(\hat{\theta}) &= q_1 K^{\theta\beta} \tilde{X}^\top D \mathbf{1} + q_2 K^{\theta\beta} \tilde{X}^\top Z_d \mathbf{1} + q_3 K^{\theta\beta} \tilde{X}^\top \tilde{X} K^{\beta\theta} + q_4 K^{\theta\theta} \mathbf{1}^\top Z_d \mathbf{1} \\
&+ q_5 K^{\theta\theta} \mathbf{1}^\top D \mathbf{1} + q_6 K^{\theta\beta} \tilde{X}^\top \mathbf{1} K^{\theta\theta} + q_7 K^{\theta\theta} \mathbf{1}^\top \tilde{X} K^{\beta\theta} + q_8 K^{\theta\theta} K^{\theta\theta}.
\end{aligned}$$

O bloco inferior de K_ξ^{-1} , de dimensão $p \times (p+1)$, é dado por $K^{\theta*} = (K^{\theta\beta} K^{\theta\theta})$, e então o viés de ordem $O(n^{-1})$ de $\hat{\theta}$ pode ser expresso por:

$$B(\hat{\theta}) = K^{\theta*} \tilde{X}^{*\top} (\delta_1 + \delta_2). \quad (13)$$

Ao combinar (12) e (13), o viés de segunda-ordem do EMV do vetor $\xi = (\beta^\top, \theta)^\top$ pode ser escrito como:

$$B(\hat{\xi}) = K_\xi^{-1} \tilde{X}^{*\top} (\delta_1 + \delta_2) = (\tilde{X}^{*\top} W \tilde{X}^*)^{-1} \tilde{X}^{*\top} (\delta_1 + \delta_2).$$

Definindo $\varepsilon_1 = W^{-1} \delta_1$ e $\varepsilon_2 = W^{-1} \delta_2$, temos que

$$B(\hat{\xi}) = (\tilde{X}^{*\top} W \tilde{X}^*)^{-1} \tilde{X}^{*\top} W (\varepsilon_1 + \varepsilon_2). \quad (14)$$

A equação (14) é o principal resultado dessa seção. O viés $B(\hat{\xi})$ pode ser obtido como o vetor de estimador de mínimos quadrados ponderado da regressão de $(\varepsilon_1 + \varepsilon_2)$ contra as colunas de \tilde{X}^* , sendo W a matriz de pesos. Ele dependerá da não linearidade da função $\eta(x_i; \beta)$ e do parâmetro escala θ . O vetor de viés $B(\hat{\xi})$ será pequeno quando $(\varepsilon_1 + \varepsilon_2)$ for quase ortogonal às colunas de \tilde{X}^* . A equação (14) é facilmente obtido algebricamente para qualquer tipo de regressão não linear, uma vez que envolve operações simples sobre matrizes e vetores. Podemos ainda expressar (14) como $B(\hat{\xi}) = B_1(\hat{\xi}) + B_2(\hat{\xi})$, em que $B_1(\hat{\xi}) = (\tilde{X}^{*\top} W \tilde{X}^*)^{-1} \tilde{X}^{*\top} W \varepsilon_1$, corresponde a parte não linear do modelo, e $B_2(\hat{\xi}) = (\tilde{X}^{*\top} W \tilde{X}^*)^{-1} \tilde{X}^{*\top} W \varepsilon_2$ representa a parte linear. Consequentemente, se temos um MLLGG, então $B_1(\hat{\xi}) = 0$.

Finalmente, no lado direito da equação (14), o EMV do vetor de parâmetros ξ pode ser inserido para definir o estimador corrigido $\tilde{\xi} = \hat{\xi} - \hat{B}(\hat{\xi})$, em que $\hat{B}(\hat{\xi})$ é o valor estimado de $B(\hat{\xi})$. Espera-se que o estimador corrigido $\tilde{\xi}$ tenha melhores propriedades amostrais do que o estimador $\hat{\xi}$.

3.3 Correção por Bootstrap

Uma estratégia diferente a obtenção de um estimador corrigido é através da utilização do método bootstrap (Efron, 1979). A idéia principal é usar reamostragem dos dados a fim de estimar a função do viés. Considere uma amostra aleatória $y = (y_1, \dots, y_n)^\top$ de uma variável aleatória Y com função de distribuição $F = F_\xi(y)$, em que ξ é o vetor de parâmetros que indexa a distribuição e é visto como função de F , ou seja, $\xi = \tau(F)$. Seja $\hat{\xi}$ o EMV para ξ baseado em y , isto é, $\hat{\xi} = s(y)$. No bootstrap paramétrico, procedemos como segue. As amostras bootstrap e todas as estatísticas calculadas a partir delas serão denotadas usando ‘*’. Obtemos, a partir da amostra original y , um grande número (digamos, N) de pseudo-amostras $y^* = (y_1^*, \dots, y_n^*)^\top$, e então usamos tais amostras artificiais para melhorar as inferências estatísticas sobre ξ . A amostra y^* é formada realizando-se a amostragem baseado em $F_{\hat{\xi}}$, ou seja, a partir da função de distribuição do modelo depois que os parâmetros desconhecidos são substituídos por suas

respectivas EMV. Para cada pseudo-amostra y^* , nós estimamos ξ , então obtemos $\hat{\xi}^{*1}, \dots, \hat{\xi}^{*N}$. Em seguida, usamos o estimador bootstrap para obter uma estimativa da função do viés. O viés de $\hat{\xi}$ é $B(\hat{\xi}) = E(\hat{\xi}) - \xi$. Denotamos o viés do estimador $\hat{\xi} = s(y)$ por $B_F(\hat{\xi}, \xi) = E_F[\hat{\xi} - \xi] = E_F[s(y)] - \tau(F)$, em que o subscrito F indica que a esperança é calculada com base em relação à F . Então, a estimativa paramétrica para o viés é dado por:

$$B_{F_{\hat{\xi}}}(\hat{\xi}, \xi) = E_{F_{\hat{\xi}}}[s(y)] - \tau(F_{\hat{\xi}}).$$

Se N amostras bootstrap y^{*1}, \dots, y^{*N} são geradas independentemente a partir da amostra original y , e as réplicas bootstrap correspondentes $\hat{\xi}^{*1}, \dots, \hat{\xi}^{*N}$ são calculadas, em que $\hat{\xi}^{*i} = s(y^{*i})$, $i = 1, \dots, N$, então aproximamos a esperança $E_{F_{\hat{\xi}}}[s(y)]$ pela média amostral $\hat{\xi}^{*(\cdot)} = \frac{1}{N} \sum_{i=1}^N \hat{\xi}^{*i}$. A estimativa do viés baseado em N réplicas de $\hat{\xi}$ é dado por $\hat{B}_{F_{\hat{\xi}}}(\hat{\xi}, \xi) = \hat{\xi}^{*(\cdot)} - s(y)$. Finalmente, definimos o estimador corrigido via bootstrap por:

$$\check{\xi} = s(y) - \hat{B}_{F_{\hat{\xi}}}(\hat{\xi}, \xi) = 2\hat{\xi} - \hat{\xi}^{*(\cdot)}.$$

3.4 Estimação Intervalar

Intervalos de confiança assintóticos são construídos baseados no fato do EMV ter normalidade assintótica, $\hat{\xi} \stackrel{a}{\sim} N(\xi, K_{\xi}^{-1})$ e $\check{\xi} \stackrel{a}{\sim} N(\xi, K_{\xi}^{-1})$, em que K_{ξ}^{-1} é a inversa da matriz de informação de Fisher dada por (5). Para n grande, o intervalo de confiança assintótico (ICA) para ξ_b , $b = 1, \dots, p+1$, ao nível de cobertura de aproximadamente $100(1 - \alpha)\%$ é dado por:

$$(\hat{\xi}_b + z_{(\alpha/2)} \sqrt{\hat{\kappa}^{bb}}; \hat{\xi}_b + z_{(1-\alpha/2)} \sqrt{\hat{\kappa}^{bb}}),$$

em que $-\hat{\kappa}^{bb} = \hat{\kappa}^{b,b}$ representa o elemento (b, b) de K_{ξ}^{-1} avaliada em $\hat{\xi}$ e z_{τ} denota o τ -ésimo ($0 < \tau < 1$) quantil da distribuição normal padrão.

O intervalo de confiança assintótico corrigido (ICAC) para ξ_b com nível de cobertura de

aproximadamente $100(1 - \alpha)\%$ é baseado no estimador corrigido por Cox e Snell, $\tilde{\xi}$, em vez de $\hat{\xi}$, denotado por:

$$(\tilde{\xi}_b + z_{(\alpha/2)}\sqrt{\tilde{\kappa}^{bb}}, \tilde{\xi}_b + z_{(1-\alpha/2)}\sqrt{\tilde{\kappa}^{bb}}), b = 1, \dots, p + 1,$$

em que $-\tilde{\kappa}^{bb} = \tilde{\kappa}^{b,b}$ representa o elemento (b, b) de K_{ξ}^{-1} avaliada em $\tilde{\xi}$.

Uma aproximação alternativa para a construção de intervalos de confiança são fornecidos por métodos computacionalmente intensivos, como o método bootstrap (Efron e Tibshirani, 1986). A ideia consiste em gerar N amostras bootstrap (y^{*1}, \dots, y^{*N}) de tamanho n da amostra original y e calcular o EMV $\hat{\xi}_b^{*l}$ de ξ_b , para $b = 1, \dots, p + 1$, para cada uma das N amostras bootstrap com $l = 1, \dots, N$. O intervalo de confiança percentil (ICP) para ξ_b ao nível de cobertura de aproximadamente $100(1 - \alpha)\%$ é definido tendo como limite inferior e superior os percentis empíricos $\alpha/2$ e $1 - \alpha/2$, respectivamente, dos valores $\hat{\xi}_b^{*l}$. Assim, podemos escrever o ICP como:

$$\left(\hat{\xi}_b^{*(\alpha/2)}, \hat{\xi}_b^{*(1-\alpha/2)} \right). \quad (15)$$

Se usarmos o estimador corrigido por Cox e Snell $\tilde{\xi}_b^{*l}$ em (15), obteremos o intervalo de confiança percentil corrigido (ICPC).

O intervalo de confiança BCa (Bias-Corrected and accelerated) também é construído a partir dos percentis das estimativas bootstrap, mas eles dependem de duas constantes a e z_0 , chamadas de constantes de aceleração e de viés, respectivamente. A constante a é estimada em termos de valores jackknife da forma:

$$\hat{a} = \frac{\sum_{i=1}^n \left(\hat{\xi}_b(\cdot) - \hat{\xi}_b(i) \right)^3}{6 \left\{ \sum_{i=1}^n \left(\hat{\xi}_b(\cdot) - \hat{\xi}_b(i) \right)^2 \right\}^{3/2}},$$

em que $\hat{\xi}_b(i)$ é a i -ésima réplica jackknife de $\hat{\xi}_b$, ou seja, é o estimador $\hat{\xi}_b$ obtido dos dados com a i -ésima observação removida, e $\hat{\xi}_b(\cdot) = \sum_{i=1}^n \hat{\xi}_b(i)/n$. A constante z_0 é estimada como a

proporção de réplicas bootstrap menores que a estimativa original $\hat{\xi}_b$ da forma:

$$\hat{z}_0 = \Phi^{-1} \left(\frac{\#\{\hat{\xi}_b^{*l} < \hat{\xi}_b\}}{B} \right),$$

em que Φ^{-1} denota a inversa da função de distribuição normal padrão. Assim, de forma análoga ao método percentil, os limites do intervalo bootstrap BCa (ICBCa) com nível de cobertura de aproximadamente $100(1 - \alpha)\%$ são os percentis $\hat{\xi}_b^{*(\alpha_1)}$ e $\hat{\xi}_b^{*(\alpha_2)}$, em que

$$\alpha_1 = \Phi \left(\hat{z}_0 + \frac{\hat{z}_0 + z_{\alpha/2}}{1 - \hat{a}(\hat{z}_0 + z_{\alpha/2})} \right) \text{ e } \alpha_2 = \Phi \left(\hat{z}_0 + \frac{\hat{z}_0 + z_{1-\alpha/2}}{1 - \hat{a}(\hat{z}_0 + z_{1-\alpha/2})} \right).$$

Finalmente, um outro método para construir um intervalo de confiança mais simples do que ICBCa é o denominado intervalo de confiança bootstrap- t (ICBt). Para o nível de cobertura de aproximadamente $100(1 - \alpha)\%$, computamos a estatística $T^{*l} = (\hat{\xi}_b^{*l} - \hat{\xi}_b) / \hat{e}p^{*l}$ a partir de cada réplica bootstrap, em que $\hat{e}p^{*l}$ é o erro padrão de $\hat{\xi}_b^{*l}$ ($l = 1, \dots, N$), que é obtido a partir da inversa da matriz de informação esperada. Sejam $\hat{t}^{(\alpha/2)}$ e $\hat{t}^{(1-\alpha/2)}$ os percentis $\alpha/2$ e $1 - \alpha/2$, respectivamente, das estatísticas T^{*l} , $l = 1, \dots, N$. Então, o ICBt para o parâmetro ξ_b , $b = 1, \dots, p + 1$, é dado por:

$$(\hat{\xi}_b - \hat{t}^{(1-\alpha/2)} \sqrt{\hat{\kappa}^{bb}}, \hat{\xi}_b - \hat{t}^{(\alpha/2)} \sqrt{\hat{\kappa}^{bb}}),$$

em que $-\hat{\kappa}^{bb} = \hat{\kappa}^{b,b}$ representa o elemento (b, b) de K_{ξ}^{-1} avaliada em $\hat{\xi}$.

3.5 Resultados Numéricos

Apresentamos através de alguns resultados de simulação de Monte Carlo, em amostras de tamanho finito, o desempenho dos seguintes estimadores: o EMV ($\hat{\xi}$), o EMV corrigido via Cox e Snell ($\check{\xi}$) e por bootstrap paramétrico ($\check{\check{\xi}}$). A fim de analisar os desempenhos desses

estimadores, foram calculados três medidas de qualidade para diferentes cenários: viés, viés relativo e erro quadrático médio (EQM). O viés relativo é definido como $100(\text{viés}/\text{valor verdadeiro do parâmetro})\%$. Consideramos também os intervalos de confiança discutidos na Seção 3.4, a saber: ICA, ICAC, ICP, ICPC, ICBt e ICBCa. Os resultados da estimação intervalar são apresentados em termos das taxas de cobertura dos intervalos (TCI) e dos percentuais dos limites inferiores serem maiores (% Esquerda) e dos limites superiores serem menores (% Direita) do que o verdadeiro valor do parâmetro.

A simulação de Monte Carlo é baseado no seguinte modelo de regressão:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \exp(\beta_2 x_{2i}) + \theta Z_i, \quad i = 1, \dots, n, \quad (16)$$

em que $Z_i \sim LGG(\lambda = 1)$, x_{1i} e x_{2i} são covariadas geradas a partir da distribuição uniforme $U(0, 1)$ e os parâmetros verdadeiros são tais que $\beta_0 = \beta_1 = \theta = 1$ e β_2 variando entre 1; 1,25; 1,5 e 2. O estudo de simulação de Monte Carlo foi realizado com 10.000 réplicas considerando os seguintes tamanhos amostrais: $n = 20, 30, 50$ e 100 . Para cada réplica de Monte Carlo, consideramos $N = 600$ réplicas bootstrap.

Primeiramente, estudamos a influência do tamanho da amostra no desempenho dos três estimadores. Consideramos $\beta_2 = 1, 25$ e $n = 20, 50$ e 100 . A Tabela 3.1 apresenta as estimativas do viés, do viés relativo e do EQM dos estimadores corrigidos e não-corrigidos dos parâmetros do modelo (16). No geral, os estimadores corrigidos são ligeiramente diferentes dos não-corrigidos mesmo quando o tamanho da amostra é grande ($n = 100$). Os resultados da Tabela 3.1 indicam que, para um n fixo, os estimadores corrigidos possuem melhores desempenhos do que os EMV, apresentando estimativas de viés e de EQM menores do que os estimadores não corrigidos, exceto para β_1 quando $n = 50$ e $n = 100$. Isto mostra que as correções de viés desenvolvidos nas seções anteriores são bastantes eficazes e produzem estimadores mais precisos. Se n aumenta, o desempenho em amostras finitas de todos os estimadores melhora, o que está de acordo com a teoria assintótica. Entre os estimadores corrigidos, o estimador $\check{\xi}$

Tabela 3.1 Comparação dos três estimadores para o MNLLGG quando $\beta_0 = \beta_1 = 1$, $\beta_2 = 1,25$, $\theta = 1$ e $\lambda = 1$ para $n = 20, 50$ e 100 .

ξ	n	Viés			Viés Relativo			EQM(10^{-4})		
		ξ	ξ	ξ	ξ	ξ	ξ	ξ	ξ	ξ
β_0	20	0,037	0,037	0,010	3,715	3,676	1,041	13,8010	13,5110	1,0841
	50	0,018	0,009	0,003	1,752	0,944	0,350	3,0694	0,8919	0,1223
	100	0,008	0,004	0,001	0,800	0,442	0,122	0,6405	0,1957	0,0148
β_1	20	0,045	0,006	0,000	4,491	0,551	0,022	20,1660	0,3037	0,0005
	50	-0,004	-0,005	-0,006	-0,437	-0,520	-0,605	0,1913	0,2708	0,3654
	100	-0,001	-0,003	-0,004	-0,146	-0,316	-0,368	0,0212	0,0996	0,1352
β_2	20	-0,086	-0,061	0,019	-6,863	-4,900	1,530	73,5920	37,5070	3,6561
	50	-0,017	-0,011	-0,001	-1,346	-0,916	-0,072	2,8289	1,3118	0,0081
	100	-0,009	-0,006	-0,001	-0,718	-0,485	-0,056	0,8047	0,3679	0,0049
θ	20	-0,102	-0,013	-0,011	-10,184	-1,258	-1,137	103,7000	1,5831	1,2930
	50	-0,041	-0,002	-0,002	-4,055	-0,241	-0,191	16,4390	0,0580	0,0366
	100	-0,019	0,001	0,001	-1,867	0,084	0,090	3,4839	0,0070	0,0081

teve um melhor desempenho no sentido de que, em muitos casos, forneceu a menor estimativa de viés em valor absoluto. As estimativas dos vieses relativos e dos EQM dos estimadores corrigidos refletem o ganho obtido com as correções de viés. Este fato é mais visível para o parâmetro θ do que para os parâmetros β . Analisamos outras simulações variando o λ num intervalo $(0,5;2)$, os resultados foram semelhantes de quando $\lambda = 1$. Em conclusão, os vieses de segunda ordem dos EMV no modelo de regressão MNLLGG não devem ser ignorados em amostras de tamanhos pequeno a moderado, uma vez que eles não são negligenciáveis.

Em segundo lugar, estudamos o comportamento dos estimadores para diferentes valores de β_2 com tamanho da amostra fixo em $n = 30$. Os resultados apresentados na Tabela 3.2 indicam que as estimativas de β_1 e θ são bastante estáveis à medida que aumentamos o verdadeiro valor de β_2 , enquanto as estimativas do vieses do estimador $\tilde{\beta}_0$ e de todos os estimadores de β_2 diminuem. Além disso, o EMV tendem a superestimar o verdadeiro valor de β_1 , enquanto a tendência dos estimadores corrigidos é de subestimar. Outro fato notório é que as estimativas apresentadas pelo estimador $\check{\xi}$ são bem melhores do que às dos demais estimadores em todos os cenários, com exceção das estimativas de β_1 .

Nas Tabelas 3.3 e 3.4 são apresentados os resultados da estimação intervalar. Na Tabela 3.3, consideramos os níveis de cobertura de aproximadamente 90%, 95% e 99% e $n = 20$. Para

Tabela 3.2 Comparação dos três estimadores para o MNLLGG quando $\beta_0 = \beta_1 = 1$, $\beta_2 = 1; 1,5; 2$, $\theta = 1$, $\lambda = 1$ e $n = 30$.

ξ	β_2	Viés			Viés Relativo			EQM(10^{-4})		
		ξ	ξ	ξ	ξ	ξ	ξ	ξ	ξ	ξ
β_0	1,0	0,026	0,024	0,006	2,594	2,444	0,612	6,729	5,973	0,375
	1,5	0,025	0,014	0,007	2,490	1,433	0,696	6,201	2,054	0,485
	2,0	0,025	0,010	0,007	2,480	1,037	0,711	6,150	1,076	0,505
β_1	1,0	0,007	-0,007	-0,009	0,706	-0,712	-0,909	0,499	0,507	0,827
	1,5	0,010	-0,007	-0,010	0,963	-0,746	-0,972	0,927	0,556	0,945
	2,0	0,011	-0,008	-0,010	1,127	-0,773	-0,980	1,270	0,598	0,959
β_2	1,0	-0,050	-0,038	0,018	-5,043	-3,835	1,807	25,430	14,704	3,266
	1,5	-0,026	-0,014	0,000	-1,742	-0,956	-0,024	6,831	2,056	0,001
	2,0	-0,015	-0,006	-0,001	-0,773	-0,308	-0,038	2,393	0,379	0,006
θ	1,0	-0,065	-0,003	-0,002	-6,519	-0,326	-0,226	42,496	0,106	0,051
	1,5	-0,065	-0,003	-0,002	-6,532	-0,340	-0,226	42,668	0,116	0,051
	2,0	-0,065	-0,004	-0,002	-6,544	-0,353	-0,228	42,828	0,125	0,052

os parâmetros β_0 , β_1 e θ , o intervalo ICBt apresentou taxa de coberturas empíricas mais próximas dos níveis de coberturas nominais. Além disso, o ICBt foi o intervalo que apresentou para todos os parâmetros o melhor balanceamento, isto é, a probabilidade do verdadeiro valor do parâmetro ser menor do que o limite inferior (% Esquerda) e a probabilidade de ser maior do que o limite superior (% Direita) do intervalo são praticamente iguais. Para o parâmetro β_2 , associado a não linearidade do modelo, os intervalos assintóticos ICA, ICAC e ICP apresentaram os melhores desempenhos em termos do nível de cobertura. Na Tabela 3.4, fixamos o de cobertura de aproximado em 95% e variamos o tamanho da amostra. O intervalo ICBt tem a menor variação na taxa de cobertura quando o tamanho da amostra n aumenta.

Os gráficos da Figura 3.1 representam os histogramas construídos a partir das 10.000 estimativas de máxima verossimilhança dos parâmetros do modelo (16) e os seis intervalos de confiança para $n = 50$. Os segmentos de reta representam os comprimentos dos intervalos entre as médias dos limites inferiores e superiores das 10.000 réplicas. O histograma do EMV de β_2 apresentou uma leve assimetria à direita, mas o ICPC correspondente apresentou o melhor balanceamento. O histograma do EMV de θ é assimétrico à esquerda, nesse caso o ICBt teve um melhor balanceamento, seguido pelo ICPC.

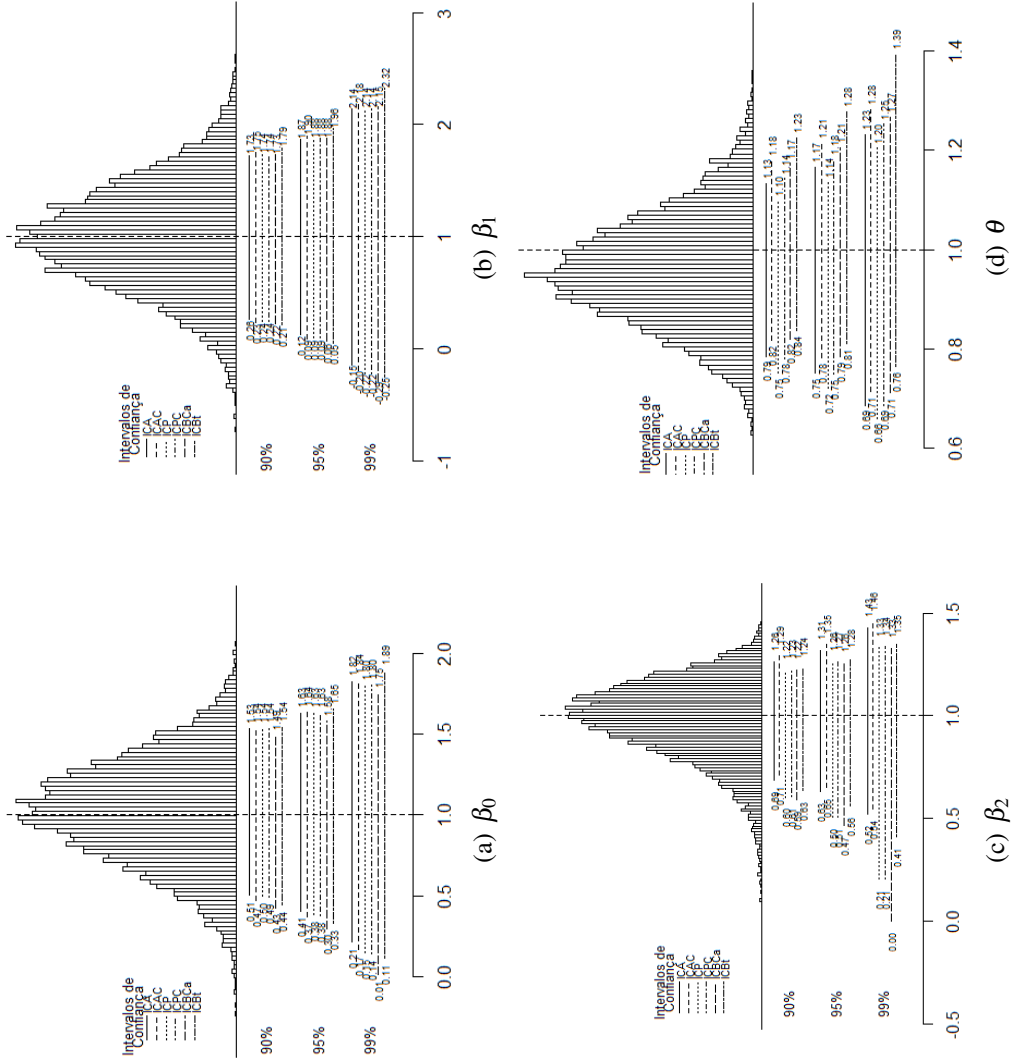
Tabela 3.3 Estimação intervalar dos parâmetros do MNLLGG quando $\beta_0 = \beta_1 = 1$, $\beta_2 = 1, 25$, $\theta = 1$, $\lambda = 1$, $n = 20$ e $(1 - \alpha)\% = 90\%$, 95% , 99% .

Intervalos	β_0		β_1		β_2		θ					
	%		%		%		%					
	TCI	Esquerda	TCI	Esquerda	TCI	Esquerda	TCI	Esquerda				
		Direita		Direita		Direita		Direita				
$\alpha = 0, 10$												
ICA	82,800	10,440	6,760	82,800	9,630	7,570	87,530	9,950	2,520	74,670	0,560	24,770
ICAC	86,210	7,960	5,830	86,050	7,210	6,740	87,270	11,920	0,810	84,630	2,080	13,290
ICP	84,260	9,370	6,370	84,430	9,610	5,960	85,600	3,840	10,560	62,500	0,020	37,480
ICPC	62,500	0,020	37,480	84,490	9,170	6,340	83,150	4,790	12,060	79,680	18,200	2,120
ICBCa	84,820	6,370	8,810	84,690	8,170	7,140	83,870	5,760	10,370	82,610	3,430	13,960
ICBt	88,940	6,210	4,850	89,740	5,140	5,120	80,230	6,200	13,570	89,810	4,860	5,330
$\alpha = 0, 05$												
ICA	89,300	6,580	4,120	88,880	6,380	4,740	92,530	6,850	0,620	80,880	0,100	19,020
ICAC	91,810	4,950	3,240	91,650	4,350	4,000	91,710	8,060	0,230	89,740	0,650	9,610
ICP	90,690	5,440	3,870	90,360	6,170	3,470	91,190	1,810	7,000	72,150	0,010	27,840
ICPC	72,150	0,010	27,840	90,790	5,380	3,830	89,860	2,640	7,500	87,830	11,040	1,130
ICBCa	90,740	3,390	5,870	90,840	4,980	4,180	89,940	2,990	7,070	88,130	1,090	10,780
ICBt	94,530	3,120	2,350	95,080	2,450	2,470	85,440	3,250	11,310	94,810	2,540	2,650
$\alpha = 0, 01$												
ICA	96,040	2,510	1,450	95,870	2,450	1,680	97,000	2,990	0,010	88,750	0,000	11,250
ICAC	97,390	1,570	1,040	97,360	1,390	1,250	96,640	3,360	0,000	94,780	0,010	5,210
ICP	96,820	1,390	1,790	96,970	1,930	1,100	97,100	0,290	2,610	84,330	0,000	15,670
ICPC	84,330	0,000	15,670	96,920	1,370	1,710	96,320	0,860	2,820	96,710	2,990	0,300
ICBCa	96,510	0,960	2,530	97,040	1,620	1,340	96,310	0,750	2,940	91,360	0,020	8,620
ICBt	98,650	0,870	0,480	98,860	0,700	0,440	90,970	1,070	7,960	98,780	0,730	0,490

Tabela 3.4 Estimação intervalar dos parâmetros do MNLLGG, ao nível de confiança de 95% quando $\beta_0 = \beta_1 = 1, \beta_2 = 1, 25, \theta = 1$ e $\lambda = 1$ para $n = 20, 50$ e 100 .

Intervalos	β_0		β_1		β_2		θ	
	TCI	%	TCI	%	TCI	%	TCI	%
$n = 20$								
ICA	89,300	6,580	88,880	6,380	92,530	6,850	80,880	0,100
ICAC	91,810	4,950	91,650	4,350	91,710	8,060	89,740	0,650
ICP	90,690	5,440	90,360	6,170	91,190	1,810	72,150	0,010
ICPC	72,150	0,010	90,790	5,380	89,860	2,640	87,830	11,040
ICBCa	90,740	3,390	90,840	4,980	89,940	2,990	88,130	1,090
ICBt	94,530	3,120	95,080	2,450	85,440	3,250	94,810	2,540
$n = 50$								
ICA	92,740	4,350	92,620	3,630	93,470	4,390	88,930	0,460
ICAC	93,760	3,520	93,620	3,140	93,840	4,850	92,820	1,250
ICP	93,220	3,720	93,170	3,290	92,730	2,290	84,780	0,090
ICPC	84,780	0,090	93,300	3,520	89,360	1,030	64,870	35,040
ICBCa	93,690	2,340	93,270	2,950	92,570	2,110	91,710	1,430
ICBt	94,770	2,740	94,590	2,770	93,070	3,020	94,750	2,760
$n = 100$								
ICA	94,100	3,440	93,980	2,870	94,160	3,630	91,860	0,760
ICAC	94,640	2,950	94,310	2,680	94,390	3,950	93,800	1,530
ICP	94,210	3,100	94,190	2,710	93,750	2,020	89,360	0,360
ICPC	89,360	0,360	94,240	3,000	86,860	0,350	43,110	56,890
ICBCa	94,300	2,180	94,090	2,700	93,780	1,860	93,070	1,610
ICBt	94,890	2,770	94,850	2,580	93,960	2,870	94,630	2,810

Figura 3.1 Estimação intervalar dos parâmetros do MNLGG quando $\beta_0 = \beta_1 = \beta_2 = 1$, $\theta = 1$ e $\lambda = 1$ para $n = 50$.



Fonte: Autoria própria

3.6 Aplicação

Nesta seção, consideramos uma aplicação ao conjunto de dados que correspondem a 46 peças de metal que foram submetidas à fadiga em ciclos, os quais foram analisados por Rieck e Nedelman (1991) e Lemonte e Cordeiro (2009). A variável resposta N denota o número de ciclos até a falha do metal e a variável explicativa x é o trabalho por ciclo (mJ/m^3). Rieck e Nedelman (1991) ajustaram o modelo de regressão linear log-Birnbaum-Saunders ('Modelo I') dado por

$$y_i = \beta_0 + \beta_1 \log(x_i) + \varepsilon_i, \quad i = 1, \dots, 46,$$

em que $y_i = \log(N_i)$ e $\varepsilon_i \sim \text{SN}(\alpha, 0, 2)$, em que SN é a distribuição seno-normal. Os EMV forneceram as seguintes estimativas (com os respectivos erros padrão em parênteses): $\hat{\beta}_0 = 12,280(0,394)$, $\hat{\beta}_1 = -1,671(0,110)$ e $\hat{\alpha} = 0,410(0,043)$.

Lemonte e Cordeiro (2009) ajustaram o modelo de regressão não linear Birnbaum-Saunders ('Modelo II')

$$y_i = \beta_0 + \beta_1 \exp(\beta_2/x_i) + \varepsilon_i, \quad i = 1, \dots, 46,$$

em que $y_i = \log(N_i)$ e $\varepsilon_i \sim \text{SN}(\alpha, 0, 2)$. As EMV são dadas por: $\hat{\beta}_0 = 8,988(0,744)$, $\hat{\beta}_1 = -5,180(0,508)$, $\hat{\beta}_2 = -22,520(7,378)$ e $\hat{\alpha} = 0,40(0,042)$.

Finalmente, utilizamos o modelo de regressão MNLLGG ('Modelo III') descrito como

$$y_i = \beta_0 + \beta_1 \exp(\beta_2/x_i) + \theta z_i, \quad i = 1, \dots, 46,$$

em que $z_i \sim \text{LGG}(\lambda)$, com $\lambda = 0,95$. Vale salientar que o valor do λ foi escolhido dentro da faixa de valores (0;2), de tal forma que minimizasse o AIC. As estimativas pontuais e intervalares, segundo o intervalo Bootstrap- t (ICBt), dos parâmetros encontram-se na Tabela 3.5.

O Critério de Informação de Akaike (AIC) para os Modelos I, II e III foram: 52,741; 52,483 e 50,162, respectivamente. Já o Critério de Informação Bayesiano (BIC) para os três modelos foram: 58,227; 59,798 e 57,477, respectivamente. Ambos os critérios apresentam o

Tabela 3.5 Resultado da estimação pontual e intervalar dos parâmetros dos Modelos III para os dados em estudo.

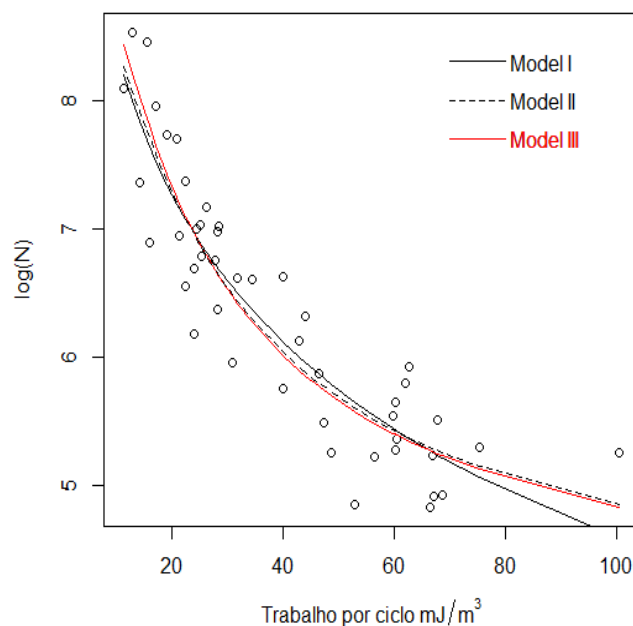
Parâmetros	Estimação pontual (erros padrão)			ICBt(95%)
	$\hat{\xi}$	$\tilde{\xi}$	$\check{\xi}$	
β_0	9,443(0,764)	9,421(0,868)	9,303(0,880)	(8,591; 12,212)
β_1	-5,619(0,563)	-5,728(0,657)	-5,428(0,669)	(-7,397; -5,008)
β_2	-19,723(5,580)	-18,769(5,736)	-18,608(6,050)	(-31,801; -7,874)
θ	0,321(0,037)	0,336(0,039)	0,335(0,039)	(0,270; 0,432)

Modelo III como àquele que fornece o melhor ajuste. Alternativamente, utilizamos o critério de validação definido por

$$CV = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_{(i)})^2,$$

em que $\hat{y}_{(i)} = \eta(x_i, \hat{\beta}_{(i)})$ e $\hat{\beta}_{(i)}$ corresponde à estimativa $\hat{\beta}$ quando a i -ésima observação é excluída. Os valores desse critério fornecidos pelos três modelos foram: 0,179; 0,178 e 0,177, respectivamente, o que indica uma leve vantagem na predição pelo Modelo III. Finalmente, a Figura 3.2 apresenta os valores observados versus os valores ajustados obtidos a partir do Modelo I, II e III. Observamos que o Modelo III fornece um melhor ajuste principalmente nos valores extremos.

Figura 3.2 Valores observados versus valores ajustados a partir dos Modelos I, II e III para os dados em estudo.



Fonte: Autoria própria

3.7 Comentários

Neste capítulo consideramos o modelo de regressão não linear log-gama generalizado com o objetivo de obter um melhor desempenho dos EMV dos parâmetros deste modelo através da correção de viés analítica ou por bootstrap. Derivamos expressões em forma fechada para os vieses de segunda ordem desses estimadores que são facilmente calculados como estimador de mínimos quadrados reponderado da regressão linear a fim de definirmos os estimadores com viés corrigido. Construímos seis tipos de intervalos de confiança com base nas correções de viés. Simulações de Monte Carlo compararam os desempenhos do EMV e suas versões corrigidas analiticamente e por bootstrap e a precisão dos intervalos de confiança. Os resultados da simulação mostraram convincentemente que os vieses de segunda-ordem dos estimadores corrigidos são menores do que os dos EMV. Por fim, apresentamos uma aplicação à dados reais. Consideramos o conjuntos de dados de fadiga de peças de metal, previamente analisados por Rieck e Nedelman (1991), e ajustamos o MNLLGG com $\lambda = 0,95$. O nosso modelo apresentou

um ajuste mais adequado do que os modelos propostos por Rieck e Nedelman (1991) e por Lemonte e Cordeiro (2009).

4

Correção de Bartlett e Tipo-Bartlett em Modelo Não Linear Log-Gama Generalizado

4.1 Introdução

Os testes estatísticos baseados nas estatísticas da razão de verossimilhanças (Wilks, 1938), de Wald (Wald, 1943) e escore (Rao, 1948) são bastantes conhecidos na literatura, sendo os mais utilizados em grandes amostras. Já o teste baseado na estatística gradiente é mais recente e foi proposto por Terrel (2002). Lemonte e Ferrari (2012) mostraram que nenhum desses testes é uniformemente superior aos outros em termos de poder local de segunda ordem. No entanto, o teste escore tem vantagem computacional sobre o teste da razão de verossimilhanças, visto que este envolve estimação dos parâmetros sob a hipótese nula (H_0) e a hipótese alternativa (H_1), enquanto o teste escore requer estimação somente sob H_0 e, no caso em que H_0 é simples, o teste escore não requer estimação alguma. Já a vantagem da estatística gradiente sobre a estatística escore é que ela não depende da matriz de informação, nem esperada e nem observada, sendo mais simples de ser computada.

As estatísticas da razão de verossimilhanças (LR), escore (S_r) e gradiente (S_g) são assintoticamente equivalentes. Isto porque em grandes amostras e sob H_0 , essas estatísticas têm

a mesma distribuição qui-quadrado (χ_q^2) com q graus de liberdade, em que q é o número de restrições impostas por H_0 . O teste baseado nessas três estatísticas são denominados assintótico de primeira ordem por serem baseados em valores críticos obtidos de uma distribuição nula limite conhecida. No entanto, quando o tamanho de amostra é pequeno ou mesmo moderado, esta aproximação pode não ser satisfatória, podendo conduzir a taxas de rejeição bastante distorcida. Portanto, uma preocupação recorrente é verificar a qualidade dessa aproximação.

Para a estatística LR , Bartlett (1937) propôs um fator de correção, visando produzir uma estatística modificada LR^* , cuja média está mais próxima do valor esperado da distribuição χ^2 de referência. Sob H_0 , a $E(LR)$ corresponde a $q\{1 + b + O(n^{-2})\}$, em que n é o tamanho da amostra e b uma constante de ordem n^{-1} que pode ser estimada consistentemente sob H_0 , enquanto que a $E(LR^*)$ corresponde a $q + O(n^{-2})$. Além disso, para testes de homogeneidade de variâncias, Bartlett (1937) mostrou que os três primeiros momentos de LR^* concordam com os momentos correspondentes da distribuição χ^2 até ordem n^{-1} . Consequentemente, temos que a distribuição de LR^* melhor se aproxima da distribuição χ^2 do que a distribuição de LR . Lawley (1956) desenvolveu um método geral de obtenção do fator de correção que envolve momentos das quatro primeiras derivadas do logaritmo da função de verossimilhança (função de log-verossimilhança) e mostrou que a estatística LR^* tem todos os momentos concordando com os respectivos da distribuição χ^2 de referência, ignorando os termos de ordem n^{-2} . Uma alternativa para a correção da estatística LR é utilizar a correção de Bartlett bootstrap (Rocke, 1989), o qual consiste em estimar numericamente o fator da correção de Bartlett. Mais detalhes sobre a correção de Bartlett são encontradas em Cordeiro e Cribari-Neto (2014).

Assim como a correção de Bartlett é utilizada para melhorar a aproximação da distribuição da estatística LR pela distribuição χ_q^2 , alguns refinamentos dos testes escore e gradiente foram obtidos com o objetivo de melhorar a aproximação da distribuição das estatísticas S_r e S_g pela distribuição χ_q^2 . Os trabalhos de Taniguchi (1991) e Chandra e Mukerjee (1991), por exemplo, apresentaram uma estatística escore modificada tendo distribuição χ_q^2 até ordem n^{-1} , mas válido apenas para certos casos especiais. Entretanto, Cordeiro e Ferrari (1991) obtiveram

uma extensão da correção de Bartlett, o qual pode ser utilizada em qualquer outra estatística assintoticamente distribuída como χ_q^2 . O fator de correção obtido por Cordeiro e Ferrari (1991) consiste em uma transformação polinomial da estatística original, por isso não é considerado um fator de correção de Bartlett genuíno, sendo denominado como correção tipo-Bartlett. Essa transformação, no entanto, não é uma transformação monótona da estatística original. Para contornar essa dificuldade, Kakizawa (1996) e Cordeiro *et al.* (1998) obtiveram fórmulas alternativas para a estatística escore modificada pelo fator tipo-Bartlett, as quais tem a vantagem de serem transformações monótonas da estatística S_r . O fator de correção tipo-Bartlett para a estatística S_g foi obtido por Vargas *et al.* (2013) através de uma expansão, sob H_0 , da função de distribuição acumulada da estatística S_g e dos resultados de Cordeiro e Ferrari (1991). Sob H_0 , as estatísticas corrigidas gradiente (S_g^*) e escore (S_r^*) têm distribuição χ_q^2 até erro de ordem n^{-1} , enquanto as estatísticas originais têm distribuição χ_q^2 até erro de ordem $n^{-1/2}$ (vide Lemonte, 2016).

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

regressão não lineares da família exponencial. Cordeiro *et al.* (2006) obtiveram uma expressão matricial para o fator de correção de Bartlett à estatística LR nos modelos lineares generalizados superdispersados. Recentemente, temos o trabalhos de Lemonte *et al.* (2010), o qual obteve a correção de Bartlett no modelo de regressão linear Birnbaum-Saunders. Adicionalmente, Lemonte *et al.* (2010) considerou o teste da razão verossimilhanças bootstrap para o mesmo modelo. Posteriormente, Lemonte *et al.* (2012) generalizou os resultados de Lemonte *et al.* (2010) para o modelo de regressão não-linear Birnbaum-Saunders. Alguns trabalhos podem ser encontrados na literatura baseada nos testes da razão verossimilhanças bootstrap e Bartlett bootstrap, entre eles estão: Bayer e Cribari-Neto (2013) no modelo de regressão beta e Cribari-Neto e Queiroz (2014) no modelo de regressão beta com dispersão variável.

Muitos resultados têm sido publicados na literatura sobre a correção tipo-Bartlett e suas aplicações. Cordeiro *et al.* (2003), por exemplo, derivam um fator de correção tipo-Bartlett para a estatística S_r para modelos lineares com covariadas na dispersão, generalizando os resultados de Cordeiro, Ferrari e Paula (1993) e Cribari-Neto e Ferrari (1995b). Uribe-Opazo *et al.* (2008) e Cysneiros *et al.* (2010) derivaram fator de correção tipo-Bartlett para a estatística S_r na classe dos modelos lineares e não lineares simétricos, respectivamente. Posteriormente os resultados de Cysneiros *et al.* (2010) foram generalizados por Nascimento (2010), que aprimorou o teste score para os modelos não lineares simétricos heteroscedásticos. Lagos *et al.* (2010) e Lemonte e Ferrari (2011) derivaram o fator de correção tipo-Bartlett para o teste score em modelos de séries temporais e Birnbaum Saunders, respectivamente. Vargas *et al.* (2014) obtiveram uma fórmula matricial para o fator de correção tipo-Bartlett da estatística S_g em modelos lineares generalizados com dispersão conhecida e desconhecida, enquanto Medeiros e Ferrari (2016) obtiveram o mesmo em modelos simétricos e log-simétricos.

4.2 Testes de Hipóteses

Considere um modelo multiparamétrico com vetor de parâmetros desconhecidos $\beta = (\beta_1^\top, \beta_2^\top)^\top$, sendo $\beta_1 = (\beta_1, \dots, \beta_q)^\top$ o vetor de parâmetros de interesse e $\beta_2 = (\beta_{q+1}, \dots, \beta_p)^\top$ o vetor de parâmetros de perturbação de dimensões q e $p - q$, respectivamente. A partição $\beta = (\beta_1^\top, \beta_2^\top)^\top$ conduz a partição do vetor escore, descrito como $U_\beta = (U_{\beta_1}^\top, U_{\beta_2}^\top)$, em que $U_{\beta_1} = \partial l(\beta) / \partial \beta_1$ e $U_{\beta_2} = \partial l(\beta) / \partial \beta_2$, sendo $l = l(\beta)$ o logaritmo da função de verossimilhança (função de log-verossimilhança). A matriz de informação de Fisher (K_β) e a sua inversa (K_β^{-1}) também podem ser particionadas da seguinte maneira, respectivamente:

$$K_\beta = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} \text{ e } K_\beta^{-1} = \begin{pmatrix} K^{11} & K^{12} \\ K^{21} & K^{22} \end{pmatrix}.$$

Nosso interesse aqui é testar $H_0 : \beta_1 = \beta_1^{(0)}$ versus $H_1 : \beta_1 \neq \beta_1^{(0)}$, sendo $\beta_1^{(0)}$ um vetor de constantes conhecidas. As estatísticas da razão de verossimilhanças, escore e gradiente são definidas, respectivamente, como:

$$\begin{aligned} LR &= 2\{l(\hat{\beta}) - l(\tilde{\beta})\}, \\ S_r &= U_{\tilde{\beta}_1}^\top \tilde{K}^{11} U_{\tilde{\beta}_1} \text{ e} \\ S_g &= U_{\tilde{\beta}_1}^\top (\hat{\beta}_1 - \beta_1^{(0)}), \end{aligned} \tag{17}$$

em que $\hat{\beta} = (\hat{\beta}_1^\top, \hat{\beta}_2^\top)^\top$ e $\tilde{\beta} = (\beta_1^{(0)\top}, \tilde{\beta}_2^\top)^\top$ são os EMV de $\beta = (\beta_1^\top, \beta_2^\top)^\top$ irrestrito e restrito, respectivamente, $U_{\tilde{\beta}_1}$ e \tilde{K}^{11} correspondem a U_{β_1} e K^{11} , respectivamente, avaliados em $\tilde{\beta}$.

No caso do MNLLGG, definido por (1) e (2), com os parâmetros λ e θ fixos, a decomposição do vetor de parâmetros $\beta = (\beta_1^\top, \beta_2^\top)^\top$ induz a correspondente partição $\tilde{X} = (\tilde{X}_1 \tilde{X}_2)$, sendo \tilde{X} a matriz de derivadas com $\tilde{X}_1 = \partial \eta / \partial \beta_1^\top$ e $\tilde{X}_2 = \partial \eta / \partial \beta_2^\top$. As estatísticas escore e gradiente para o teste de H_0 , assumindo valores fixos para θ e λ , respectivamente, são dadas

por:

$$S_g = -\frac{\lambda}{\theta} \mathbf{1}^\top \tilde{M}_1 \tilde{X} (\hat{\beta} - \tilde{\beta}) \text{ e } S_r = \lambda^2 \mathbf{1}^\top \tilde{M}_1 \tilde{X} (\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top \tilde{M}_1 \mathbf{1},$$

em que \tilde{M}_1 é a matriz M_1 avaliada sob H_0 .

Em grandes amostras e sob H_0 , as três estatísticas têm distribuição χ_q^2 . No entanto, a distribuição χ_q^2 pode não ser uma aproximação satisfatória, podendo conduzir a graves distorções de tamanho dos testes em amostras de tamanho pequeno e moderado. A fim de diminuir esta imprecisão, uma alternativa estratégica é a utilização de uma teoria assintótica de alta ordem.

Adotando r, s, t, u, v, w como indexadores do espaço paramétrico, as derivadas da função de log-verossimilhança podem ser denotadas da seguinte maneira: $U_r = \partial l / \partial \beta_r$, $U_{rs} = \partial^2 l / \partial \beta_r \partial \beta_s$, $U_{rst} = \partial^3 l / \partial \beta_r \partial \beta_s \partial \beta_t$ e assim por diante. Consequentemente, os cumulantes conjuntos dessas derivadas são definidos como $\kappa_{rs} = E(U_{rs})$, $\kappa_{r,s} = E(U_r U_s)$, $\kappa_{rst} = E(U_{rst})$, $\kappa_{r,s,t} = E(U_{rs} U_t)$, etc. Denotamos as derivadas dos momentos em relação aos componentes do vetor β por $\kappa_{rs}^{(t)} = \partial \kappa_{rs} / \partial \beta_t$ e $\kappa_{rs}^{(tu)} = \partial^2 \kappa_{rs} / \partial \beta_t \partial \beta_u$.

4.3 Correção de Bartlett

Sob condições gerais de regularidade, Lawley (1956) obteve uma expansão de $l(\hat{\beta})$ em série de Taylor sob H_0 até termos de ordem n^{-1} envolvendo derivadas até de quarta ordem da função de log-verossimilhança. Assim, ele mostrou que:

$$2E\{l(\hat{\beta}_1, \hat{\beta}_2) - l(\beta_1, \beta_2)\} = p + \varepsilon_p + O(n^{-2}), \quad (18)$$

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

$$\varepsilon_p = \sum (\lambda_{rstu} - \lambda_{rstuvw}), \quad (19)$$

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

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

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

$$2E\{l(\beta_1^{(0)}, \tilde{\beta}_2) - l(\beta_1, \beta_2)\} = p - q + \varepsilon_{p-q} + O(n^{-2}), \quad (20)$$

sendo o termo ε_{p-q} de ordem n^{-1} obtido analogamente ao termo ε_p dado em (19) com o somatório \sum estendendo-se apenas sobre os componentes do vetor β_2 , ou seja, sobre os $p - q$ parâmetros de perturbação. A estatística LR definida em (17) pode ser reescrita como:

$$LR = 2 \left[\{l(\hat{\beta}_1, \hat{\beta}_2) - l(\beta_1, \beta_2)\} - \{l(\beta_1^{(0)}, \tilde{\beta}_2) - l(\beta_1, \beta_2)\} \right].$$

A partir de (18) e (20) segue que, sob H_0 , o valor esperado de LR é dado por:

$$\begin{aligned}E(LR) &= 2E \left[\{l(\hat{\beta}_1, \hat{\beta}_2) - l(\beta_1, \beta_2)\} - \{l(\beta_1^{(0)}, \tilde{\beta}_2) - l(\beta_1, \beta_2)\} \right] \\ &= q + \varepsilon_p - \varepsilon_{p-q} + O(n^{-2}) \\ &= q \left(1 + \frac{\varepsilon_p - \varepsilon_{p-q}}{q} \right) + O(n^{-2}).\end{aligned}$$

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

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

ou, equivalentemente,

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

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

$$d = \frac{\varepsilon_p - \varepsilon_{p-q}}{q}. \quad (21)$$

As estatísticas modificadas LR^* e LR_1^* possuem distribuição χ_q^2 até ordem n^{-1} sob H_0 e sob as condições de regularidade, segundo Hayakawa (1977) (vide correção de Chesher e Smith, 1995). Um teste da razão de verossimilhanças aperfeiçoado compara as estatísticas LR^* e LR_1^* com a distribuição χ_q^2 de referência. Deve-se destacar que os fatores de correção não dependem do valor da estatística LR , mas podem depender de parâmetros desconhecidos e, neste caso, estes devem ser substituídos por suas respectivas estimativas de máxima verossimilhança sob H_0 , o que não afeta a ordem de aproximação resultante. Vale a pena ressaltar que no caso do teste da hipótese nula simples $H_0 : \beta = \beta^{(0)}$, a quantidade d dada em (21) que determina os fatores de correção de Bartlett se reduz a $d = \varepsilon_p/p$, em que ε_p é calculado pela expressão dada em (19).

4.3.1 Correção de Bartlett em MNLLGG com θ fixo

Os fatores de correção de Bartlett dependem da quantidade ε_p , dada em (19), que é uma função aparentemente complicada dos cumulantes conjuntos κ 's de derivadas da função de log-verossimilhança. O objetivo desta seção é apresentar ε_p em forma matricial e de fácil computação para a classe dos MNLLGG.

Utilizamos a notação $\tilde{x}_{ir} = \partial \eta_i / \partial \beta_r$, $\tilde{x}_{irs} = \partial^2 \eta_i / \partial \beta_r \partial \beta_s$, $\tilde{x}_{irst} = \partial^3 \eta_i / \partial \beta_r \partial \beta_s \partial \beta_t$, com $i = 1, \dots, n$, e denotamos por \sum_i o somatório sobre os dados. Temos, então, os seguintes cumulantes

para os MNLLGG (vide Apêndice A):

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

As derivadas dos cumulantes dados acima são:

$$\begin{aligned}
\kappa_{rs}^{(t)} &= \frac{\partial \kappa_{rs}}{\partial \beta_t} = -\frac{1}{\theta^2} \sum_i (\tilde{x}_{irt} \tilde{x}_{is} + \tilde{x}_{ir} \tilde{x}_{ist}), \\
\kappa_{rs}^{(tu)} &= -\frac{1}{\theta^2} \sum_i (\tilde{x}_{irtu} \tilde{x}_{is} + \tilde{x}_{irt} \tilde{x}_{isu} + \tilde{x}_{iru} \tilde{x}_{ist} + \tilde{x}_{ir} \tilde{x}_{istu}) \quad e \\
\kappa_{rst}^{(u)} &= \frac{1}{\lambda \theta^3} \sum_i (\tilde{x}_{iru} \tilde{x}_{is} \tilde{x}_{it} + \tilde{x}_{ir} \tilde{x}_{isu} \tilde{x}_{it} + \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu}) - \frac{1}{\theta^2} \sum_i (\tilde{x}_{irtu} \tilde{x}_{is} + \tilde{x}_{irt} \tilde{x}_{isu} + \tilde{x}_{iru} \tilde{x}_{ist} \\
&\quad + \tilde{x}_{ir} \tilde{x}_{istu} + \tilde{x}_{irsu} \tilde{x}_{it} + \tilde{x}_{irs} \tilde{x}_{itu}).
\end{aligned}$$

Sejam \tilde{x}_i^\top a i -ésima linha da matriz \tilde{X} , \tilde{X}_i uma matriz $p \times p$ cujo elemento (r, s) é \tilde{x}_{irs} , $i = 1, \dots, n$, e K_β^{-1} a inversa da matriz de informação de Fisher K_β dada pela equação (7). Definimos $Z = \theta^{-2} \tilde{X} (\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top$, uma matriz de dimensão $n \times n$ positiva semi-definida de posto p com elementos $z_{ij} = \tilde{x}_i^\top K_\beta^{-1} \tilde{x}_j$, as matrizes quadradas B e C de dimensão $n \times n$, com elementos dados por $b_{ij} = tr(K_\beta^{-1} \tilde{X}_i K_\beta^{-1} \tilde{X}_j)$ e $c_{ij} = \tilde{x}_i^\top K_\beta^{-1} \tilde{X}_j K_\beta^{-1} \tilde{x}_i$, respectivamente, e a matriz diagonal $D = \text{diag}\{d_{11}, \dots, d_{1n}\}$ com $d_{1i} = tr(K_\beta^{-1} \tilde{X}_i)$. Utilizamos a notação Z_d , B_d e C_d para representar matrizes diagonais formadas pelos correspondentes elementos das diagonais das matrizes Z , B e C , respectivamente. Denotamos $Z^{(3)} = Z^{(2)} \odot Z$, $Z^{(2)} = Z \odot Z$, em que \odot denota o produto de Hadamard (Rao, 1973, p. 30), ou seja, o elemento (i, j) de $Z^{(3)}$ é z_{ij}^3 .

Uma expressão simples para o termo ε_p , em notação matricial, pode ser obtida substituindo

os κ 's na expressão (19) e efetuando as somas sobre os parâmetros seguidas das somas sobre as amostras. Ao procedermos assim, aparecerão termos da forma $-\sum \tilde{x}_{ir} \kappa^{rs} \tilde{x}_{js}$, $\sum \kappa^{rs} \tilde{x}_{jsu} \kappa^{ut} \tilde{x}_{itr}$, $\sum \tilde{x}_{it} \kappa^{tu} \tilde{x}_{jus} \kappa^{sr} \tilde{x}_{ir}$ e $-\sum \kappa^{rs} \tilde{x}_{irs}$, em que $-\kappa^{rs} = \kappa^{r,s}$ representa o elemento (r, s) da matriz K_{β}^{-1} , $r, s = 1, \dots, p$. Esses termos correspondem aos elementos das matrizes Z , B , C e D , respectivamente.

Detalhamos a obtenção das parcelas $\sum \lambda_{rstu}$ e $\sum \lambda_{rstuvw}$ de ε_p no Apêndice C. Das definições dos elementos das matrizes Z , B , C e D , estas parcelas são dados por:

$$\sum \lambda_{rstu} = -\frac{1}{4\theta^4 \lambda^2} \sum_i z_{ii}^2 - \frac{1}{2\theta^3 \lambda} \sum_i z_{ii} d_i - \frac{1}{\theta^3 \lambda} \sum_i c_{ii} + \frac{1}{2\theta^2} \sum_i b_{ii} - \frac{1}{4\theta^2} \sum_i d_i^2$$

e

$$\begin{aligned} \sum \lambda_{rstuvw} = & -\frac{1}{6\lambda^2 \theta^6} \sum_{i,j} z_{ij}^3 - \frac{1}{\lambda \theta^5} \sum_{i,j} z_{ij} c_{ji} + \frac{1}{2\theta^4} \sum_{i,j} z_{ij} b_{ji} - \frac{1}{4\lambda^2 \theta^6} \sum_{i,j} z_{ij} z_{ii} z_{jj} \\ & - \frac{1}{2\lambda \theta^5} \sum_{i,j} z_{ij} d_i z_{jj} - \frac{1}{4\theta^4} \sum_{i,j} z_{ij} d_i d_j. \end{aligned}$$

Logo,

$$\begin{aligned} \varepsilon_p = & \sum (\lambda_{rstu} - \lambda_{rstuvw}) \\ = & -\frac{1}{4\theta^4 \lambda^2} \sum_i z_{ii}^2 - \frac{1}{2\theta^3 \lambda} \sum_i z_{ii} d_i - \frac{1}{\theta^3 \lambda} \sum_i c_{ii} + \frac{1}{2\theta^2} \sum_i b_{ii} - \frac{1}{4\theta^2} \sum_i d_i^2 \\ & + \frac{1}{6\lambda^2 \theta^6} \sum_{i,j} z_{ij}^3 + \frac{1}{\lambda \theta^5} \sum_{i,j} z_{ij} c_{ji} - \frac{1}{2\theta^4} \sum_{i,j} z_{ij} b_{ji} + \frac{1}{4\lambda^2 \theta^6} \sum_{i,j} z_{ij} z_{ii} z_{jj} \\ & + \frac{1}{2\lambda \theta^5} \sum_{i,j} z_{ij} d_i z_{jj} + \frac{1}{4\theta^4} \sum_{i,j} z_{ij} d_i d_j. \end{aligned}$$

Em notação matricial temos:

$$\begin{aligned}\varepsilon_p = & -\frac{1}{4\theta^4\lambda^2}\mathbf{1}^\top Z_d Z_d \mathbf{1} - \frac{1}{2\theta^3\lambda}\mathbf{1}^\top Z_d D \mathbf{1} - \frac{1}{\theta^3\lambda}\mathbf{1}^\top C_d \mathbf{1} + \frac{1}{2\theta^2}\mathbf{1}^\top B_d \mathbf{1} - \frac{1}{4\theta^2}\mathbf{1}^\top D D \mathbf{1} \\ & + \frac{1}{6\lambda^2\theta^6}\mathbf{1}^\top Z^{(3)} \mathbf{1} + \frac{1}{\lambda\theta^5}tr(ZC) - \frac{1}{2\theta^4}tr(ZB) + \frac{1}{4\lambda^2\theta^6}Z_d Z Z_d + \frac{1}{2\lambda\theta^5}\mathbf{1}^\top D Z Z_d \mathbf{1} \\ & + \frac{1}{4\theta^4}\mathbf{1}^\top D Z D \mathbf{1}.\end{aligned}$$

A expressão de ε_p resultante é decomposta em

$$\varepsilon_p = \varepsilon_p^{(L)} + \varepsilon_p^{(NL)}, \quad (22)$$

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

$$\varepsilon_p^{(L)} = -\frac{1}{4\lambda^2\theta^4}\mathbf{1}^\top Z_d \left(I - \frac{1}{\theta^2}Z \right) Z_d \mathbf{1} + \frac{1}{6\lambda^2\theta^6}\mathbf{1}^\top Z^{(3)} \mathbf{1}, \quad (23)$$

que reflete a parte linear do modelo e

$$\begin{aligned}\varepsilon_p^{(NL)} = & -\frac{1}{4\lambda\theta^3}\mathbf{1}^\top (2Z_d + \lambda\theta D) \left(I - \frac{1}{\theta^2}Z \right) D \mathbf{1} - \frac{1}{2\lambda\theta^3}\mathbf{1}^\top (2C_d - \lambda\theta B_d) \mathbf{1} \\ & + \frac{1}{2\lambda\theta^5}tr[Z(2C - \lambda\theta B)],\end{aligned} \quad (24)$$

que pode ser interpretado como um termo devido à não linearidade na componente sistemática do modelo. Se $\eta(\cdot; \cdot)$ for linear, as quantidades d_{1i} , c_{ij} e b_{ij} se anulam e, conseqüentemente, $\varepsilon_p^{(NL)} = 0$.

Considere as matrizes $\tilde{X}_{2i} = \partial^2 \eta_i / \partial \beta_2 \partial \beta_2^\top$, $i = 1, \dots, n$, e $K_{\beta_2} = \tilde{X}_2^\top W \tilde{X}_2$. A fórmula de ε_{p-q} é definida analogamente à de ε_p dada em (22) com \tilde{X} , \tilde{X}_i e K_β substituídos por \tilde{X}_2 , \tilde{X}_{2i} e K_{β_2} , respectivamente.

Os fatores de correção de Bartlett para o teste LR discutidos nesta seção são obtidos de (21) com as quantidades ε_p e ε_{p-q} deduzidas de (23) e (24). É importante salientar que a fórmula

dada em (22) somente envolve operações simples de matrizes e pode ser facilmente implementada em pacotes de computação simbólica e linguagens que permitam executar operações simples de álgebra linear, tais como OX, MAPLE, MATHEMATICA, S-Plus, R, etc.

4.4 Correção de Bartlett Bootstrap

Uma alternativa ao cálculo analítico do fator de correção de Bartlett para a estatística LR é utilizar o fator de correção calculado através do método bootstrap introduzido por Efron (1979). Baseado neste método, Rocke (1989) propôs o teste da razão de verossimilhanças bootstrap corrigida, o qual apresenta inferência confiável e não envolve cálculos complexos, apesar de ser computacionalmente custoso.

Denote uma amostra aleatória $y = (y_1, \dots, y_n)^\top$, a correção de Bartlett bootstrap pode ser descrita nos seguintes passos:

1. Geramos N amostras bootstrap (y^{*1}, \dots, y^{*N}) a partir do modelo assumido, considerando H_0 verdadeira e substituindo os parâmetros por suas respectivas estimativas restritas calculadas usando a amostra original y ;
2. Calculamos o valor da estatística LR para cada pseudo-amostra bootstrap, ou seja, para cada pseudo-amostra y^{*b} calculamos a estatística correspondente, LR^{*b} , com $b = 1, \dots, N$, tal que

$$LR^{*b} = 2 \left\{ l \left(\hat{\beta}^{*b}; y^{*b} \right) - l \left(\tilde{\beta}^{*b}; y^{*b} \right) \right\},$$

em que $\hat{\beta}^{*b}$ e $\tilde{\beta}^{*b}$ são EMV de β sob H_1 and H_0 , respectivamente, obtidos a partir da maximização de $l(\beta; y^{*b})$;

3. O teste baseado na estatística da razão de verossimilhanças bootstrap, consiste em rejeitar H_0 se $LR > \hat{q}_{(1-\alpha)}$, em que $\hat{q}_{(1-\alpha)}$ é o percentil estimado $1 - \alpha$ de $LR^* = (LR^{*1}, \dots, LR^{*N})^\top$, tal que

$$\frac{1}{N} \sum_{b=1}^N I\left(\left\{LR^{*b} \leq \widehat{q}_{(1-\alpha)}\right\}\right) = 1 - \alpha,$$

em que $I(\cdot)$ é uma função indicadora. Isto é, $I(\{LR^{*b} \leq \widehat{q}_{(1-\alpha)}\}) = 1$ se $LR^{*b} \leq \widehat{q}_{(1-\alpha)}$ e zero, caso contrário.

A partir de agora, denotamos por LR_{boot} a estatística obtida a partir deste processo, ou seja, LR_{boot} é a mesma estatística LR , mas é comparado com o percentil estimado em vez do quantil da distribuição χ^2 . Já a estatística da razão de verossimilhanças Bartlett bootstrap proposto por Roche (1989) é calculada por

$$LR_{boot}^* = \frac{LR}{\overline{LR^*}} q,$$

em que $\overline{LR^*} = N^{-1} \sum_{b=1}^N LR^{*b}$. Sob H_0 , a estatística LR_{boot}^* segue uma distribuição χ_q^2 aproximadamente. Vale ressaltar que no método bootstrap descrito acima a estatística LR é assintoticamente pivotal, ou seja, a sua distribuição assintótica não depende de parâmetros desconhecidos. Segundo Hall (1992, pp. 83–91) quando a estatística não é assintoticamente pivotal, o erro de aproximação é de magnitude $O(n^{-1})$, enquanto que para a estatística assintoticamente pivotal o erro é de magnitude $O(n^{-2})$.

4.5 Correção Tipo-Bartlett

Nesta seção, apresentaremos fatores de correção tipo-Bartlett para as estatísticas escore e gradiente, respectivamente.

4.5.1 Correção Tipo-Bartlett à Estatística Escore

Seja $F_s(\cdot)$ a função distribuição acumulada de uma variável aleatória χ^2 com s graus de liberdade. Harris (1985), seguindo as idéias de Hayakawa (1977), obteve uma expansão

assintótica até ordem n^{-1} , sob H_0 , para a estatística S_r , tal que

$$P(S_r \leq z) = F_q(z) + \frac{1}{24} \left[A_3 F_{q+6}(z) + (A_2 - 3A_3) F_{q+4}(z) + (3A_3 - 2A_2 + A_1) F_{q+2}(z) + (A_2 - A_1 - A_3) F_q(z) \right] + O(n^{-2}), \quad (25)$$

em que as expressões algébricas para os termos A_1 , A_2 e A_3 são definidos como:

$$A_1 = 3 \sum (\kappa_{rst} + 2\kappa_{r,st}) (\kappa_{uvw} + 2\kappa_{u,v,w}) a^{rs} a^{vw} m^{tu} - 6 \sum (\kappa_{rst} + 2\kappa_{r,st}) \kappa_{u,v,w} a^{rs} a^{tu} m^{vw} + 6 \sum (\kappa_{r,st} - \kappa_{r,s,t}) (\kappa_{uvw} + 2\kappa_{u,v,w}) a^{sv} a^{tw} m^{ru} - 6 \sum (\kappa_{r,s,t,u} + \kappa_{r,s,tu}) a^{tu} m^{rs}, \quad (26)$$

$$A_2 = -3 \sum \kappa_{r,s,t} \kappa_{u,v,w} a^{tu} m^{rs} m^{vw} + 6 \sum (\kappa_{rst} + 2\kappa_{r,st}) \kappa_{u,v,w} a^{rs} m^{tu} m^{vw} - 6 \sum \kappa_{r,s,t} \kappa_{u,v,w} a^{tw} m^{ru} m^{sv} + 3 \sum \kappa_{r,s,t,u} m^{rs} m^{tu} \quad (27)$$

e

$$A_3 = 3 \sum \kappa_{r,s,t} \kappa_{u,v,w} m^{rs} m^{tu} m^{vw} + 2 \sum \kappa_{r,s,t} \kappa_{u,v,w} m^{ru} m^{sv} m^{tw}, \quad (28)$$

sendo \sum o somatório sobre todas as componentes do vetor β e, a^{rs} e m^{rs} , respectivamente, os elementos (r,s) das matrizes

$$\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 0 & K_{22}^{-1} \end{bmatrix} \text{ e } M = K_{\beta}^{-1} - A.$$

No entanto, Harris (1985) e, posteriormente, Cox (1988) observaram que não é possível obter um fator de correção simples para a estatística S_r de tal forma que a estatística modificada, dada por S_r multiplicada por este fator, tenha distribuição χ^2 até ordem n^{-1} . Dessa forma, Harris (1985) propôs um teste escore aperfeiçoado baseado em uma expansão, até ordem n^{-1} , para

seus quantis. Ou seja, em lugar de modificar a estatística de teste S_r , ele propôs a modificação do valor crítico obtido da distribuição χ^2 de referência para um nível de significância fixo.

Mais tarde, Cordeiro e Ferrari (1991), escreveram a expansão assintótica (25) em função da densidade de S_r e usou a relação de recorrência $f_{q+2}(z) = zq^{-1}f_q(z)$, em que $f_q(\cdot)$ denota a função densidade de probabilidade de uma variável aleatória χ^2 com q graus de liberdade, para mostrar que até ordem n^{-1}

$$g(z) = f_q(z) (1 + B_0 + B_1z + B_2z^2 + B_3z^3),$$

em que

$$B_0 = \frac{A_2 - A_1 - A_3}{24}, B_1 = \frac{3A_3 - 2A_2 + A_1}{24q}, B_2 = \frac{A_2 - 3A_3}{24q(q+2)} \text{ e } B_3 = \frac{A_3}{12q(q+2)(q+4)}.$$

Dessa forma, Cordeiro e Ferrari (1991) definem uma estatística S_r modificada da forma

$$S_r^* = S_r [1 - (c + bS_r + aS_r^2)], \quad (29)$$

em que o fator de correção entre chaves é uma função da própria estatística S_r . Eles obtiveram as constantes a , b e c de tal forma que a função densidade da estatística S_r apresentada em (29) coincida com $f_q(z)$ até ordem n^{-1} . Os autores mostraram que

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

A distribuição da estatística S_r^* , sob H_0 , é uma χ_q^2 com o erro aproximado reduzido da ordem n^{-1} para ordem $n^{-3/2}$. Contudo, a não ser que $a = b = 0$, a estatística S_r^* nem sempre será uma transformação monótona de S_r . Em virtude disto, algumas fórmulas alternativas à estatística score modificada foram propostas na literatura. Kakisawa (1996) sugeriu uma transformação

monótona dada por $K(S_r) = S_r^* + P(S_r)$, com

$$P(S_r) = \frac{1}{4} \left\{ c^2 S_r + 2bc S_r^2 + \left(2ac + \frac{4}{3} b^2 \right) S_r^3 + 3ab S_r^4 + \frac{9}{5} a^2 S_r^5 \right\}.$$

Outra estatística alternativa de uma transformação monótona de S_r , expressa em termos da função de distribuição normal padrão, $\Phi(\cdot)$, foi proposta por Cordeiro *et al.* (1998) e é dada por:

$$\tilde{S}_r = \begin{cases} \sqrt{\frac{\pi}{3a}} \exp\left(\frac{b^2}{3a} - c\right) \left\{ \Phi\left(\sqrt{6a}S_r + \sqrt{\frac{2}{3a}}b\right) - \Phi\left(\sqrt{\frac{2}{3a}}b\right) \right\} & \text{se } a > 0, \\ \frac{1}{2b} \exp(-c) \{1 - \exp(-2bS_r)\} & \text{se } a = 0 \text{ e } b \neq 0. \end{cases}$$

As três estatísticas S_r^* , $K(S_r)$ e \tilde{S}_r são equivalentes até segunda ordem, isto é, elas diferem em termos de ordem $n^{-3/2}$ (Cordeiro *et al.*, 2003).

4.5.2 Correção Tipo-Bartlett à Estatística Gradiente

Para obter a correção tipo-Bartlett para estatística S_g , Vargas *et al.* (2013), através de uma expansão assintótica sob H_0 , mostrou que a função distribuição acumulada de S_g pode ser escrita como

$$P(S_g \leq x) = F_q(x) + \frac{1}{24} \sum_{i=0}^3 R_i F_{q+2i}(x) + o(n^{-1}),$$

em que $F_s(\cdot)$ é a função distribuição acumulada de uma variável aleatória χ^2 com s graus de liberdade, $R_1 = 3A_3^g - 2A_2^g + A_1^g$, $R_2 = A_2^g - 3A_3^g$, $R_3 = A_3^g$, $R_0 = -(R_1 + R_2 + R_3)$, com

$$\begin{aligned}
A_1^g &= 3 \sum \kappa_{rst} \kappa_{uvw} \left[m^{rs} a^{vw} (m^{tu} + 2a^{tu}) + a^{rs} m^{tu} a^{vw} + 2m^{ru} a^{sv} a^{tw} \right] \\
&\quad - 12 \sum \kappa_{rs}^{(t)} \kappa_{uv}^{(w)} \left(\kappa^{tr} \kappa^{su} \kappa^{vw} + a^{tr} a^{su} a^{vw} + \kappa^{tu} \kappa^{vr} \kappa^{sw} + a^{tu} a^{vr} a^{sw} \right) \\
&\quad - 6 \sum \kappa_{rst} \kappa_{uv}^{(w)} \left[(a^{tw} - \kappa^{tw}) (\kappa^{ru} \kappa^{vs} - a^{ru} a^{vs}) + m^{rs} (a^{tu} a^{vw} + \kappa^{tu} \kappa^{vw}) \right. \\
&\quad \left. + 2a^{st} (\kappa^{ru} \kappa^{vw} - a^{ru} a^{vw}) + 2a^{su} a^{vt} m^{rw} \right] \\
&\quad + 6 \sum \kappa_{rstw} m^{rs} a^{tw} - 6 \sum \kappa_{rst}^{(w)} \left[m^{rs} (a^{tw} - \kappa^{tw}) + 2m^{rw} a^{st} \right] \\
&\quad + 12 \sum \kappa_{st}^{(rw)} (\kappa^{rs} \kappa^{tw} - a^{rs} a^{tw}), \tag{30}
\end{aligned}$$

$$\begin{aligned}
A_2^g &= -3 \sum \kappa_{rst} \kappa_{uvw} \left[m^{rs} m^{tu} a^{vw} + m^{rs} a^{tu} m^{vw} + 2m^{ru} m^{sv} a^{tw} \right. \\
&\quad \left. + \frac{1}{4} (3m^{rs} m^{tu} m^{vw} + 2m^{ru} m^{sv} m^{tw}) \right] \\
&\quad + 6 \sum \kappa_{rst} \kappa_{uv}^{(w)} \left[m^{tw} (\kappa^{ru} \kappa^{vs} - a^{ru} a^{vs}) + m^{rs} (\kappa^{tu} \kappa^{vw} - a^{tu} a^{vw}) \right] \\
&\quad + 6 \sum \kappa_{rst}^{(w)} m^{rs} m^{tw} - 3 \sum \kappa_{rstw} m^{rs} m^{tw} \tag{31}
\end{aligned}$$

e

$$A_3^g = \frac{1}{4} \sum \kappa_{rst} \kappa_{uvw} (3m^{rs} m^{tu} m^{vw} + 2m^{ru} m^{sv} m^{tw}), \tag{32}$$

sendo \sum o somatório sobre todas as componentes do vetor β , a^{rs} e m^{rs} são, respectivamente, os elementos (r, s) das matrizes A e M . A partir desse resultado e baseado nos resultados de Cordeiro e Ferrari (1991), Vargas *et al.* (2013) obteve uma correção tipo-Bartlett para a estatística S_g , tal que a estatística modificada possui uma distribuição χ_q^2 , sob H_0 , até um erro

de ordem n^{-1} e definida como

$$S_g^* = S_g [1 - (c_g + b_g S_g + a_g S_g^2)], \quad (33)$$

em que

$$a_g = \frac{A_3^g}{12q(q+2)(q+4)}, b_g = \frac{A_2^g - 2A_3^g}{12q(q+2)} \text{ e } c_g = \frac{A_1^g - A_2^g + A_3^g}{12q}.$$

4.5.3 Correção Tipo-Bartlett em MNLLGG com θ fixo

O objetivo desta seção é apresentar, em forma matricial e de fácil computação, os fatores da correção tipo-Bartlett para a estatística S_r , dada em (29), e para a estatística S_g , dada em (33), para a classe dos MNLLGG. Estas correções dependem das quantidades A 's, dadas em (26)–(28) e (30)–(32), que são funções dos cumulantes conjuntos κ 's de derivadas da função de log-verossimilhança (vide Apêndice A). Ao substituir os κ 's nas expressões dos A 's e efetuando as somas sobre os parâmetros seguidas das somas sobre as amostras, aparecerão termos que correspondem aos elementos das matrizes Z , B , C e D , definidas na Seção 4.3.1, e o termo $z_{aij} = \sum \tilde{x}_{is} a^{su} \tilde{x}_{ju}$, o qual corresponde ao elemento (i, j) da matriz $Z_2 = \tilde{X}_2 (\tilde{X}_2^\top \tilde{X}_2)^{-1} \tilde{X}_2^\top$. Adicionalmente, utilizamos a notação Z_{2d} para representar a matriz diagonal formada pelos elementos da diagonal da matriz Z_2 . Os detalhes da obtenção das parcelas das quantidades A 's estão descritos no Apêndice D.

Assim, as quantidades dadas em (26)–(28) necessárias no cálculo do fator da correção tipo-Bartlett para estatística S_r , podem ser descritas na forma matricial como:

$$A_1 = A_1^{(L)} + A_1^{(NL)}, \quad A_2 = A_2^{(L)} + A_2^{(NL)} \text{ e } A_3 = A_3^{(L)},$$

em que $A_1^{(L)}$, $A_2^{(L)}$ e $A_3^{(L)}$ refletem a parte linear do modelo e têm as seguintes formas matriciais:

$$A_1^{(L)} = \frac{3}{\lambda^2 \theta^6} \mathbf{1}^\top Z_{2d} (Z - Z_2) Z_{2d} \mathbf{1} + \frac{12}{\lambda^2 \theta^6} \mathbf{1}^\top Z_{2d} Z_2 (Z_d - Z_{2d}) \mathbf{1} + \frac{18}{\lambda^2 \theta^6} \mathbf{1}^\top (Z - Z_2) \otimes Z_2^{(2)} \mathbf{1} - \frac{24}{\theta^4 \lambda^2} \mathbf{1}^\top (Z_d - Z_{2d}) Z_{2d} \mathbf{1},$$

$$A_2^{(L)} = -\frac{12}{\lambda^2 \theta^6} \mathbf{1}^\top (Z_d - Z_{2d}) Z_2 (Z_d - Z_{2d}) \mathbf{1} - \frac{12}{\lambda^2 \theta^6} \mathbf{1}^\top Z_{2d} (Z - Z_2) (Z_d - Z_{2d}) \mathbf{1} - \frac{24}{\lambda^2 \theta^6} \mathbf{1}^\top (Z - Z_2)^{(2)} \otimes Z_2 \mathbf{1} + \frac{18}{\theta^4 \lambda^2} \mathbf{1}^\top (Z_d - Z_{2d})^2 \mathbf{1}$$

e

$$A_3^{(L)} = \frac{12}{\lambda^2 \theta^6} \mathbf{1}^\top (Z_d - Z_{2d}) (Z - Z_2) (Z_d - Z_{2d}) \mathbf{1} + \frac{8}{\lambda^2 \theta^6} \mathbf{1}^\top (Z - Z_2)^{(3)} \mathbf{1}.$$

As expressões $A_1^{(NL)}$ e $A_2^{(NL)}$ podem ser interpretadas como termos devido à não linearidade na componente sistemática do modelo, são dadas por:

$$A_1^{(NL)} = \frac{6}{\lambda \theta^5} \mathbf{1}^\top Z_{2d} (Z - Z_2) D_2 \mathbf{1} + \frac{3}{\theta^4} \mathbf{1}^\top D_2 (Z - Z_2) D_2 \mathbf{1} + \frac{12}{\lambda \theta^5} \mathbf{1}^\top D_2 Z_2 (Z_d - Z_{2d}) \mathbf{1} + \frac{6}{\lambda \theta^5} \mathbf{1}^\top (Z - Z_2) [2C_2 - \lambda \theta B_2] \mathbf{1} - \frac{12}{\lambda \theta^3} \mathbf{1}^\top (Z_d - Z_{2d}) D_2 \mathbf{1}$$

e

$$A_2^{(NL)} = -\frac{12}{\lambda \theta^5} \mathbf{1}^\top D_2 (Z - Z_2) (Z_d - Z_{2d}) \mathbf{1}.$$

Se $\eta(\cdot; \cdot)$ for linear, as quantidades d_{1i} , c_{ij} e b_{ij} se anulam e, conseqüentemente, $A_1^{(NL)} = A_2^{(NL)} = 0$.

Já as quantidades (30)–(32) necessárias para o cálculo do fator da correção tipo-Bartlett para estatística S_g no MNLLGG são apresentadas no Apêndice F. Para o modelo linear log–

gama generalizado (MLLGG), essas quantidades são expressas na forma matricial da seguinte forma:

$$A_1^g = \frac{3}{\lambda^2 \theta^6} \mathbf{1}^\top \{Z_d Z Z_{2d} + Z_d Z_2 Z_{2d} - 2Z_{2d} Z_2 Z_{2d}\} \mathbf{1} + \frac{6}{\lambda^2 \theta^6} \mathbf{1}^\top (Z - Z_2) \odot Z_2^{(2)} \mathbf{1} - \frac{6}{\theta^4 \lambda^2} \mathbf{1}^\top (Z_d - Z_{2d}) Z_{2d} \mathbf{1},$$

$$A_2^g = -\frac{3}{4\lambda^2 \theta^6} \mathbf{1}^\top (Z_d - Z_{2d}) (Z - Z_2) (3Z_d + Z_{2d}) \mathbf{1} - \frac{3}{\lambda^2 \theta^6} \mathbf{1}^\top (Z_d - Z_{2d}) Z_2 (Z_d - Z_{2d}) \mathbf{1} - \frac{6}{\lambda^2 \theta^6} \mathbf{1}^\top (Z - Z_2)^{(2)} \odot Z_2 \mathbf{1} - \frac{6}{4\lambda^2 \theta^6} \mathbf{1}^\top (Z - Z_2)^{(3)} \mathbf{1} + \frac{3}{\theta^4 \lambda^2} \mathbf{1}^\top (Z_d - Z_{2d})^{(2)} \mathbf{1}$$

e

$$A_3^g = \frac{3}{4\lambda^2 \theta^6} \mathbf{1}^\top (Z_d - Z_{2d}) (Z - Z_2) (Z_d - Z_{2d}) \mathbf{1} + \frac{1}{2\lambda^2 \theta^6} \mathbf{1}^\top (Z - Z_2)^{(3)} \mathbf{1}.$$

4.6 Resultados Numéricos

Com o objetivo de avaliar a eficácia das correções de Bartlett e tipo-Bartlett nos MLLGG e MNLLGG, apresentamos, nesta seção, os resultados de simulações de Monte Carlo. Comparamos os desempenhos de onze testes baseados nas seguintes estatísticas: razão de verossimilhanças (LR), razão de verossimilhanças Bootstrap (LR_{boot}), escore (S_r), gradiente (S_g) e suas respectivas versões corrigidas, a saber: LR^* , LR_1^* , LR_{boot}^* , S_r^* , $K(S_r)$, \tilde{S}_r , S_g^* . Estes desempenhos são avaliados em função da proximidade das probabilidades de rejeição de H_0 , sendo esta verdadeira (probabilidade do erro tipo I), aos seus respectivos níveis nominais dos testes. Avaliamos também os poderes dos testes em estudo sob algumas situações.

O estudo de simulação de Monte Carlo é baseado em 10.000 réplicas considerando os seguintes tamanhos amostrais: $n = 20, 30, 40$ e 50 . Para cada réplica de Monte Carlo, considera-

mos $N = 600$ réplicas bootstrap. Os valores do parâmetro λ variam entre $\{0,5;0,75;1;1,5\}$, já o parâmetro θ é fixo em 1. Os níveis nominais considerados são $\alpha = 1\%, 5\%$ e 10% . Para cada tamanho da amostra e cada nível considerado, calculamos as taxas de rejeição de cada estatística de teste, isto é, estimamos via simulação $P(T \geq \chi_{(\alpha;q)}^2)$, em que $T = LR, LR^*, LR_1^*, LR_{boot}^*, S_r, S_r^*, K(S_r), \tilde{S}_r, S_g, S_g^*$ e $\chi_{(\alpha;q)}^2$ é o percentil $(1 - \alpha)$ da distribuição χ_q^2 , com $q = 2$. Para o teste bootstrap, a taxa de rejeição é obtida a partir do cálculo da probabilidade $P(LR > \hat{q}_{(1-\alpha)})$, em que $\hat{q}_{(1-\alpha)}$ é o percentil $(1 - \alpha)$ estimado da distribuição empírica de LR^* .

4.6.1 Modelos Lineares

Nesta seção, apresentamos os resultados numéricos baseados nos seguintes MLLGG:

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

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

Assumindo diversos valores para λ , na Tabela 4.1 temos as taxas de rejeição dos testes nos modelos com $p = 8$ e $n = 30$. O teste baseado na estatística LR mostra ser mais liberal do que os testes baseados nas estatísticas corrigidas, rejeitando mais do que deveria, principalmente para valores menores de λ . Por sua vez, os testes baseados em LR^*, LR_1^*, LR_{boot} e LR_{boot}^* apresentam taxas de rejeição bem mais próximas aos níveis nominais e mais estáveis em relação à variação do valor do λ , evidenciando a eficácia da correção. Por exemplo, para $\alpha = 10\%$ e $\lambda = (0,5;0,75;1;1,5)$ as taxas de rejeição do teste baseado na estatística LR variam de $15,9\%$ à $10,5\%$, enquanto que as taxas de rejeição dos testes baseados em LR^*, LR_1^*, LR_{boot} e LR_{boot}^* variam entre $9,3\%$ e $10,1\%$, $7,3\%$ e $10,0\%$, $9,9\%$ e $10,7\%$, e $9,8\%$ e $10,7\%$, respectivamente. Já o teste baseado na estatística S_r é o mais conservativo, apresentando taxas

de rejeição inferiores aos níveis nominais correspondentes, sendo tal tendência atenuada por suas versões corrigidas (S_r^* , $K(S_r)$ e \tilde{S}_r). No entanto, as taxas de rejeição de S_r aproximam-se do nível nominal com o aumento do valor de λ . O teste baseado em S_g , por sua vez, apresenta taxas de rejeição bastante próximas aos níveis nominais adotados, independente do valor do λ , de forma que o ganho com a correção tipo-Bartlett não é tão notória.

Tabela 4.1 Taxas de rejeição dos testes LR , LR^* , LR_1^* , LR_{boot} , LR_{boot}^* , S_r , S_r^* , $K(S_r)$, \tilde{S}_r , S_g e S_g^* em MLLGG com $\theta = 1$, $n = 30$, $p = 8$ e diversos valores de λ .

$\alpha(\%)$	λ	LR	LR^*	LR_1^*	LR_{boot}	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r	S_g	S_g^*
1	0,50	2,1	0,6	0,4	1,2	0,8	0,3	1,6	1,7	2,0	0,6	0,8
	0,75	1,7	1,1	0,9	1,3	1,1	0,5	0,9	1,0	1,0	0,9	1,0
	1,00	1,5	1,0	1,0	1,2	1,1	0,7	1,1	1,1	1,2	1,0	1,0
	1,50	1,3	1,1	1,0	1,2	1,0	0,9	1,1	1,1	1,1	1,0	1,1
5	0,50	9,1	4,3	3,1	5,1	4,7	2,1	4,1	4,3	4,5	5,2	4,2
	0,75	7,1	5,2	4,9	5,0	5,0	2,5	4,0	4,2	4,2	5,3	4,9
	1,00	6,1	5,1	5,0	5,3	5,3	3,6	4,7	4,8	4,8	5,3	5,1
	1,50	5,4	5,0	5,0	5,5	5,4	4,2	4,8	4,9	4,9	5,1	5,0
10	0,50	15,9	9,3	7,3	10,4	10,1	4,0	6,9	7,1	7,5	11,4	8,4
	0,75	13,2	10,0	9,5	9,9	9,8	6,3	8,9	9,0	9,2	11,0	9,7
	1,00	11,9	10,1	10,0	10,3	10,1	8,0	9,5	9,6	9,6	10,9	10,1
	1,50	10,5	9,7	9,7	10,7	10,7	9,0	9,7	9,7	9,7	10,2	9,8

Para avaliar o comportamento dos testes quando o tamanho da amostra aumenta, fixamos o número de parâmetros $p = 8$, $\lambda = 1$ e variamos $n = (20, 30, 40, 50)$. Com base nos resultados apresentados na Tabela 4.2, podemos observar que os testes baseados nas estatísticas não corrigidas LR , S_r e S_g tendem à produzir taxas de rejeição mais próximas aos níveis nominais conforme aumentamos o tamanho da amostra. Por outro lado, as taxas de rejeição dos testes baseados nas estatísticas corrigidas e em LR_{boot} tendem a permanecer estáveis mesmo aumentando o tamanho amostral. Por exemplo, para $\alpha = 10\%$ e $n = (20, 30, 40, 50)$, as taxas de rejeição do teste baseado na estatística LR variam de 11,1% à 12,8%, enquanto que as taxas de rejeição dos testes baseados nas versões corrigidas LR^* e LR_1^* variam entre 9,7% e 10,3%, já as taxas de rejeição dos testes baseados em LR_{boot} e LR_{boot}^* variam entre 10,1% e 11,0%. Neste mesmo cenário, as taxas de rejeição do teste S_r variam de 7,2% à 8,9%, já taxas de rejeição das suas três versões corrigidas variam entre 9,5% e 10,0%. Mesmo produzindo taxas de rejeição

mais estáveis que os demais testes não corrigidos, o teste S_g apresenta taxas de rejeição, ainda considerando $\alpha = 10\%$, que variam de 10,4% à 11,3%, enquanto que as taxas de rejeição do teste baseado na S_g^* variam entre 9,9% e 10,1%.

Tabela 4.2 Taxas de rejeição dos testes LR , LR^* , LR_1^* , LR_{boot} , LR_{boot}^* , S_r , S_r^* , $K(S_r)$, \tilde{S}_r , S_g e S_g^* em MLLGG com $\theta = 1$, $p = 8$, $\lambda = 1$ e diversos valores de n .

$\alpha(\%)$	n	LR	LR^*	LR_1^*	LR_{boot}	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r	S_g	S_g^*
1	20	1,7	1,0	0,9	1,4	1,1	0,6	1,2	1,2	1,3	0,9	1,0
	30	1,5	1,0	1,0	1,2	1,1	0,7	1,1	1,1	1,2	1,0	1,0
	40	1,3	1,1	1,0	1,1	0,9	0,7	1,1	1,1	1,1	1,0	1,0
	50	1,4	1,2	1,1	1,3	1,1	0,8	1,1	1,1	1,1	1,1	1,1
5	20	7,2	5,5	5,3	5,4	5,2	3,1	4,9	4,9	5,0	5,7	5,3
	30	6,1	5,1	5,0	5,3	5,3	3,6	4,7	4,8	4,8	5,3	5,1
	40	5,7	5,0	5,0	5,2	5,1	3,9	4,9	4,9	4,9	5,3	5,1
	50	5,7	5,0	5,0	5,3	5,3	4,2	5,1	5,1	5,1	5,1	5,0
10	20	12,8	10,3	10,0	10,5	10,4	7,2	9,6	9,7	9,7	11,3	10,1
	30	11,9	10,1	10,0	10,3	10,1	8,0	9,5	9,6	9,6	10,9	10,1
	40	11,2	9,8	9,7	10,4	10,4	8,5	9,8	9,8	9,8	10,4	9,9
	50	11,1	10,1	10,1	11,0	10,9	8,9	10,0	10,0	10,0	10,4	10,0

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

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

Com o número fixo de parâmetros de interesse em $q = 2$, uma vez que a hipótese nula a ser testada é $H_0 : \beta_5 = \beta_6 = 0$, apresentamos na Tabela 4.3 as taxas de rejeição dos testes nos modelos acima. Podemos notar que o número de parâmetros de perturbação ($p - 2$) tem um impacto considerável sobre a aproximação por χ^2 das distribuições das estatísticas não corrigidas LR , S_r

e S_g . À medida que aumentamos o número de parâmetros de perturbação, as taxas de rejeição do teste LR tornam-se consideravelmente superiores ao nível de significância correspondente, o mesmo ocorre, embora de maneira menos nítida, com o teste S_g , enquanto o teste baseado na estatística S_r tende a apresentar taxas de rejeição bem abaixo do nível nominal especificado. Os testes baseados nas estatísticas corrigidas apresentam desempenhos superiores, visto que mesmo variando o número de parâmetros de perturbação, as suas taxas de rejeição tendem a permanecer mais próximas dos níveis nominais. Dentre os testes corrigidos, LR_{boot} e LR_{boot}^* são os que apresentam tendência conservativa quando o número de parâmetros de perturbação é pequeno.

Tabela 4.3 Taxas de rejeição dos testes LR , LR^* , LR_1^* , LR_{boot} , LR_{boot}^* , S_r , S_r^* , $K(S_r)$, \tilde{S}_r , S_g e S_g^* em MLLGG com $\theta = 1$, $n = 30$, $\lambda = 1$ e diversos valores de p .

$\alpha(\%)$	p	LR	LR^*	LR_1^*	LR_{boot}	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r	S_g	S_g^*	
1	3	1,1	1,1	1,1	1,1	1,0	1,1	1,0	1,0	1,0	1,0	1,1	
	4	1,2	1,1	1,1	1,1	1,0	1,0	1,2	1,2	1,2	1,1	1,2	
	5	1,3	1,1	1,1	1,1	0,9	1,0	1,2	1,2	1,2	1,1	1,1	
	6	1,3	1,1	1,1	1,2	1,1	0,6	1,1	1,1	1,1	1,0	1,1	
	7	1,4	1,0	1,0	1,3	1,1	0,6	1,0	1,1	1,1	0,9	1,1	
	8	1,5	1,0	1,0	1,2	1,1	0,7	1,1	1,1	1,1	1,0	1,0	
	5	3	5,5	5,2	5,2	4,7	4,6	4,9	5,2	5,2	5,2	5,1	5,2
		4	5,5	5,2	5,2	4,9	4,8	4,3	5,2	5,2	5,2	5,1	5,2
5		5,7	5,2	5,2	4,8	4,7	4,5	5,3	5,3	5,3	5,3	5,3	
6		5,8	5,1	5,1	5,2	5,1	3,7	5,0	5,0	5,0	5,1	5,1	
7		6,1	5,2	5,1	5,2	5,2	3,7	5,0	5,0	5,1	5,3	5,1	
8		6,1	5,1	5,0	5,3	5,3	3,6	4,7	4,8	4,8	5,3	5,1	
10		3	10,5	10,2	10,2	9,4	9,1	9,3	10,2	10,2	10,2	10,1	10,2
		4	10,9	10,4	10,4	9,2	9,1	8,9	10,1	10,1	10,1	10,3	10,3
	5	11,2	10,5	10,4	9,3	9,1	9,3	10,3	10,3	10,3	10,7	10,4	
	6	11,3	10,4	10,3	9,8	9,8	8,2	9,9	10,0	10,0	10,5	10,4	
	7	11,8	10,6	10,5	9,6	9,5	8,4	10,0	10,0	10,1	10,9	10,5	
	8	11,9	10,1	10,0	10,3	10,1	8,0	9,5	9,6	9,6	10,9	10,1	

Para o estudo do poder, fixamos $\lambda = 1$, $n = 20$, $p = 4$ e $\alpha = 5\%$. Os resultados apresentados na Tabela 4.4 foram obtidos levando em consideração a hipótese alternativa $H_1 : \beta_5 = \beta_6 \neq 0$ e diferentes valores de $\beta_5 = \beta_6 = \beta^{(0)}$, com $\beta^{(0)}$ variando de 0,0 a 3,5. Os resultados indicam que não há uma perda expressiva de poder derivada do fato de usar o fator de correção de Bartlett, uma vez que os poderes dos testes baseados nas estatísticas LR , LR^* , LR_1^* , LR_{boot} e LR_{boot}^* são

bem próximos. Por outro lado, o uso do fator de correção tipo-Bartlett nas estatísticas S_r e S_g resulta num ganho no poder. Como observado na Figura 4.1, o desempenho da estatística S_r é bem inferior aos dos demais testes. Os testes baseados nas estatísticas S_r^* , $K(S_r)$, \tilde{S}_r apresentam poderes semelhantes, mas também ligeiramente inferiores aos dos demais testes corrigidos.

Tabela 4.4 Poder dos testes LR , LR^* , LR_1^* , LR_{boot} , LR_{boot}^* , S_r , S_r^* , $K(S_r)$, \tilde{S}_r , S_g e S_g^* , em MLLGG com $\theta = 1$, $\lambda = 1$, $p = 4$, $n = 20$.

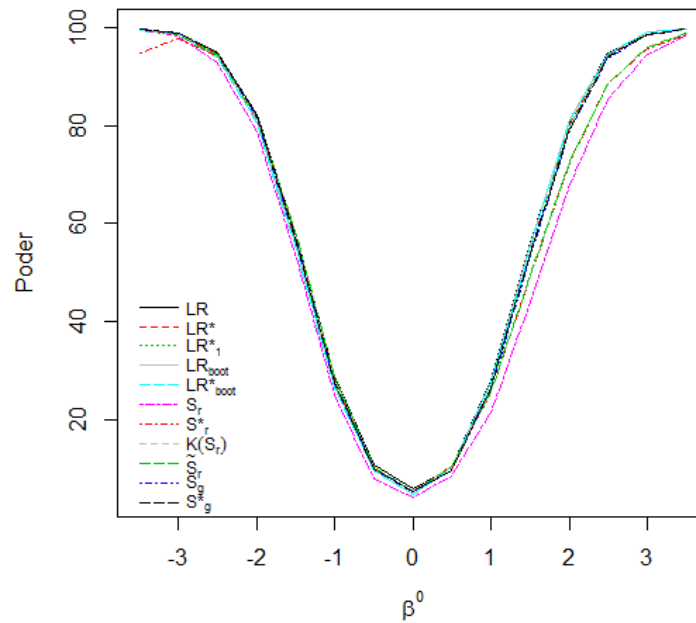
β^0	LR	LR^*	LR_1^*	LR_{boot}	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r	S_g	S_g^*
0,0	5,8	5,3	5,4	4,7	4,7	4,1	5,3	5,4	5,4	5,2	5,3
0,5	10,3	9,5	9,5	10,3	10,1	8,6	10,5	10,5	10,6	9,5	9,7
1,0	28,1	26,5	26,6	27,6	27,3	21,4	25,5	25,5	25,6	26,3	26,6
1,5	56,0	54,1	54,1	55,8	55,3	44,0	49,2	49,3	49,5	53,2	53,8
2,0	81,1	80,0	80,1	81,2	80,9	67,8	72,5	72,6	72,7	79,1	79,3
2,5	94,7	94,1	94,2	94,6	94,6	85,3	88,4	88,5	88,6	93,7	93,9
3,0	98,6	98,6	98,6	99,0	99,0	94,5	95,8	95,9	96,0	98,4	98,5
3,5	99,8	99,8	99,8	99,8	99,8	98,1	98,6	98,7	98,7	99,8	99,8

Na Tabela 4.5 apresentamos comparações das médias e variâncias das estatísticas LR , LR^* , LR_1^* , LR_{boot}^* , S_r , S_r^* , $K(S_r)$, \tilde{S}_r , S_g e S_g^* com a da distribuição χ_2^2 , para o modelo MLLGG com $n = 20$, $\lambda = 0,5$ e $p = 4$. Os resultados mostram que as estatísticas corrigidas apresentam médias e variâncias bem mais próximas da distribuição χ_2^2 do que as estatísticas não corrigidas. Este mesmo comportamento é observado na Tabela 4.6 ao comparar os quantis amostrais dessas estatísticas com os da distribuição χ_2^2 . No geral, as estatísticas que apresentam os momentos e os quantis empíricos mais afastados da distribuição de referência são as estatísticas S_r e \tilde{S}_r . Na Figura 4.2 construímos o gráfico de discrepância relativa de quantis, definida como

$$\frac{T(1 - \alpha) - \chi_{(\alpha;q)}^2}{\chi_{(\alpha;q)}^2},$$

em que $T(1 - \alpha)$ denota o quantil amostral de ordem $(1 - \alpha)$ do conjunto de valores simulados da estatística do teste T , sendo $T = LR$, LR^* , LR_1^* , LR_{boot}^* , S_r , S_r^* , $K(S_r)$, \tilde{S}_r , S_g , S_g^* . Sendo assim, quanto mais próxima da ordenada nula a curva de discrepância estiver, melhor aproxi-

Figura 4.1 Curva de poder dos testes para o MLLGG com $\theta = 1$, $\lambda = 1$, $p = 4$, $n = 20$.



Fonte: Autoria própria

mada está a distribuição nula da estatística do teste pela distribuição assintótica χ_2^2 . A análise do gráfico mostra que a curva de discrepância da estatística S_r está bastante abaixo da ordenada nula, o que ratifica a tendência conservadora do teste baseado nesta estatística. Também podemos observar que apenas as estatísticas LR^* , LR_1^* , LR_{boot}^* e S_g^* têm curvas de discrepância mais próximas à ordenada nula. Por outro lado, a estatística LR apresenta curva de discrepância ligeiramente acima da ordenada nula, confirmando a tendência liberal do teste baseado em tal estatística.

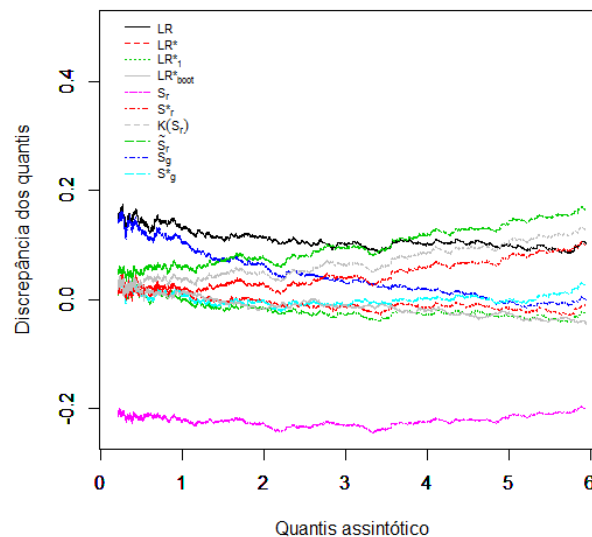
Tabela 4.5 Média e variância da distribuição χ_2^2 e das estatísticas $LR, LR^*, LR_1^*, LR_{boot}^*, S_r, S_r^*, K(S_r), \tilde{S}_r, S_g$ e S_g^* em MLLGG com $p = 4, n = 20$ e $\lambda = 0,5$.

momentos	χ_2^2	LR	LR^*	LR_1^*	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r	S_g	S_g^*
média	2,000	2,213	1,982	1,955	1,969	1,557	2,090	2,137	2,196	2,057	2,010
variância	4,000	4,792	3,843	3,739	3,759	2,461	4,464	4,689	4,980	3,756	4,257

Tabela 4.6 Quantis da distribuição χ_2^2 e das estatísticas $LR, LR^*, LR_1^*, LR_{boot}^*, S_r, S_r^*, K(S_r), \tilde{S}_r, S_g$ e S_g^* em MLLGG com $p = 4, n = 20$ e $\lambda = 0,5$.

quantis(%)	χ_2^2	LR	LR^*	LR_1^*	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r	S_g	S_g^*
85,0	3,8	4,2	3,8	3,7	3,7	2,9	4,0	4,1	4,2	3,9	3,8
90,0	4,6	5,0	4,5	4,5	4,5	3,6	4,9	5,0	5,2	4,6	4,6
95,0	6,0	6,6	5,9	5,8	5,9	4,8	6,5	6,7	6,9	6,0	6,1
97,0	7,0	7,7	6,9	6,8	6,8	5,6	7,7	7,9	8,2	6,9	7,3
99,0	9,2	10,2	9,2	9,0	9,0	7,3	9,8	10,0	10,3	8,7	9,5
99,5	10,6	11,5	10,3	10,2	10,4	8,2	10,7	10,9	11,2	10,0	11,2

Figura 4.2 Discrepância relativa de quantis para o MLLGG com $\theta = 1, \lambda = 0,5, p = 4, n = 20$.



Fonte: Autoria própria

4.6.2 Modelos Não-Lineares

Nesta seção, apresentamos os resultados numéricos baseados nos seguintes MNLLGG:

$$Y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \exp(\beta_5 x_{i5}) + \theta Z_i, \quad i = 1, \dots, n \text{ e } k = 1, \dots, 4,$$

em que $Z_i \sim LGG(\lambda)$. Consideramos a hipótese nula $H_0 : \beta_3 = \beta_4 = 0$ e a variável resposta gerada assumindo que $\beta_0 = \beta_1 = \beta_2 = \beta_5 = 0,05$. As covariáveis x_1, \dots, x_5 foram tomadas como amostras aleatórias das seguintes distribuições: $LN(0, 1)$, $F(2, 5)$, $Beta(2, 3)$, $U(0, 1)$ e $N(0, 2)$.

Assumindo diversos valores para λ , na Tabela 4.7 temos as taxas de rejeição dos testes nos modelos com $p = 5$ e $n = 20$. Podemos observar que o teste baseado na estatística LR , assim como nos modelos lineares, apresenta taxas de rejeição superiores aos níveis nominais correspondentes, principalmente para $\lambda = 0,5$. Por outro lado, os testes baseados nas versões corrigidas de LR e LR_{boot} apresentam tendência conservativa para valores pequenos de λ . Observamos também que o teste LR_{boot} apresenta taxas de rejeição próximas dos respectivos níveis nominais. Já o teste S_r apresenta taxas de rejeição inferiores aos níveis nominais correspondentes, porém essas taxas aproximam-se dos níveis nominais à medida que o valor de λ aumenta. Entre os testes baseados nas versões corrigidas da S_r , S_r^* é o que apresenta taxas de rejeição mais próximas aos níveis nominais mesmo com a variação do valor de λ .

Na Tabela 4.8, fixamos $\lambda = 1$ e $p = 5$ e variamos o tamanho da amostra. Os resultados mostram que, em quase todos os cenários, as taxas de rejeição do teste LR são superiores aos dos demais testes, enquanto que o teste S_r fornecem as menores taxas de rejeição, abaixo dos níveis nominais correspondente. Os testes corrigidos, no entanto, apresentam taxas de rejeição mais próximas aos níveis nominais mesmo com a variação do tamanho amostral. Por exemplo, para $\alpha = 5\%$ e $n = (20, 30, 40, 50)$ as taxas de rejeição do teste baseado na estatística LR variam entre 5,2% e 5,7%, já as taxas de rejeição dos testes LR^* e LR_1^* variam entre 4,9% e 5,4%, e as dos testes LR_{boot} e LR_{boot}^* variam entre 4,5% e 5,2%. Por outro lado, neste mesmo cenário,

Tabela 4.7 Taxas de rejeição dos testes LR , LR^* , LR_1^* , LR_{boot}^* , LR_{boot} , S_r , S_r^* , $K(S_r)$ e \tilde{S}_r em MNLLGG com $\theta = 1$, $p = 5$, $n = 20$ e diversos valores de λ .

$\alpha(\%)$	λ	LR	LR^*	LR_1^*	LR_{boot}	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r
1	0,50	1,6	0,8	0,7	1,0	0,9	0,1	0,8	0,9	1,1
	0,75	1,1	0,8	0,8	0,9	0,8	0,4	1,0	1,1	1,1
	1,00	1,3	1,1	1,1	1,2	1,0	0,7	1,1	1,1	1,1
	1,50	1,0	1,0	1,0	1,2	1,1	0,8	1,0	1,0	1,0
5	0,50	6,6	4,4	4,3	5,0	4,7	1,1	5,1	5,6	6,2
	0,75	5,9	5,0	4,9	4,8	4,5	2,8	5,1	5,1	5,3
	1,00	5,7	5,1	5,0	4,7	4,5	3,9	5,3	5,3	5,3
	1,50	5,2	5,0	5,0	5,3	5,2	4,3	5,0	5,0	5,0
10	0,50	12,8	9,2	8,8	9,6	9,4	3,8	10,1	10,8	11,8
	0,75	11,1	9,4	9,3	9,7	9,6	6,5	10,4	10,6	10,8
	1,00	11,3	10,4	10,4	10,2	10,2	8,0	10,0	10,1	10,1
	1,50	10,6	10,2	10,2	10,3	10,2	9,0	10,2	10,2	10,2

as taxas de rejeição do teste S_r variam entre 3,9% e 4,7% e as taxas de rejeição dos testes S_r^* , $K(S_r)$ e \tilde{S}_r variam entre 5,1% e 5,4%.

Tabela 4.8 Taxas de rejeição dos testes LR , LR^* , LR_1^* , LR_{boot}^* , LR_{boot} , S_r , S_r^* , $K(S_r)$ e \tilde{S}_r em MNLLGG com $\theta = 1$, $p = 5$, $\lambda = 1$ e diversos valores de n .

$\alpha(\%)$	n	LR	LR^*	LR_1^*	LR_{boot}	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r
1	20	1,3	1,1	1,1	1,2	1,0	0,7	1,1	1,1	1,1
	30	1,3	1,1	1,1	1,1	0,9	0,9	1,2	1,2	1,2
	40	1,0	0,9	0,9	1,3	1,0	0,7	0,9	0,9	0,9
	50	1,1	1,0	1,0	1,1	1,0	0,7	0,9	0,9	0,9
5	20	5,7	5,1	5,0	4,7	4,5	3,9	5,3	5,3	5,3
	30	5,7	5,4	5,4	4,8	4,7	4,3	5,2	5,2	5,2
	40	5,2	4,9	4,9	5,2	5,1	4,5	5,1	5,2	5,2
	50	5,5	5,3	5,3	5,0	4,9	4,7	5,4	5,4	5,4
10	20	11,3	10,4	10,4	10,2	10,2	8,0	10,0	10,1	10,1
	30	10,8	10,3	10,3	9,2	9,2	8,8	10,2	10,2	10,2
	40	10,3	9,9	9,9	10,2	10,0	8,7	9,9	9,9	9,9
	50	10,8	10,4	10,4	10,5	10,3	9,5	10,4	10,4	10,4

Com o objetivo de analisar a influência do número de parâmetros de perturbação no desempenho dos testes, fixando o tamanho da amostra em $n = 30$ e $\lambda = 1$, consideramos os seguintes preditores:

1. $\eta_i = \beta_3 x_{i3} + \beta_4 x_{i4} + \exp(\beta_5 x_{i5}) \quad (p = 3),$
2. $\eta_i = \beta_0 + \beta_3 x_{i3} + \beta_4 x_{i4} + \exp(\beta_5 x_{i5}) \quad (p = 4),$
3. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_3 x_{i3} + \beta_4 x_{i4} + \exp(\beta_5 x_{i5}) \quad (p = 5),$
4. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \exp(\beta_5 x_{i5}) \quad (p = 6).$

Vale lembrar que a hipótese nula a ser testada é $H_0 : \beta_3 = \beta_4 = 0$, ou seja, temos fixo o número de parâmetros de interesse em $q = 2$. A Tabela 4.9 apresenta as taxas de rejeição dos testes. Podemos notar que à medida que aumentamos o número de parâmetros de perturbação ($p - 2$), as taxas de rejeição do teste LR tornam-se consideravelmente superiores ao nível de significância correspondente. Já o teste baseado na estatística S_r torna-se conservativo pois, produz taxas de rejeição bem menores do que os níveis nominais à medida que o número de parâmetros de perturbação aumenta. O teste LR_{boot} apresenta esta mesma tendência de uma maneira menos nítida. O impacto do número de parâmetros de perturbação é bem menos marcantes nos testes baseados em LR^* , LR_1^* e LR_{boot}^* , visto que suas taxas de rejeição tendem a permanecerem mais estáveis.

Tabela 4.9 Taxas de rejeição dos testes LR , LR^* , LR_1^* , LR_{boot}^* , LR_{boot} , S_r , S_r^* , $K(S_r)$ e \tilde{S}_r em MNLLGG com $\lambda = 1$, $\theta = 1$, $n = 20$ e diversos valores de p .

$\alpha(\%)$	p	LR	LR^*	LR_1^*	LR_{boot}	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r
1	3	1,1	1,0	1,0	1,3	1,1	1,7	0,0	0,3	0,0
	4	1,2	1,0	1,0	1,2	1,0	0,9	1,0	1,0	1,0
	5	1,3	1,1	1,1	1,2	1,0	0,7	1,1	1,1	1,1
	6	1,5	1,2	1,1	1,0	0,9	0,5	1,0	1,0	1,0
5	3	5,3	5,0	5,0	5,3	5,1	5,2	4,3	5,2	5,3
	4	5,6	5,3	5,3	5,0	4,9	4,3	5,4	5,4	5,4
	5	5,7	5,1	5,0	4,7	4,5	3,9	5,3	5,3	5,3
	6	6,8	5,6	5,4	4,3	4,6	3,7	4,9	5,0	5,0
10	3	10,8	10,4	10,4	10,3	10,3	9,1	9,6	10,2	10,3
	4	11,2	10,5	10,5	10,1	10,0	8,9	10,4	10,5	10,5
	5	11,3	10,4	10,4	10,2	10,2	8,0	10,0	10,1	10,1
	6	12,5	10,7	10,3	9,8	10,2	8,1	9,7	9,8	9,8

Os resultados apresentados na Tabela 4.10 foram obtidos levando em consideração a hipótese alternativa $H_1 : \beta_3 = \beta_4 \neq 0$ para $\lambda = 1$, $n = 30$, $p = 4$, $\alpha = 5\%$ e diferentes valores de $\beta_3 = \beta_4 = \beta^{(0)}$, com $\beta^{(0)}$ variando de 0,0 a 3,5. A partir desses resultados, observamos que o poder dos testes baseados nas estatísticas LR , LR_{boot} e LR_{boot}^* é ligeiramente superior aos dos testes baseados nas demais estatísticas. Também podemos notar o ganho no poder com a correção tipo-Batlett para a estatística S_r , visto que os testes baseados em suas versões corrigidas apresentam poderes superiores ao seu.

Tabela 4.10 Poder dos testes LR , LR^* , LR_1^* , LR_{boot}^* , LR_{boot} , S_r , S_r^* , $K(S_r)$ e \tilde{S}_r em MNLLGG com $\theta = 1$, $\lambda = 1$, $p = 4$, $n = 20$.

β^0	LR	LR^*	LR_1^*	LR_{boot}	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r
0,0	5,6	5,3	5,3	4,9	5,0	4,1	5,3	5,4	5,4
0,5	9,8	9,1	9,1	9,2	9,4	8,6	10,5	10,5	10,6
1,0	26,0	24,8	24,7	25,4	25,5	21,4	25,5	25,5	25,6
1,5	54,2	52,8	52,7	53,8	54,3	44,0	49,2	49,3	49,5
2,0	80,6	79,6	79,6	80,2	80,5	67,8	72,5	72,6	72,7
2,5	94,5	94,2	94,2	94,5	94,4	85,3	88,4	88,5	88,6
3,0	98,9	98,8	98,8	99,1	99,1	94,5	95,8	95,9	96,0
3,5	99,8	99,8	99,8	99,8	99,8	98,1	98,6	98,7	98,7

Na Tabela 4.11 apresentamos comparações da média e variância das estatísticas LR , LR^* , LR_1^* , LR_{boot}^* , LR_{boot} , S_r , S_r^* , $K(S_r)$ e \tilde{S}_r com as da distribuição χ_2^2 , para o modelo com $n = 20$, $\lambda = 0,5$ e $p = 4$. Os resultados mostram que as versões corrigidas das estatísticas LR e S_r apresentam médias e variâncias bem mais próximas da distribuição χ_2^2 do que as respectivas estatísticas não-corrigidas. O mesmo ocorre quando comparamos os quantis amostrais de LR , S_r e das suas versões corrigidas com os da distribuição χ_2^2 , como observado na Tabela 4.12. Por exemplo, para o quantil de ordem 85%, o qual corresponde ao valor 3,8 na distribuição χ_2^2 , corresponde ao valor 4,2 da estatística LR , ao valor 3,7 das estatísticas corrigidas LR^* , LR_1^* e LR_{boot}^* , ao valor 3,0 da estatística S_r , 3,9 da estatística S_r^* e 4,0 das estatísticas $K(S_r)$ e \tilde{S}_r .

Tabela 4.11 Média e variância da distribuição χ_2^2 e das estatísticas $LR, LR^*, LR_1^*, LR_{boot}^*, LR_{boot}, S_r, S_r^*, K(S_r)$ e \tilde{S}_r em MNLLGG com $p = 4, n = 20$ e $\lambda = 0,5$.

momentos	χ_2^2	LR	LR^*	LR_1^*	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r
média	2,000	2,213	1,992	1,968	1,983	1,657	2,047	2,087	2,112
variância	4,000	4,729	3,834	3,739	3,905	2,815	3,930	4,180	4,357

Tabela 4.12 Quantis da distribuição χ_2^2 e das estatísticas $LR, LR^*, LR_1^*, LR_{boot}^*, LR_{boot}, S_r, S_r^*, K(S_r)$ e \tilde{S}_r em MNLLGG com $p = 4, n = 20$ e $\lambda = 0,5$.

quantis (%)	χ_2^2	LR	LR^*	LR_1^*	LR_{boot}^*	S_r	S_r^*	$K(S_r)$	\tilde{S}_r
85,0	3,8	4,2	3,7	3,7	3,7	3,0	3,9	4,0	4,0
90,0	4,6	5,1	4,6	4,5	4,6	3,7	4,8	4,9	5,0
95,0	6,0	6,6	5,9	5,9	6,0	4,9	6,3	6,4	6,6
97,0	7,0	7,6	6,8	6,8	7,0	5,8	7,3	7,5	7,6
99,0	9,2	9,9	8,9	8,8	9,1	7,9	9,1	9,4	9,6
99,5	10,6	11,3	10,2	10,1	10,4	9,0	9,6	10,1	10,3

4.7 Aplicações

4.7.1 Exemplo 1

Apresentamos a seguir uma aplicação dos testes apresentados neste capítulo utilizando o conjunto de dados analisados por Freund (1983), referentes a um estudo do tamanho de lulas predadas por tubarões e atuns. A variável resposta P é o peso em libras de $n = 22$ espécimes de lulas capturadas, as covariáveis correspondem as características do bico ou boca das lulas, em polegadas, são elas: comprimento do bico (x_1), comprimento dos tentáculos (x_2), comprimento do entalhe do bico (x_3), comprimento do entalhe dos tentáculos (x_4) e largura da lula (x_5). O modelo usado é:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \theta z_i, i = 1, \dots, 22, \quad (34)$$

em que $y_i = \log(P_i)$ e $z_i \sim \text{LGG}(\lambda)$, com $\lambda = 1,9$ e $\theta = 0,12$. Os valores de θ e λ foram escolhido dentro da faixa de valores (0;2), de tal forma que minimizassem o AIC. As esti-

mativas de máximas verossimilhanças (com os respectivos erros padrão em parênteses) foram: $\hat{\beta}_0 = -2,278(0,159)$, $\hat{\beta}_1 = -5,704(1,307)$, $\hat{\beta}_2 = 0,534(0,439)$, $\hat{\beta}_3 = -0,410(0,473)$, $\hat{\beta}_4 = 0,810(1,083)$, $\hat{\beta}_5 = 1,808(0,624)$ e $\hat{\beta}_6 = 2,409(0,828)$.

Estamos interessados em testar três hipóteses, a saber: (i) $H_0 : \beta_1 = \beta_2 = 0$, (ii) $H_0 : \beta_3 = \beta_4 = 0$ e (iii) $H_0 : \beta_5 = 0$. Realizamos estes testes de hipótese utilizando as estatísticas LR , LR_{boot} , S_r , S_g e suas versões modificadas ao nível de significância α . Conforme apresentado na Tabela 4.13, ao testarmos a hipótese (i), observamos que nenhum dos testes rejeitam H_0 . Como podemos notar, as correções não têm grande efeito. Mas, ao testarmos a hipótese (ii), ao nível de significância de 10%, observamos que apenas o teste S_r não rejeita H_0 , enquanto que os demais testes (inclusive os baseados nas estatísticas score corrigidas) tem maior evidência contra a hipótese de que não há efeito dos comprimentos dos entalhes do bico e dos tentáculos no peso das lulas. Neste caso, observamos ainda que os efeitos das correções para a estatística score foram bem mais significativas. Por outro lado, para a hipótese apresentada em (iii), ao nível de significância de 1%, observamos que apenas os testes LR , LR_{boot}^* , S_g e S_g^* levam a rejeição da hipótese de nulidade do efeito da largura da lula em seu peso. Podemos observar que os testes baseados na estatística score e suas versões modificadas, assim como os testes baseados nas versões modificadas da estatística da razão de verossimilhanças e na estatística da razão de verossimilhanças bootstrap levam a não rejeição de H_0 ao nível de significância de 1%.

4.7.2 Exemplo 2

Consideramos uma segunda aplicação dos testes estudados neste capítulo utilizando o conjunto de dados que correspondem a 46 peças de metal submetidas à fadiga em ciclos analisado na Seção 3.6. A variável resposta N denota o número de ciclos até a falha da amostra de metal e, x , o trabalho por ciclo (mJ/m^3), é a variável explicativa. O modelo utilizado é dado por

$$y_i = \beta_0 + \beta_1 \exp(\beta_2/x_i) + \theta z_i, i = 1, \dots, 46, \quad (35)$$

Tabela 4.13 P-valores dos testes de hipóteses no modelo 34.

Estatísticas	H_0		
	$\beta_1 = \beta_2 = 0$	$\beta_3 = \beta_4 = 0$	$\beta_5 = 0$
LR	0,4546	0,0364	0,0097
LR^*	0,4580	0,0376	0,0101
LR_1^*	0,4580	0,0377	0,0101
LR_{boot}	0,4700	0,0300	0,0133
LR_{boot}^*	0,4740	0,0376	0,0095
S_r	0,5266	0,1106	0,0222
S_r^*	0,5100	0,0982	0,0178
$K(S_r)$	0,5098	0,0980	0,0178
\tilde{S}_r	0,5095	0,0979	0,0177
S_g	0,4530	0,0431	0,0099
S_g^*	0,4553	0,0430	0,0096

em que $z_i \sim \text{LGG}(\lambda)$, com $\lambda = 0,95$. O valor do λ foi escolhido dentro da faixa de valores (0;2), de tal forma que minimizasse o AIC. As estimativas de máximas verossimilhanças (com os respectivos erros padrão em parênteses) foram: $\hat{\beta}_0 = 9,443(0,764)$, $\hat{\beta}_1 = -5,619(0,563)$, $\hat{\beta}_2 = -19,723(5,580)$ e $\hat{\theta} = 0,321(0,037)$.

Nosso objetivo aqui é testar a hipótese $H_0 : \beta_1 = 1$ contra $H_0 : \beta_1 \neq 1$. Para estas hipóteses, todos os testes LR , LR^* , LR_1^* , LR_{boot} , LR_{boot}^* , S_r , S_r^* , $K(S_r)$ e \tilde{S}_r forneceram p -valores menores que 0,001, ou seja, todos os testes considerados levam à rejeição de H_0 . Para avaliar o efeito das correções em amostras pequenas, consideramos 26 observações selecionadas entre as 46 peças de metal submetidas à fadiga em ciclos e reajustamos o modelo (35). Nesta sub-amostra as EMV (com os respectivos erros padrão em parênteses) passam a ser: $\hat{\beta}_0 = 10,020(2,111)$, $\hat{\beta}_1 = -5,704(1,307)$ e $\hat{\beta}_2 = -15,030(11,889)$. Testando novamente $H_0 : \beta_1 = 1$ versus $H_0 : \beta_1 \neq 1$, os testes LR , LR^* , LR_1^* , LR_{boot} , LR_{boot}^* , S_r , S_r^* , $K(S_r)$ e \tilde{S}_r forneceram os seguintes p -valores: 0,009; 0,0120; 0,0121; 0,015, 0,0128, 0,0284, 0,0164, 0,0158 e 0,0153. De modo que ao nível de significância de 5%, todos os testes levam a rejeição de H_0 . No entanto, ao adotarmos o nível de significância de 1%, o teste LR leva a rejeição de H_0 , enquanto que a não rejeição de H_0 é obtida quando os testes são baseados em LR^* , LR_1^* , LR_{boot} , LR_{boot}^* , S_r , S_r^* , $K(S_r)$

e \tilde{S}_r . Em outras palavras, os testes baseados nas versões corrigidas da estatística da razão de verossimilhanças, na metodologia bootstrap, na estatística score e em suas versões corrigidas conduzem a conclusão diferente daquela obtida através do teste da razão de verossimilhanças original.

4.8 Comentários

Neste capítulo, derivamos o fator da correção de Bartlett para a estatística LR , assim como os fatores da correção tipo-Bartlett para as estatísticas S_r e S_g nos MNLLGG. Abordamos o caso em que os parâmetros λ e θ são fixos. Resultados numéricos mostraram o desempenho, em amostras finitas, dos testes baseados nas estatísticas não corrigidas e nas estatísticas corrigidas LR^* , LR_1^* , S_r^* , $K(S_r)$, \tilde{S}_r e S_g^* , também foram considerados os testes da razão de verossimilhanças bootstrap e sua versão corrigida (via Rocke).

Concluimos desses resultados que o teste da razão de verossimilhanças usual e o teste score nos MNLLGG e MLLGG podem apresentar taxas de rejeição empíricas sob H_0 bastante distorcidas quando valores de λ são pequenos e em amostras de tamanho pequeno e moderado. O teste LR mostrou ser bastante liberal, enquanto o teste score mais conservativo. Essas tendências foram acentuadas conforme aumentamos o número de parâmetros de perturbação. Nestes casos, o ganho com as correções de Bartlett e tipo-Bartlett foi evidente, uma vez que os testes corrigidos LR^* , LR_1^* , LR_{boot}^* , S_r^* , $K(S_r)$ e \tilde{S}_r apresentaram taxas de rejeição mais próxima do nível nominal correspondente do que as suas respectivas versões não corrigidas. Outra vantagem dos testes corrigidos é que, mesmo com a variação do valor do λ , do tamanho da amostra e do número de parâmetros de perturbação, as suas taxas de rejeição foram mais estáveis do que dos testes não-corrigidos. Já o teste gradiente, por sua vez, apresentou no MLLGG taxas de rejeição bastante próximas aos níveis nominais adotados, independente do valor do λ . No entanto, quando aumentamos o número de parâmetros de perturbação ou diminuimos o tamanho da amostra, esse teste apresentou tendência liberal, de forma que o uso da correção tipo-Bartlett

estabilizou as taxas de rejeição, deixando-as mais próximas aos níveis nominais considerados. Quanto ao poder, os resultados indicam que os poderes dos testes baseados nas estatísticas corrigidas foram próximos ou até mesmos superiores ao dos demais testes. No geral, os poderes dos testes baseados nas estatísticas score e suas versões corrigidas são inferiores aos dos demais testes, além de apresentarem os momentos e os quantis empíricos mais afastados da distribuição de referência χ_2^2 . Diante de todos os resultados obtidos, recomendamos para aplicações práticas o uso das versões corrigidas dos testes da razão de verossimilhanças, score e gradiente.

5

Técnicas de Diagnóstico no Modelo Não Linear Log-Gama Generalizado

5.1 Introdução

A eficácia de um modelo de regressão paramétrico depende da escolha correta da distribuição da variável aleatória dependente. Neste contexto, os resíduos do modelo de regressão tornam-se uma ferramenta importante para validar a distribuição proposta para a variável resposta, uma vez que fornece evidências sobre as observações cujo comportamento não está completamente descrito pelo modelo ajustado e avalia se a diferença entre o valor observado e explicado pelo modelo é estatisticamente significativa. Na validação do modelo, também se faz necessário verificar a existência de observações extremas que exerçam alguma interferência desproporcional nos resultados do ajuste através da análise de diagnóstico.

Os resíduos propostos para o modelo não linear log-gama generalizado (MNLGG) são baseados nas metodologias de Cox e Snell (1968), de Pregibon (1981) e a partir da função desvio. Na literatura, é vasto o uso destas e outras metodologias na análise de resíduos em modelos de regressão, por exemplo, Cysneiros e Vanegas (2008) apresentam resíduos para classe dos modelos de regressão não lineares simétricos, Silva *et al.* (2010) comparam três resíduos para o modelo de regressão log-Burr XII para dados com censuras, Vanegas *et al.* (2012) propõem

uma definição geral para os resíduos do modelo não linear Birnbaum-Saunders. Por outro lado, os procedimentos de identificação de observações influentes nas estimativas dos parâmetros para MNLLGG foram desenvolvidos com base no conceito de alavancagem generalizada sob a abordagem de Wei *et al.* (1998), em que eles generalizam a definição de pontos de alavanca para modelos gerais, na abordagem influência local proposta por Cook (1986), em que a ideia básica consiste em avaliar a influência conjunta das observações sob pequenas perturbações nos dados ou no próprio modelo, e sob enfoque de influência global (Cook, 1977). Alguns autores têm estudado os procedimentos de diagnóstico com base nestas abordagens, por exemplo, Galea *et al.* (2000), Tang *et al.* (2000), Fung *et al.* (2002), Ortega *et al.* (2003), Sun e Wei (2004), Xie e Wei (2007) e Vanegas e Cysneiros (2010).

5.2 Análise de Resíduos

Uma etapa importante na análise do ajuste de um modelo de regressão é verificar a qualidade do ajuste, o qual pode ser feita iniciando com a análise de resíduos. Esta, por sua vez, tem por objetivo avaliar a adequação da distribuição proposta para a variável resposta através da discrepância entre os valores ajustados e o conjunto de dados, como também de detectar a presença de pontos aberrantes, ou seja, pontos que interfere nos valores ajustados, embora não estejam muito afastado dos demais pontos. Considere o modelo definido por (1) e (2). Se denotarmos por $\hat{\xi} = \xi(\mathbf{y})$ a estimativa de máxima verossimilhança de ξ e por μ o vetor de valores esperados, então $\hat{y} = \widehat{E(Y)} = \mu(\hat{\xi})$ é o vetor de respostas preditas. Utilizando a metodologia apresentada em Cox e Snell (1968), propomos o resíduo de Pearson para MNLLGG baseado na diferença $y_i - \hat{y}_i$ e definido como:

$$\hat{r}_i = \frac{y_i - \hat{\mu}_i}{\sqrt{\widehat{Var}(Y)}} = \frac{y_i - \hat{\mu}_i}{\lambda \hat{\theta} \sqrt{\psi^{(1)}(\lambda^2)}} = \frac{\hat{z}_i}{\lambda \sqrt{\psi^{(1)}(\lambda^2)}},$$

em que $\hat{z}_i = (y_i - \hat{\mu}_i)/\hat{\theta}$ e $\hat{\mu}_i = \eta(x_i, \hat{\beta})$, para $i = 1, \dots, n$. Lembrando que λ^2 é o parâmetro de forma da distribuição LGG considerado fixo, e $\psi^{(1)}(\cdot)$ denota a primeira derivada da função digama.

Considerando o parâmetro θ fixo, uma segunda opção seria utilizar o resíduo definido por Pregibon (1981). Este último, é derivado do processo iterativo de estimação dos parâmetros do modelo. A partir da convergência do processo iterativo apresentado em (8), a expressão do estimador de β é dado por:

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

com $\hat{\delta} = \tilde{X} \hat{\beta} - \lambda \theta \hat{M}_1 \mathbf{1}$, $\hat{M}_1 = \text{diag}\{\hat{m}_{11}, \dots, \hat{m}_{1n}\}$ sendo uma matriz diagonal de dimensão $n \times n$ cujo i -ésimo elemento é $\hat{m}_{1i} = 1 - \exp\{\lambda^{-1} \hat{z}_i + c(\lambda)\}$ e $\hat{z}_i = (y_i - \hat{\mu}_i)/\theta$. Assim, $\hat{\beta}$ pode ser interpretado como a solução de mínimos quadrados da regressão linear de $\hat{\delta}$ contra as colunas de \tilde{X} e o resíduo ordinário é definido, a partir dessa regressão, como:

$$r^* = \hat{\delta} - \tilde{X} \hat{\beta} = (I - H) \hat{\delta},$$

em que $H = \tilde{X} (\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top$. Assumindo que $\text{Var}(\hat{\delta}) \cong \theta^2 I$, temos que aproximadamente $\text{Var}(r^*) \cong \theta^2 (I - H)$. Como $E(r^*) = 0$, o resíduo padronizado é dado por:

$$t_{S_i} = \frac{r_i^* - E(r_i^*)}{\sqrt{\text{Var}(r_i^*)}} = -\frac{\lambda m_{1i}}{\sqrt{1 - h_{ii}}}, \quad (36)$$

em que h_{ii} é o i -ésimo elemento da diagonal principal da matriz H .

Um outra proposta, considerando o parâmetro θ fixo, para avaliar a qualidade do ajuste é através dos resíduos calculados a partir da função desvio. Esta função corresponde à distância, para cada observação, entre o logaritmo da função de verossimilhança (função de log-verossimilhança) do modelo saturado (com n parâmetros) e do modelo sob investigação (com

p parâmetros). O resíduo componente do desvio nos MNLLGG é dado, então, por

$$t_{D_i} = \text{sinal}(\tilde{\mu}_i - \hat{\mu}_i) \sqrt{2} \{l(\tilde{\mu}_i) - l(\hat{\mu}_i)\}^{1/2},$$

em que $\text{sinal}(\tilde{\mu}_i - \hat{\mu}_i)$ uma função que representa o sinal da diferença entre $\tilde{\mu}_i$ e $\hat{\mu}_i$, $l(\mu_i)$ corresponde a função de log-verossimilhança dada por (3) avaliado em $\mu_i = \eta(x_i, \beta)$, $\tilde{\mu}_i$ é a estimativa de μ_i no modelo saturado, ou seja, $\tilde{\mu}_i = y_i + \theta \lambda c(\lambda)$, e $\hat{\mu}_i$ é a estimativa de μ_i no modelo investigado. Assim, t_{D_i} pode ser expresso por

$$t_{D_i} = \text{sinal}(\tilde{\mu}_i - \hat{\mu}_i) \sqrt{2} \{-\lambda^2(\hat{m}_{1i} + c(\lambda)) - \lambda \hat{z}_i\}^{1/2}.$$

É possível padronizar o resíduo componente do desvio utilizando os elementos da diagonal principal da matriz H da seguinte maneira:

$$t'_{D_i} = \frac{t_{D_i}}{\sqrt{1 - h_{ii}}}.$$

Com o objetivo de verificar as propriedades empíricas dos três resíduos propostos, \hat{r}_i , t_{S_i} e t'_{D_i} , realizamos simulações Monte Carlo com 10.000 réplicas considerando o seguinte modelo:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \exp(\beta_2 x_{2i}) + \theta Z_i, \quad i = 1, \dots, n,$$

em que $Z_i \sim LGG(\lambda)$ com λ variando entre 1, 1,5 e 2, as variáveis x_{1i} e x_{2i} foram tomadas como amostras aleatórias da distribuição uniforme $U(0, 1)$. Enquanto que a variável resposta foi gerada assumindo que $\beta_0 = \beta_1 = \beta_2 = \theta = 1$. Na primeira etapa fixamos $n = 15$, para os resíduos propostos foram consideradas as seguintes medidas: média, erro padrão (EP), assimetria (AS) e curtose (K). Os resultados apresentados na Tabela 5.1 demonstram que a média do resíduo \hat{r} é próxima de zero, em todos os cenários. Já a assimetria negativa está de acordo com a assimetria da distribuição LGG e, como era de se esperar, o aumento do valor de λ faz com que

as estimativas da medida de assimetria aproximem-se de zero. Quanto a curtose, a distribuição de resíduo \hat{r} apresenta-se leptocúrtica para $\lambda = 1$ e aproxima-se do valor da distribuição normal padrão ao passo que o valor do λ aumenta. Já os resultados para os resíduos t_{S_i} , apresentados na Tabela 5.2 mostram que a sua média é próxima de zero e a sua assimetria é positiva. Da mesma forma que o resíduo \hat{r} , a assimetria do t_{S_i} diminui com o aumento do valor de λ , enquanto que a sua curtose aproxima-se de 3. Quanto ao resíduo t'_{D_i} , segundo os resultados apresentados na Tabela 5.3, entre os três resíduos estudados, esse é o que apresenta médias mais distantes do zero, embora apresente erros padrão próximo a 1 e medidas de assimetria e curtose bem mais próxima da distribuição normal.

Tabela 5.1 Medidas descritivas das estimativas do resíduo de Pearson no MNLLGG.

Obs.	$\lambda = 1$				$\lambda = 1,5$				$\lambda = 2$			
	Média	EP	AS	K	Média	EP	AS	K	Média	EP	AS	K
1	0,09	0,87	-1,50	6,89	0,05	0,83	-0,96	4,56	0,04	0,82	-0,70	3,80
2	0,07	0,98	-1,30	5,89	0,10	0,94	-0,86	4,25	0,08	0,93	-0,65	3,92
3	0,13	0,82	-1,68	7,31	0,08	0,78	-1,14	5,37	0,06	0,76	-0,87	4,20
4	0,05	0,97	-1,29	5,64	0,05	0,93	-0,88	4,27	0,04	0,93	-0,65	3,82
5	0,08	0,97	-1,32	6,01	0,08	0,95	-0,85	4,27	0,07	0,95	-0,61	3,61
6	0,09	0,90	-1,42	6,13	0,06	0,87	-1,02	5,14	0,02	0,87	-0,72	3,81
7	0,05	0,97	-1,35	5,97	0,04	0,95	-0,89	4,40	0,03	0,93	-0,68	3,89
8	0,10	0,96	-1,48	7,81	0,08	0,96	-0,86	4,31	0,08	0,93	-0,63	3,71
9	0,11	0,91	-1,42	6,19	0,08	0,91	-1,04	5,00	0,08	0,89	-0,71	4,06
10	0,06	0,97	-1,38	6,23	0,05	0,94	-0,92	4,60	0,04	0,94	-0,59	3,59
11	0,08	0,90	-1,32	5,74	0,05	0,89	-0,89	4,20	0,03	0,91	-0,72	3,90
12	0,07	0,95	-1,34	5,71	0,07	0,92	-0,86	4,19	0,05	0,94	-0,69	3,84
13	0,08	0,97	-1,32	5,79	0,07	0,95	-0,81	4,08	0,05	0,95	-0,64	3,70
14	0,02	0,94	-1,40	6,59	0,03	0,92	-0,83	4,19	0,03	0,92	-0,70	3,92
15	0,05	0,92	-1,44	6,31	0,02	0,89	-0,97	4,66	0,01	0,88	-0,69	3,73

Na segunda etapa fixamos $n = 50$ e construímos histogramas (Figuras 5.1 – 5.3), a partir dos quais visualizamos a assimetria da distribuição dos três resíduos propostos. Enquanto a distribuição do resíduo t'_{D_i} apresenta uma diminuição na assimetria, as distribuições dos resíduos \hat{r} e t_{S_i} vão ficando mais simétricas à medida que o valor de λ aumenta. Isso ocorre porque a distribuição LGG converge para a distribuição normal quando $\lambda \rightarrow \infty$.

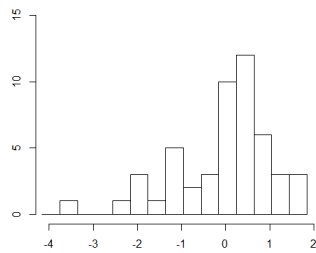
Tabela 5.2 Medidas descritivas das estimativas do resíduo padronizado no MNLLGG.

Obs.	$\lambda = 1$				$\lambda = 1,5$				$\lambda = 2$			
	Média	EP	AS	K	Média	EP	AS	K	Média	EP	AS	K
1	-0,07	0,89	0,91	3,48	-0,07	0,94	0,64	3,11	-0,06	0,96	0,52	3,26
2	0,05	0,99	1,34	4,86	0,07	1,00	0,97	4,15	0,05	1,00	0,75	3,59
3	-0,08	0,89	0,62	2,77	-0,07	0,95	0,44	2,83	-0,07	0,97	0,28	2,82
4	0,00	0,94	1,23	4,36	0,00	0,96	0,91	3,92	0,01	0,98	0,72	3,54
5	0,05	0,98	1,35	4,93	0,05	1,00	1,04	4,26	0,05	1,00	0,77	3,71
6	-0,03	0,89	1,00	3,71	-0,04	0,93	0,73	3,38	-0,05	0,96	0,55	3,17
7	0,01	0,95	1,31	4,80	0,00	0,97	0,94	3,94	0,00	0,97	0,72	3,62
8	0,06	0,98	1,30	4,66	0,05	1,00	0,95	3,89	0,05	0,99	0,74	3,62
9	0,02	0,92	1,03	3,81	0,02	0,97	0,77	3,46	0,03	0,99	0,65	3,51
10	0,01	0,94	1,26	4,46	0,00	0,95	0,95	4,05	0,01	0,99	0,77	3,68
11	-0,02	0,90	1,12	4,14	-0,02	0,94	0,76	3,40	-0,02	0,97	0,62	3,31
12	0,01	0,93	1,19	4,31	0,02	0,97	0,89	3,80	0,02	0,99	0,73	3,70
13	0,05	0,98	1,36	5,08	0,05	1,01	1,08	4,61	0,03	0,99	0,73	3,54
14	-0,07	0,88	1,27	4,61	-0,03	0,96	0,94	3,94	-0,02	0,97	0,66	3,41
15	-0,05	0,89	1,10	4,05	-0,07	0,93	0,78	3,53	-0,06	0,96	0,59	3,23

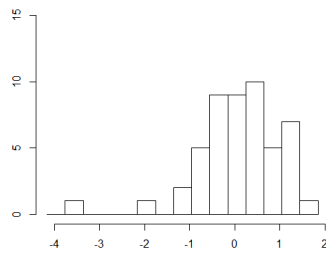
Tabela 5.3 Medidas descritivas das estimativas do resíduo componente do desvio padronizado no MNLLGG.

Obs.	$\lambda = 1$				$\lambda = 1,5$				$\lambda = 2$			
	Média	EP	AS	K	Média	EP	AS	K	Média	EP	AS	K
1	-0,37	1,07	-0,47	3,05	-0,26	1,03	-0,34	3,05	-0,20	1,02	-0,27	3,04
2	-0,28	1,06	-0,20	2,80	-0,15	1,02	-0,18	2,92	-0,10	1,01	-0,13	3,01
3	-0,37	1,12	-0,67	3,32	-0,26	1,07	-0,51	3,28	-0,20	1,05	-0,45	3,21
4	-0,33	1,05	-0,25	2,77	-0,20	1,01	-0,21	2,93	-0,15	1,01	-0,15	2,98
5	-0,27	1,04	-0,20	2,84	-0,16	1,02	-0,14	2,99	-0,11	1,01	-0,12	2,92
6	-0,33	1,05	-0,40	2,92	-0,23	1,02	-0,32	3,08	-0,20	1,02	-0,26	2,99
7	-0,32	1,05	-0,24	2,85	-0,21	1,02	-0,20	2,92	-0,16	1,00	-0,17	3,03
8	-0,26	1,04	-0,23	2,91	-0,17	1,03	-0,17	2,89	-0,10	1,00	-0,14	2,97
9	-0,29	1,05	-0,37	2,89	-0,18	1,04	-0,32	3,08	-0,12	1,02	-0,21	3,07
10	-0,31	1,04	-0,25	2,86	-0,21	1,01	-0,21	3,01	-0,15	1,01	-0,11	2,92
11	-0,32	1,03	-0,31	2,85	-0,22	1,01	-0,26	2,88	-0,17	1,02	-0,23	3,00
12	-0,30	1,04	-0,30	2,83	-0,19	1,01	-0,20	2,90	-0,14	1,02	-0,18	3,03
13	-0,27	1,04	-0,22	2,85	-0,17	1,02	-0,12	2,95	-0,13	1,01	-0,14	2,93
14	-0,38	1,03	-0,28	2,93	-0,24	1,02	-0,17	2,93	-0,17	1,01	-0,20	3,00
15	-0,36	1,05	-0,38	2,95	-0,27	1,02	-0,30	3,03	-0,21	1,02	-0,23	2,97

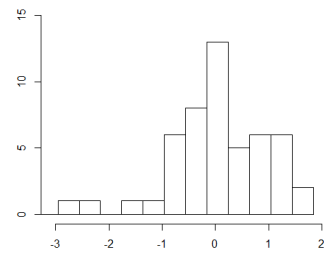
Figura 5.1 Histogramas das estimativas do resíduo de Pearson.



(a) $\lambda = 1$



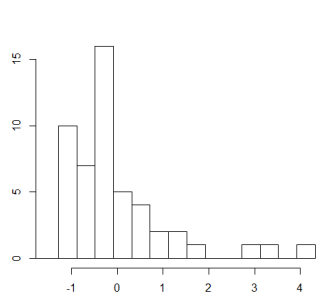
(b) $\lambda = 1,5$



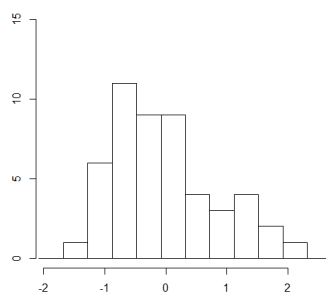
(c) $\lambda = 2$

Fonte: Autoria própria

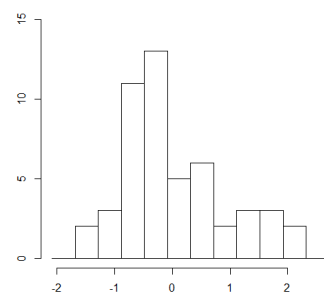
Figura 5.2 Histogramas das estimativas do resíduo padronizado.



(a) $\lambda = 1$



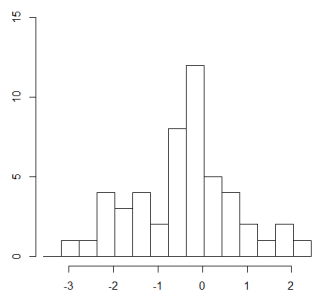
(b) $\lambda = 1,5$



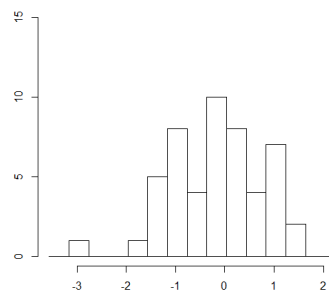
(c) $\lambda = 2$

Fonte: Autoria própria

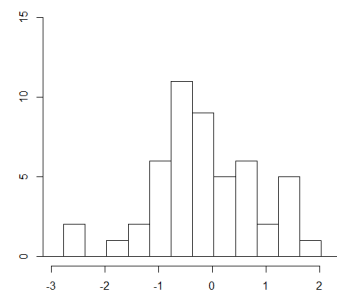
Figura 5.3 Histogramas das estimativas do resíduo componente do desvio padronizado.



(a) $\lambda = 1$



(b) $\lambda = 1,5$



(c) $\lambda = 2$

Fonte: Autoria própria

5.3 Análise de Diagnóstico

O objetivo na análise de diagnóstico é verificar a existência de observações que exerçam um efeito desproporcional nas estimativas dos coeficientes do modelo, levando muitas vezes a conclusões errôneas quanto aos resultados inferenciais. Estas observações são conhecidas como pontos influentes e são caracterizadas por terem um perfil diferente dos demais pontos no que diz respeito aos valores das variáveis explicativas.

5.3.1 Alavancagem Generalizada

O conceito de alavancagem consiste em conhecer a influência que cada observação y_i exerce no próprio valor ajustado \hat{y}_i e é apresentada pela derivada $\partial \hat{y}_i / \partial y_i$. No caso do modelo normal linear, essa influência coincide com o i -ésimo elemento da diagonal principal da matriz de projeção $H = X(X^\top X)^{-1}X^\top$, sendo X a matriz modelo. No entanto, Wei *et al.* (1998) propuseram uma expressão para alavancagem em modelos mais gerais. Segundo estes, alavancagem generalizada de $\hat{\xi}$ pode ser definida como:

$$GL(\hat{\xi}) = \{(D_\xi)(-\ddot{L}_{\xi\xi})^{-1}(\ddot{L}_{\xi y})\}|_{\xi=\hat{\xi}}, \quad (37)$$

em que $D_\xi = \partial \mu / \partial \xi^\top$, $\ddot{L}_{\xi\xi} = \partial^2 l(\xi) / \partial \xi \partial \xi^\top$ e $\ddot{L}_{\xi y} = \partial^2 l(\xi) / \partial \xi \partial y^\top$.

No caso dos MNLLGG, temos que $\xi = (\beta^\top, \theta)^\top$, logo $D_\xi = (D_\beta \ D_\theta) = (\tilde{X} \ \mathbf{0}_1)$, em que $\tilde{X} = \tilde{X}(\beta) = \partial \eta / \partial \beta^\top$ é a matriz de derivadas $n \times p$ e $\mathbf{0}_1$ é um vetor de zeros de tamanho n . A matriz $\ddot{L}_{\xi\xi}$ é definida em (4), enquanto que a matriz $\ddot{L}_{\xi y}$ é dada por:

$$\ddot{L}_{\xi y} = \begin{pmatrix} \ddot{L}_{\beta y} \\ \ddot{L}_{\theta y} \end{pmatrix} = \frac{1}{\theta^2} \begin{pmatrix} \tilde{X}^\top (I - M_1) \\ z^\top (I - M_1) - \lambda \mathbf{1}^\top M_1 \end{pmatrix},$$

em que $M_1 = \text{diag}\{m_{11}, \dots, m_{1n}\}$ sendo uma matriz diagonal de dimensão $n \times n$ cujo i -ésimo elemento é $m_{1i} = 1 - \exp\{\lambda^{-1}z_i + c(\lambda)\}$ e $z = (z_1, \dots, z_n)^\top$ é um vetor de dimensão $n \times 1$.

Então, podemos simplificar (37) da seguinte forma:

$$GL(\hat{\xi}) = -\frac{1}{\theta^2} \tilde{X} \ddot{L}^{\beta\beta} \tilde{X}^\top (I - M_1) - \frac{1}{\theta^2} \tilde{X} \ddot{L}^{\beta\theta} z^\top (I - M_1) + \frac{\lambda}{\theta^2} \tilde{X} \ddot{L}^{\beta\theta} \mathbf{1}^\top M_1 |_{\xi=\hat{\xi}},$$

em que $\ddot{L}^{\beta\beta}$ e $\ddot{L}^{\beta\theta}$ correspondem as submatrizes do bloco superior de $\ddot{L}_{\xi\xi}^{-1}$ de dimensão $p \times (p+1)$. Ou seja, $\ddot{L}^{\beta\beta} = (\ddot{L}_{\beta\beta}^{-1} + FE^{-1}F^\top)$ e $\ddot{L}^{\beta\theta} = -\ddot{L}^{\beta\beta} \ddot{L}_{\beta\theta} \ddot{L}_{\theta\theta}^{-1}$, em que $F = \ddot{L}_{\beta\beta}^{-1} \ddot{L}_{\beta\theta}$ e $E = \ddot{L}_{\theta\theta} - \ddot{L}_{\theta\beta} \ddot{L}_{\beta\beta}^{-1} \ddot{L}_{\beta\theta}$. Desta forma, a matriz de alavancagem generalizada é dada por:

$$GL(\hat{\xi}) = GL_{\beta}(\hat{\xi}) + GL_{\beta,\theta}(\hat{\xi}),$$

em que

$$GL_{\beta}(\hat{\xi}) = -\frac{1}{\theta^2} \tilde{X} \ddot{L}_{\beta\beta}^{-1} \tilde{X}^\top (I - M_1) |_{\xi=\hat{\xi}}$$

e

$$GL_{\beta,\theta}(\hat{\xi}) = -\frac{1}{\theta^2} \tilde{X} F E^{-1} F^\top \tilde{X}^\top (I - M_1) + \frac{1}{\theta^2} \tilde{X} (\ddot{L}_{\beta\beta}^{-1} - F E^{-1} F^\top) \ddot{L}_{\beta\theta} \ddot{L}_{\theta\theta}^{-1} z^\top (I - M_1) - \lambda \mathbf{1}^\top M_1 |_{\xi=\hat{\xi}}.$$

Quando o parâmetro θ for considerado fixo, o estimador do vetor de parâmetros $\xi = (\beta^\top, \theta)^\top$ será igual $\hat{\xi} = (\hat{\beta}^\top, \theta)$ e a matriz de alavancagem generalizada reduzir-se-á a $GL(\hat{\xi}) = GL_{\beta}(\hat{\xi})$. Um gráfico dos elementos da diagonal de $GL(\hat{\xi})$ versus índices pode revelar pontos com alta influência no seu próprio valor predito.

5.3.2 Influência Local

Com o objetivo de verificar a existência de pontos que após modificações pequenas no modelo e/ou nos dados causam variações desproporcionais nas conclusões inferenciais, Cook (1986) propôs um método de influência local, o qual avalia a influência conjunta de pontos sob pequenas perturbações nas estimativas de máxima verossimilhança.

Seja $l(\xi)$ a função de log-verossimilhança dada em (3) e o vetor de perturbação $w =$

$(w_1, \dots, w_n)^\top$, restrito a algum subconjunto Ω do \mathbb{R}^n . A fim de avaliar os efeitos das perturbações nas estimativas fornecidas pelo modelo, o afastamento de verossimilhança, também conhecido como *desvio local* é definido como:

$$LD(w) = 2\{l(\hat{\xi}) - l(\hat{\xi}_w)\},$$

em que $LD(w) \geq 0$, $\hat{\xi}$ e $\hat{\xi}_w$ são os estimadores de máxima verossimilhança (EMV) de ξ dos modelos não perturbado e perturbado, respectivamente. Considere w_0 o vetor de não perturbação, ou seja, $l(\hat{\xi}_{w_0}) = l(\hat{\xi})$. Temos que w_0 é o ponto mínimo da função $LD(w)$, uma vez que $LD(w_0) = 0$. Cook (1986) sugere estudar o comportamento do $LD(w)$ em torno de w_0 através da curvatura normal da linha projetada $LD(w_0 + ad)$, em que $a \in \mathbb{R}$ e d é um vetor fixo não nulo unitário, ou seja, $\|d\| = 1$. Cook (1986) mostra que a curvatura normal $C_d(\xi)$ possui a seguinte forma geral:

$$C_d(\xi) = 2|d^\top \Delta^\top \ddot{L}_{\xi\xi}^{-1} \Delta d|,$$

em que Δ é uma matriz $(p+1) \times n$ que depende do esquema de perturbação, dada por $\Delta = \partial^2 l(\hat{\xi}_w) / \partial \xi \partial w^\top$ avaliada em $\xi = \hat{\xi}$ e $w = w_0$.

Uma sugestão de Cook (1986) é utilizar d_{max} , o autovetor correspondente ao maior autovalor $C_{d_{max}}$ da matriz $\ddot{F} = \Delta^\top (-\ddot{L}_{\xi\xi})^{-1} \Delta$, como medida de influência local. O gráfico de d_{max} contra a ordem das observações pode revelar quais os pontos que, sob pequenas perturbações, exercem maior influência em $\hat{\xi}$. Outra possibilidade de análise foi proposta por Escobar e Meeker (1992), o qual sugerem que a medida de influência a ser tomada seja os elementos da diagonal principal da matriz \ddot{F} . Lesaffre e Verbeke (1998) sugerem que a curvatura normal seja avaliada na direção da i -ésima observação, que consiste em avaliarmos $C_d(\xi)$ no vetor d_i de dimensão $n \times 1$ composto por um na i -ésima linha e zero nas demais. Essa curvatura é denominada por C_i que é igual a $2|f_{ii}|$, em que f_{ii} é o i -ésimo elemento da diagonal principal da matriz \ddot{F} . É sugerido um gráfico de C_i versus índice, em que as observações com $C_i > 2\bar{C}$, sendo $\bar{C} = \sum_{j=1}^n C_j / n$, merecem atenção especial.

Neste estudo, apresentamos três tipos de esquema de perturbação no MNLLGG definido em (1) e (2). Obteremos para cada esquema de perturbação a matriz Δ avaliada em $\xi = \hat{\xi}$ e $w = w_0$, apresentada sob a forma:

$$\Delta = \begin{pmatrix} \Delta_\beta \\ \Delta_\theta \end{pmatrix} = \begin{pmatrix} \partial^2 l(\xi_w) / \partial \beta \partial w^\top \\ \partial^2 l(\xi_w) / \partial \theta \partial w^\top \end{pmatrix}, \quad \hat{\Delta} = \Delta|_{\xi=\hat{\xi}}.$$

Perturbação Aditiva na Resposta

Com objetivo de analisar a sensibilidade do modelo quando a variável resposta é submetida a uma perturbação aditiva, considere o esquema $y_{iw} = y_i + w_i s_y$, em que $i = 1, \dots, n$, s_y é um fator de escala, utilizado com o intuito de padronizar os componentes de w e pode ser estimado pelo desvio padrão da variável resposta. Assim, a função de log-verossimilhança do modelo perturbado assume a forma:

$$l(\xi_w) = (2\lambda^2 - 1)n \log \lambda - n \log \Gamma(\lambda^2) + \lambda \sum_{i=1}^n z_{iw} + n\lambda^2 c(\lambda) - \lambda^2 \sum_{i=1}^n \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} - n \log \theta,$$

em que $z_{iw} = (y_i + w_i s_y - \eta_i) / \theta$, sendo $\eta_i = \eta(x_i, \beta)$. Derivando $l(\xi_w)$ em relação à w_i , $i = 1, \dots, n$, obtemos:

$$\frac{\partial l(\xi_w)}{\partial w_i} = \frac{\partial l(\xi_w)}{\partial z_{iw}} \frac{\partial z_{iw}}{\partial w_i} = \left[\lambda - \lambda \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} \right] \frac{s_y}{\theta}. \quad (38)$$

Derivando (38) em relação ao β_s , $s = 1, \dots, p$, obtemos:

$$\begin{aligned} \frac{\partial^2 l(\xi_w)}{\partial \beta_s \partial w_i} &= -\frac{s_y}{\theta} \lambda \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} \frac{1}{\lambda} \frac{\partial z_{iw}}{\partial \beta_s} \\ &= \frac{s_y}{\theta^2} \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} \tilde{x}_{is} \\ &= \frac{s_y}{\theta^2} (1 - m_{1iw}) \tilde{x}_{is}, \end{aligned}$$

em que $m_{1iw} = 1 - \exp\{\lambda^{-1}z_{iw} + c(\lambda)\}$ e $\tilde{x}_{is} = \partial\eta_i/\partial\beta_s$, $i = 1, \dots, n$. Ao derivar (38) em relação ao θ obtemos:

$$\begin{aligned} \frac{\partial^2 l(\xi_w)}{\partial\theta\partial w_i} &= -\lambda \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} \frac{1}{\lambda} \frac{\partial z_{iw}}{\partial\theta} \frac{s_y}{\theta} + \left[\lambda - \lambda \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\}\right] \left(-\frac{s_y}{\theta^2}\right) \\ &= -\exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} \frac{z_{iw}s_y}{\theta^2} - \lambda \left[1 - \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\}\right] \frac{s_y}{\theta^2} \\ &= -\{(1 - m_{1iw})z_{iw} + \lambda m_{1iw}\} \frac{s_y}{\theta^2}, \quad i = 1, \dots, n. \end{aligned}$$

Para a perturbação aditiva na resposta, $w_0 = (0, \dots, 0)^\top$ e a matriz Δ avaliada em $\xi = \hat{\xi}$ e $w = w_0$ fica da forma:

$$\hat{\Delta} = \begin{pmatrix} \hat{\Delta}_\beta \\ \hat{\Delta}_\theta \end{pmatrix} = \frac{s_y}{\hat{\theta}^2} \begin{pmatrix} \hat{X}^\top (I - \hat{M}_1) \\ \hat{z}^\top (I - \hat{M}_1) - \lambda \mathbf{1}^\top \hat{M}_1 \end{pmatrix} = \frac{s_y}{\hat{\theta}^2} \hat{X}^{*\top} \hat{M}_3,$$

em que $\hat{M}_3 = (I - \hat{M}_1, \hat{z}^\top (I - \hat{M}_1) - \lambda \mathbf{1}^\top \hat{M}_1)^\top$, as matrizes \hat{X}^* e \hat{M}_1 correspondem respectivamente às matrizes \tilde{X}^* e M_1 avaliadas em $\xi = \hat{\xi}$, assim como o vetor \hat{z} é formado pelos elementos $\hat{z}_i = (y_i - \hat{\eta}_i)/\hat{\theta}$, $i = 1, \dots, n$.

No caso em que $s_y = 1$, existe uma relação direta da matriz $\ddot{F} = 2\Delta^\top (-\ddot{L}_{\xi\xi})^{-1} \Delta$ com a matriz de alavancagem generalizada $GL(\xi)$. Note que neste caso, $\Delta = \ddot{L}_{\xi y}$. Definimos o bloco inferior de $\ddot{L}_{\xi\xi}^{-1}$, de dimensão $1 \times (p+1)$, como $\ddot{L}^{\theta*} = (\ddot{L}^{\theta\beta} \ddot{L}^{\theta\theta})$. Com isso podemos mostrar que:

$$\ddot{F} = 2\Delta^\top (-\ddot{L}_{\xi\xi})^{-1} \Delta = \frac{2}{\theta^2} (I - M_1) GL(\xi) - 2\Delta_\theta^\top \ddot{L}^{\theta*} \Delta.$$

Em particular, quando parâmetro θ for considerado fixo, teremos:

$$\ddot{F} = 2\theta^{-2} (I - \hat{M}_1) GL(\hat{\xi}).$$

Perturbação de Casos Ponderados

Para verificar qual a influência exercida pela i -ésima observação no modelo, considere agora um esquema de perturbação no qual a função de log-verossimilhança do modelo perturbado seja expressa na forma:

$$\begin{aligned} l(\xi_w) &= \sum_{i=1}^n w_i l(y_i; \xi) \\ &= (2\lambda^2 - 1) \log \lambda \sum_{i=1}^n w_i - \log \Gamma(\lambda^2) \sum_{i=1}^n w_i + \lambda \sum_{i=1}^n w_i z_i + \lambda^2 c(\lambda) \sum_{i=1}^n w_i \\ &\quad - \lambda^2 \sum_{i=1}^n w_i \exp \left\{ \frac{z_i}{\lambda} + c(\lambda) \right\} - \sum_{i=1}^n w_i \log \theta. \end{aligned}$$

Derivando $l(\xi_w)$ em relação à w_i , $i = 1, \dots, n$, obtemos:

$$\frac{\partial l(\xi_w)}{\partial w_i} = (2\lambda^2 - 1) \log \lambda - \log \Gamma(\lambda^2) + \lambda z_i + \lambda^2 c(\lambda) - \lambda^2 \exp \left\{ \frac{z_i}{\lambda} + c(\lambda) \right\} - \log \theta.$$

Consequentemente, para $s = 1, \dots, p$, temos que:

$$\begin{aligned} \frac{\partial^2 l(\xi_w)}{\partial \beta_s \partial w_i} &= \left[\lambda - \lambda \exp \left\{ \frac{z_i}{\lambda} + c(\lambda) \right\} \right] \left(-\frac{\tilde{x}_{is}}{\theta} \right) \\ &= -\frac{\lambda}{\theta} m_{1i} \tilde{x}_{is}, \quad i = 1, \dots, n. \end{aligned}$$

Além disso,

$$\frac{\partial^2 l(\xi_w)}{\partial \theta \partial w_i} = \left[\lambda - \lambda \exp \left\{ \frac{z_i}{\lambda} + c(\lambda) \right\} \right] \left(-\frac{z_i}{\theta} \right) - \frac{1}{\theta} = -\frac{\lambda}{\theta} m_{1i} z_i - \frac{1}{\theta} = -\frac{\lambda}{\theta} m_{2i},$$

com $m_{2i} = m_{1i} z_i + 1/\lambda$, $i = 1, \dots, n$. Para a perturbação de casos ponderados, $w_0 = (1, \dots, 1)^\top$ e a matriz Δ avaliada em $\xi = \hat{\xi}$ e $w = w_0$ fica da forma:

$$\hat{\Delta} = -\frac{\lambda}{\hat{\theta}} \hat{X}^{*\top} \hat{M},$$

em que \hat{M} corresponde a matriz $M = (M_1, m_2^\top)^\top$ avaliada em $\xi = \hat{\xi}$, sendo $m_2 = (m_{21}, \dots, m_{2n})^\top$ um vetor de tamanho n .

Perturbação Aditiva no Preditor

Agora, o objetivo é avaliar a sensibilidade do modelo à pequenas perturbações em uma determinada variável explicativa x_r . Considere o esquema em que é adicionado à x_r um vetor de perturbação w ponderado por um fator escala s_{x_r} , que pode ser o desvio padrão da variável em questão. Ou seja,

$$x_{irw} = x_{ir} + w_i s_{x_r}, \quad i = 1, \dots, n.$$

Seja $\eta_{iw} = \eta(x_{iw}, \beta)$ com $x_{iw} = (x_{i1}, \dots, x_{irw}, \dots, x_{ik})^\top$. A função de log-verossimilhança para o MNLLGG com perturbação aditiva na r -ésima variável explicativa é expressa por:

$$l(\xi_w) = (2\lambda^2 - 1)n \log \lambda - n \log \Gamma(\lambda^2) + \lambda \sum_{i=1}^n z_{iw} + n\lambda^2 c(\lambda) - \lambda^2 \sum_{i=1}^n \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} - n \log \theta,$$

em que $z_{iw} = (y_i - \eta_{iw})/\theta$, $i = 1, \dots, n$. Derivando $l(\xi_w)$ em relação à w_i , $i = 1, \dots, n$, obtemos:

$$\begin{aligned} \frac{\partial l(\xi_w)}{\partial w_i} &= \frac{\partial l(\xi_w)}{\partial z_{iw}} \frac{\partial z_{iw}}{\partial \eta_{iw}} \frac{\partial \eta_{iw}}{\partial w_i} \\ &= \left[\lambda - \lambda \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} \right] \left(-\frac{1}{\theta}\right) f_{1i}, \end{aligned} \quad (39)$$

em que $f_{1i} = \partial \eta_{iw} / \partial w_i$. Derivando (39) em relação ao β_s , $s = 1, \dots, p$, obtemos:

$$\begin{aligned} \frac{\partial^2 l(\xi_w)}{\partial \beta_s \partial w_i} &= \frac{\lambda}{\theta} \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} \frac{1}{\lambda} \frac{\partial z_{iw}}{\partial \eta_{iw}} \frac{\partial \eta_{iw}}{\partial \beta_s} f_{1i} - \frac{\lambda}{\theta} \left[1 - \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\}\right] \frac{\partial f_{1i}}{\partial \beta_s} \\ &= \frac{1}{\theta} \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} \left(-\frac{1}{\theta}\right) \tilde{x}_{isw} f_{1i} - \frac{\lambda}{\theta} \left[1 - \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\}\right] f_{2is} \\ &= -\frac{1}{\theta^2} (1 - m_{1iw}) \tilde{x}_{isw} f_{1i} - \frac{\lambda}{\theta} m_{1iw} f_{2is}, \end{aligned}$$

em que $m_{1iw} = 1 - \exp\{z_{iw}/\lambda + c(\lambda)\}$, $\tilde{x}_{isw} = \partial\eta_{iw}/\partial\beta_s$ e $f_{2is} = \partial f_{1i}/\partial\beta_s$, $i = 1, \dots, n$. Analogamente, ao derivar (39) em relação ao θ obtemos:

$$\begin{aligned} \frac{\partial^2 l(\xi_w)}{\partial\theta\partial w_i} &= -\lambda \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} \frac{1}{\lambda} \frac{\partial z_{iw}}{\partial\theta} \left(-\frac{1}{\theta}\right) f_{1i} + \left[\lambda - \lambda \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\}\right] \frac{1}{\theta^2} f_{1i} \\ &= \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\} \left(-\frac{z_{iw}}{\theta}\right) \left(\frac{1}{\theta}\right) f_{1i} + \frac{\lambda}{\theta^2} \left[1 - \exp\left\{\frac{z_{iw}}{\lambda} + c(\lambda)\right\}\right] f_{1i} \\ &= -(1 - m_{1iw}) \frac{z_{iw}}{\theta^2} f_{1i} + \frac{\lambda}{\theta^2} m_{1iw} f_{1i}, \quad i = 1, \dots, n. \end{aligned}$$

Considere as matrizes $F_1 = \text{diag}\{f_{11}, \dots, f_{1n}\}$, F_2 de dimensão $n \times p$, cujo (i, s) -ésimo elemento é f_{2is} , $M_4 = (I, z^\top)^\top$ e $M_5 = (\theta F_2^\top, -\mathbf{1}^\top F_1)^\top$. Para perturbação aditiva no preditor, $w_0 = (0, \dots, 0)^\top$. A matriz Δ avaliada em $\xi = \hat{\xi}$ e $w = w_0$ dependerá da forma de $\eta(x_{iw}, \beta)$ através das quantidades f_{1i} e f_{2is} e será dada por:

$$\hat{\Delta} = \begin{pmatrix} \hat{\Delta}_\beta \\ \hat{\Delta}_\theta \end{pmatrix} = -\frac{1}{\hat{\theta}^2} \begin{pmatrix} \hat{X}^\top (I - \hat{M}_1) F_1 + \lambda \hat{\theta} F_2^\top \hat{M}_1 \\ \hat{z}^\top (I - \hat{M}_1) F_1 - \lambda \mathbf{1}^\top \hat{F}_1 \hat{M}_1 \end{pmatrix} = -\frac{1}{\hat{\theta}^2} \hat{X}^{*\top} \hat{M}_4 (I - \hat{M}_1) \hat{F}_1 - \frac{\lambda}{\hat{\theta}^2} \hat{M}_5 \hat{M}_1,$$

em que \hat{F}_1 , \hat{F}_2 , \hat{M}_4 e \hat{M}_5 correspondem as matrizes F_1 , F_2 , M_4 e M_5 avaliadas em $\xi = \hat{\xi}$, as quais dependerão da não linearidade da função $\eta(x_i, \beta)$.

5.3.3 Influência Global

Quando o parâmetro θ é fixo, ainda podemos utilizar uma outra medida bastante conhecida para avaliar o impacto nas estimativas dos parâmetros de regressão quando a i -ésima observação é excluída, o afastamento da verossimilhança $LD_i = 2\{l(\hat{\beta}) - l(\hat{\beta}_{(i)})\}$, em que $\hat{\beta}_{(i)}$ corresponde ao estimador $\hat{\beta}$ quando a i -ésima observação é excluída. Essa medida foi introduzida por Cook (1977) e é também conhecida como influência global.

Uma vez que não é possível obter uma forma analítica para LD_i , é comum utilizar a segunda

aproximação em série de Taylor em torno de $\hat{\beta}$, obtendo:

$$LD_i \cong (\beta - \hat{\beta})^\top \{-\ddot{L}_{\beta\beta}(\hat{\beta})\}(\beta - \hat{\beta}).$$

Substituindo $-\ddot{L}_{\beta\beta}(\hat{\beta})$ pelo seu valor esperado $K_\beta(\hat{\beta})$ e β por $\hat{\beta}_{(i)}$, temos que:

$$LD_i^\beta \cong \theta^{-2}(\hat{\beta} - \hat{\beta}_{(i)})^\top (\hat{X}^\top \hat{X})(\hat{\beta} - \hat{\beta}_{(i)}), \quad (40)$$

em que \hat{X} corresponde a matriz \tilde{X} avaliada em $\hat{\beta}$.

Em geral, não é possível obter uma forma fechada para $\hat{\beta}_{(i)}$ e, então, aproximações têm sido utilizadas. Pregibon(1981) sugere a aproximação de um passo, que consiste em tomar a primeira iteração do processo iterativo pelo método scoring de Fisher, iniciando em $\hat{\beta}$. Tal aproximação é dada por:

$$\hat{\beta}_{(i)}^1 = \hat{\beta} + \{-\ddot{L}_{\beta\beta}(\hat{\beta})\}^{-1} l_{(i)}(\hat{\beta}),$$

em que $l_{(i)}(\hat{\beta})$ é a função de log-verossimilhança sem a i -ésima observação. Substituindo novamente $-\ddot{L}_{\beta\beta}(\hat{\beta})$ por $K_\beta(\hat{\beta})$ de acordo com os cálculos apresentados no Apêndice E, temos:

$$\hat{\beta}_{(i)} = \hat{\beta} - \frac{\hat{r}_i^*}{(1 - \hat{h}_{ii})} (\hat{X}^\top \hat{X})^{-1} \hat{x}_i, \quad (41)$$

em que \hat{x}_i^\top corresponde a i -ésima linha matriz \hat{X} .

Substituindo (41) em (40), obtemos:

$$\begin{aligned} LD_i^\beta &\cong \left[\frac{\hat{r}_i^*}{(1 - \hat{h}_{ii})} (\hat{X}^\top \hat{X})^{-1} \hat{x}_i \right]^\top \left(\frac{1}{\theta^2} \hat{X}^\top \hat{X} \right) \left[\frac{\hat{r}_i^*}{(1 - \hat{h}_{ii})} (\hat{X}^\top \hat{X})^{-1} \hat{x}_i \right] \\ &= \frac{\hat{r}_i^{*2}}{\theta^2 (1 - \hat{h}_{ii})^2} \hat{x}_i^\top (\hat{X}^\top \hat{X}) \hat{x}_i \\ &= \frac{\hat{r}_i^{*2}}{\theta^2 (1 - \hat{h}_{ii})^2} \hat{h}_{ii}, \end{aligned}$$

com $\hat{h}_{ii} = \hat{x}_i^\top \left(\hat{X}^\top \hat{X} \right)^{-1} \hat{x}_i$. Logo,

$$LD_i^\beta \cong \frac{\hat{h}_{ii}}{(1 - \hat{h}_{ii})} \hat{t}_{S_i}^2, \quad (42)$$

em que t_{S_i} encontra-se definido em (36). Aqui, é adotada a aproximação definida em (42) como a distância de Cook para a classe dos MNLLGG.

A aproximação de um passo, em geral, subestima o verdadeiro valor de LD_i , no entanto é suficiente para chamar a atenção de pontos aberrantes e influentes. Gráficos de LD_i^β contra os índices das observações e de \hat{h}_{ii} contra os valores ajustados são recomendados.

5.4 Aplicação

Nesta seção, consideramos novamente o conjunto de dados que correspondem a 46 peças de metal submetidas à fadiga em ciclos analisados na Seção 3.6. A variável resposta N denota o número de ciclos até a falha da amostra de metal e, x , o trabalho por ciclo (mJ/m^3), é a variável explicativa. Entre os modelos analisados na Seção 3.6, vimos que o modelo MNLLGG ('Modelo III') é o que melhor se ajusta aos dados segundo a metodologia de validação cruzada, o AIC e o BIC, o mesmo sendo expresso por

$$y_i = \beta_0 + \beta_1 \exp(\beta_2/x_i) + \theta_{z_i}, \quad i = 1, \dots, 46, \quad (43)$$

em que $z_i \sim \text{LGG}(\lambda)$, com $\lambda = 0,95$. O valor do λ foi escolhido dentro da faixa de valores (0;2), de tal forma que minimizasse o AIC. As EMV (com os respectivos erros padrão em parênteses) foram: $\hat{\beta}_0 = 9,443(0,764)$, $\hat{\beta}_1 = -5,619(0,563)$, $\hat{\beta}_2 = -19,723(5,580)$ e $\hat{\theta} = 0,321(0,037)$. Nos gráficos dos resíduos, apresentados na Figura 5.4 (a,e), a observação 5 aparece como possível ponto aberrante nos gráficos dos resíduos \hat{r} e t'_D , este último ainda apresenta a observação 12 com um alto resíduo. Já as observações 4, 25 e 39, aparecem como possíveis pontos aberran-

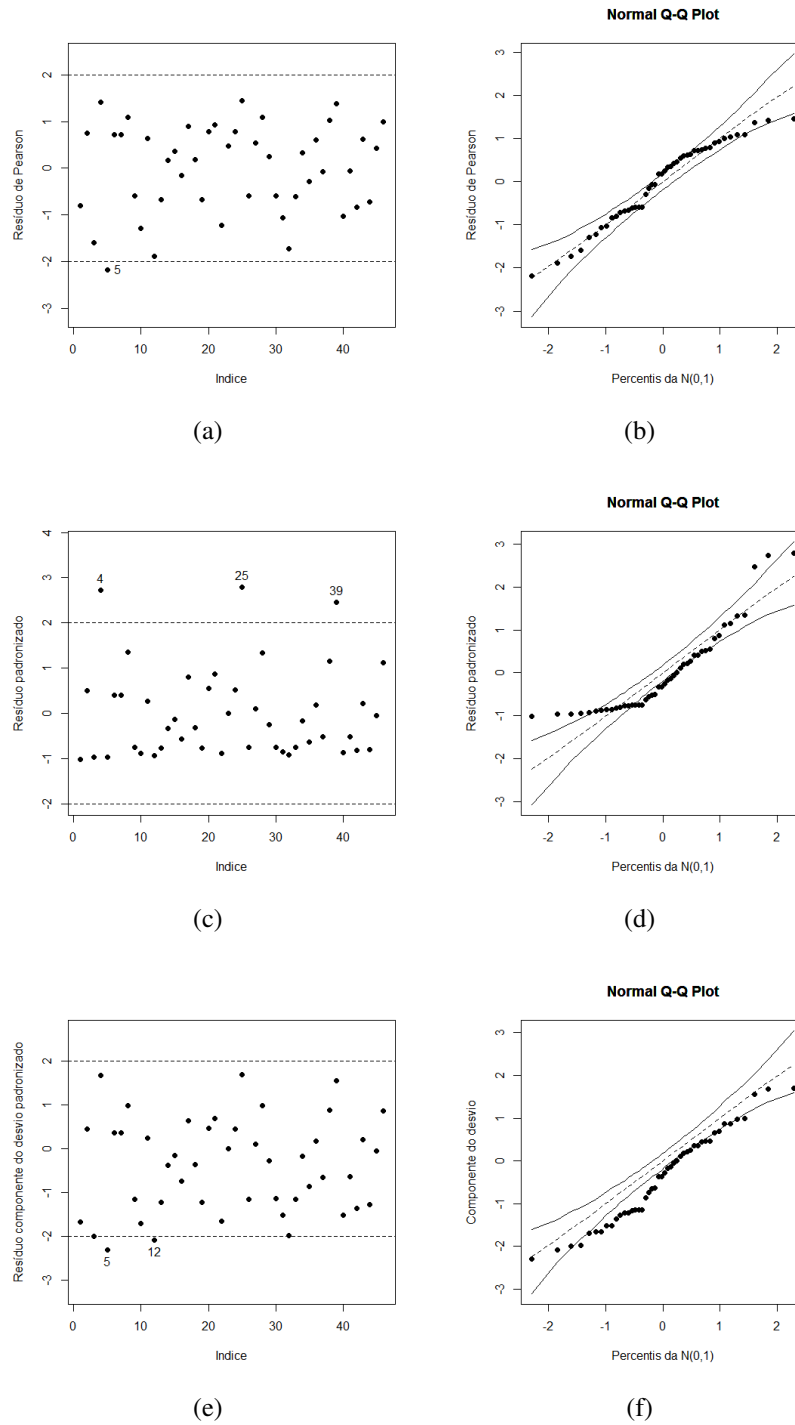
tes no gráfico do resíduo t_S (ver Figura 5.4 (c)). Analisando os três resíduos, não visualizamos presença de heteroscedasticidade. Por outro lado, através dos envelopes apresentados para os resíduos t_S e t'_D (Figuras 5.4 (d,f)), temos indícios de afastamento das suposições do modelo, o que não é visto no envelope do resíduo \hat{r} (Figura 5.4 (b)).

As observações 2, 4 e 46 aparecem com alta alavancagem na Figura 5.5, conseqüentemente para os valores próximos a estes pontos teremos uma menor precisão em suas previsões. Considerando os três esquemas de perturbação: perturbação aditiva na resposta, perturbação de casos ponderados e perturbação aditiva no preditor; construímos os gráficos de influência local C_i , apresentados na Figura 5.6, contra os índices das observações. Podemos notar que, em todos os esquemas de perturbação estudados, a observação 4 é destacada. A Figura 5.6(a) indica que pequenas mudanças nos valores da variável resposta podem produzir mudanças substanciais nas predições das observações 2, 4 e 46. Sob o esquema de perturbação de casos apresentado na Figura 5.6(b), encontramos as observações 1, 4, 25 e 39 como observações potencialmente influentes. A Figura 5.6(c) mostra o gráfico de C_i quando a perturbação ocorre nos valores de x . Neste caso, as observações 1, 2 e 4 são destacadas. Afim de verificar se estas observações são valores atípicos realizamos uma análise de sensibilidade destas observações. Esta análise consiste em remover uma determinada observação, reajustar o modelo e, em seguida, calcular as taxas de variação das estimativas dos parâmetros pela expressão:

$$\text{taxa}(\xi) = \left| \frac{\hat{\xi}^{(i)} - \hat{\xi}}{\hat{\xi}} \right| \times 100,$$

em que $\hat{\xi}$ e $\hat{\xi}^{(i)}$ são os EMV no modelo completo e no modelo sem a i -ésima observação, respectivamente. Os resultados dessa análise estão apresentados na Tabela 5.4. Observamos que as retiradas das observações 1 e 4 apresentaram maiores impactos nas estimativas dos parâmetros, principalmente nas estimativas dos coeficientes de regressão. O mesmo pode ser visto no gráfico de influência global apresentado na Figura 5.7, o qual apresenta as observações

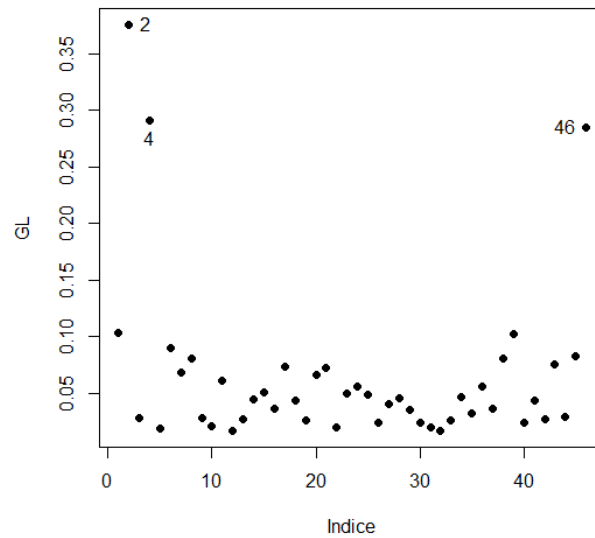
Figura 5.4 Gráfico dos resíduos versus o índice de observações e envelope para o Modelo III.



Fonte: Autoria própria

1 e 4 com os maiores valores da medida de influência LD .

Figura 5.5 Gráfico de alavancagem generalizada para o Modelo III.

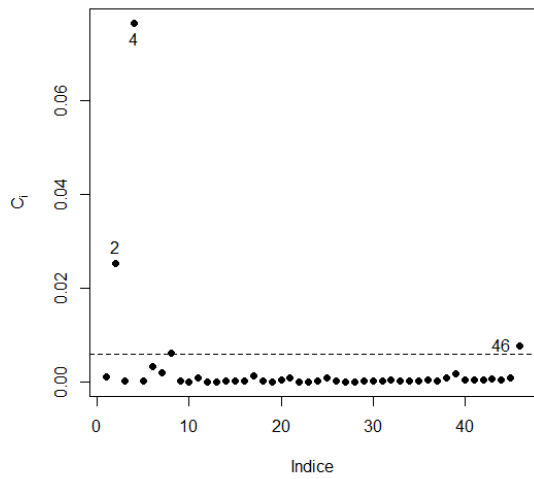


Fonte: Autoria própria

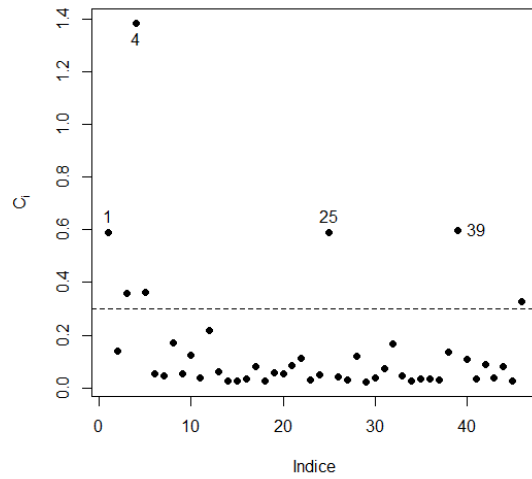
Tabela 5.4 Análise de sensibilidade ao remover as observações influentes no Modelo III.

Observações	1	2	4	25	39	46
β_0	6,47	2,59	5,14	3,20	0,80	2,67
β_1	8,80	3,34	7,77	4,50	0,60	1,59
β_2	16,69	8,25	15,62	9,56	4,91	11,95
θ	1,56	1,71	2,76	2,78	1,82	0,92

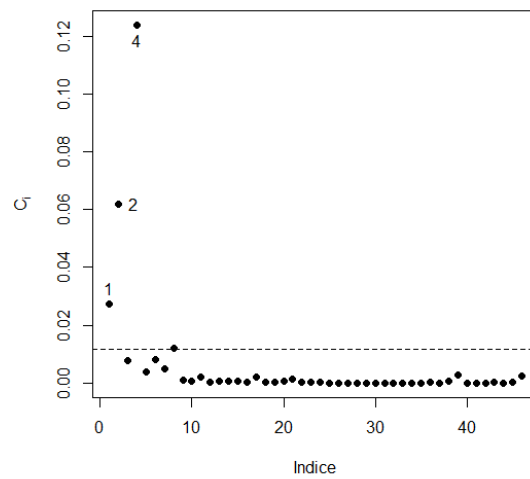
Figura 5.6 Gráfico de C_i contra o índice de observações segundo os esquemas de perturbação (a) aditiva na resposta (b) casos ponderados e (c) aditiva no preditor para o Modelo III.



(a)

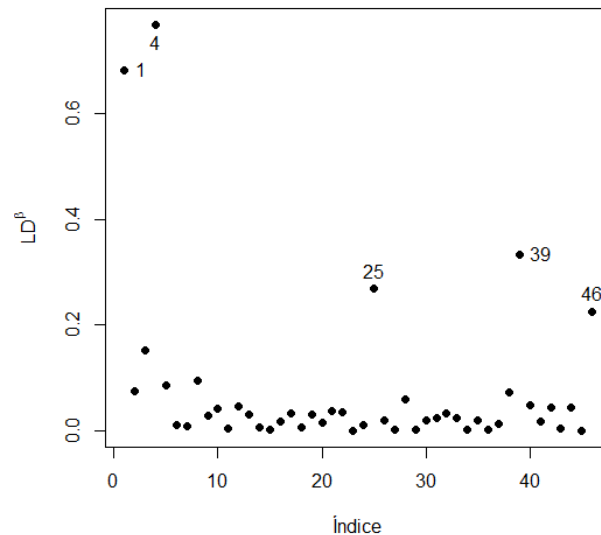


(b)



(c)

Fonte: Autoria própria

Figura 5.7 Gráfico de Influência Global para o Modelo III.

Fonte: Autoria própria

5.5 Comentários

Na análise residual do modelo não linear log-gama generalizado, propomos três resíduos: resíduo de Pearson, o resíduo padronizado e o componente do desvio. Realizamos um estudo de simulação para avaliar os desempenhos destes resíduos. Observamos que a distribuição dos resíduos de Pearson e o padronizado tem média próxima de zero, desvio padrão próximo a 1, é leptocúrtica e ligeiramente assimétrica. A diferença desses dois resíduos é que o resíduo de Pearson apresentou assimetria negativa, enquanto o resíduo padronizado assimetria positiva. Já o resíduo do componente do desvio é, entre os resíduos propostos, o que apresenta médias mais distantes do zero, além de apresentar erros padrão próximos a 1 e medidas de assimetria e curtoses menores do que os demais resíduos. Como era de esperar, com o aumento do valor de λ , a assimetria e a curtose dos três resíduos vão se aproximando da distribuição normal. Dos três resíduos propostos, sugerimos o uso do resíduo de Pearson. Apresentamos também algumas técnicas de diagnóstico para os MNLLGG, como as seguintes medidas: alavancagem

generalizada, influência global e influência local. Para esta última, adotamos a metodologia de Cook (1986), considerando os esquemas de perturbação na resposta, de casos e no preditor. Mostramos que há uma relação entre a matriz de influência local sob perturbação na resposta e a matriz de alavancagem generalizada. Por último, avaliamos as técnicas de diagnóstico, através do conjunto de dados que correspondem à peças de metal que foram submetidas à fadiga em ciclos, as quais mostraram-se ser eficientes em detectar observações influentes.

6

Considerações Finais

Resumimos as principais contribuições teóricas deste trabalho nos seguintes itens:

- (i) No Capítulo 2, introduzimos a classe de modelos não lineares log-gama generalizados (MNLLGG), o qual abrange dois importantes modelos: o modelo linear log-gama generalizado (Young e Bakir, 1987) e o modelo de regressão não linear valor-extremo tipo I (Barreto-Souza e Vasconcellos, 2011).
- (ii) No Capítulo 3, derivamos uma expressão em forma matricial para o viés de segunda ordem de Cox e Snell (1968) para os estimadores de máxima verossimilhança (EMV) dos parâmetros dos MNLLGG. Os resultados cobrem a situação em que o parâmetro de forma λ^2 é fixo.
- (iii) No Capítulo 4 derivamos uma expressão em forma matricial do fator de correção de Bartlett usado para aperfeiçoar o teste baseado na estatística da razão de verossimilhanças (LR), como também do fator de correção tipo-Bartlett para as estatísticas gradiente (S_g) e escore (S_r) na classe dos MNLLGG, considerando fixos parâmetro de forma λ^2 e o parâmetro de escala θ .
- (iv) No Capítulo 5, propomos uma metodologia de análise de resíduos e diagnóstico, sob o enfoque de alavancagem generalizada, influência local e influência global para a classe dos MNLLGG com o parâmetro de forma λ^2 fixo.

Além dessas contribuições, estudos de simulação de Monte Carlo e aplicações foram feitos com a finalidade de verificar a eficácia das correções e da metodologia proposta para a análise de resíduos e diagnóstico nos MNLLGG, dos quais podemos tirar as seguintes conclusões:

- (a) Os EMV corrigidos via Cox e Snell e via bootstrap apresentaram um bom desempenho. As correções nos estimadores se fazem necessárias mesmo com tamanhos de amostras consideráveis, uma vez que estas produzem estimadores mais eficientes tanto em termos de viés, como em termos de viés relativo e erro quadrático médio, do que o EMV usual.
- (b) O ganho com as correções de Bartlett na estatística LR e tipo-Bartlett nas estatísticas S_r e S_g foram bem evidente, de modo que com os testes corrigidos as taxas de rejeição ficaram bem mais próximas do nível nominal. As versões corrigidas dos testes LR e S_r apresentaram taxas de rejeição mais estáveis mesmo com a variação do valor do λ , do tamanho da amostra e do número de parâmetros de perturbação. Já no teste gradiente, o ganho com a correção tipo-Bartlett foi mais evidente quando variamos o tamanho da amostra e o número de parâmetros de perturbação.
- (c) As técnicas de diagnóstico desenvolvidas para os MNLLGG mostram-se eficientes em detectar os pontos influentes. Dos três resíduos propostos, sugerimos o uso do resíduo de Pearson, uma vez que a sua distribuição tem média próxima de zero, desvio padrão próximo à 1, assimetria negativa, conforme a assimetria da distribuição LGG.

Referências Bibliográficas

- AGOSTINELLI, C.; MARAZZI, A.; YOHAI, V. J. Robust Estimators of the Generalized Log-Gamma Distribution. *Technometrics*, v. 56, p. 92–101, 2014.
- BARRETO-SOUZA, W.; VASCONCELLOS, K. L. P. Bias and skewness in a general extreme-value regression model. *Computational Statistics and Data Analysis*, v. 55, p. 1379–1393, 2011.
- BARTLETT, M. S. Properties of sufficiency and statistical tests. *Proceedings of the Royal Society A*, v. 160, p. 268–282, 1937.
- BATES, D. M.; WATTS, D. G. *Nonlinear Regression Analysis and Its Applications*. New York: Wiley, 1988.
- BAYER, F. M.; CRIBARI-NETO, F. Bartlett corrections in beta regression models. *Journal of Statistical Planning and Inference*, v. 143, p. 531–547, 2013.
- BOTTER, D. A.; CORDEIRO, G. M. Improved estimators for generalized linear models with dispersion covariates. *Journal of Statistical Computation and Simulation*, v. 62, p. 91–104, 1998.
- CHANDRA, T. K.; MUKERJEE, R. Bartlett-type modification for Rao's efficient score statistic. *Journal of Multivariate Analysis*, v. 36, p. 103–112, 1991.
- CHESHER, A.; SMITH, R. Bartlett corrections to likelihood ratio tests. *Biometrika*, v. 82, p. 433–436, 1995.
- COOK, R. D. Assessment of local influence. *Journal of the Royal Statistical Society B*, v. 48, p. 133–169, 1986.
- COOK, R. D. Detection of influential observations in linear regressions. *Technometrics*, v. 19, p. 15–18, 1977.
- COOK, D. R.; TSAI, C. L.; WEI, B. C. Bias in nonlinear regression. *Biometrika*, v. 73, p. 615–623, 1986.
- CORDEIRO, G.M. Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistical Society B*, v. 45, p. 404–413, 1983.

- CORDEIRO, G. M. On the corrections to the likelihood ratio statistics. *Biometrika*, v. 74, 265–274, 1987.
- CORDEIRO, G. M. *Introdução à Teoria Assintótica*. IMPA Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1999.
- CORDEIRO, G. M. Corrected likelihood ratio tests in symmetric nonlinear regression models. *Journal of Statistical Computation and Simulation*, v. 74, p. 600–620, 2004.
- CORDEIRO, G. M.; BOTTER, D. A.; BARROSO, L. P.; FERRARI, S. L. P. Three corrected score tests for generalized linear models with dispersion covariates. *Statistica Neerlandica*, v. 57, p. 391–409, 2003.
- CORDEIRO, G. M.; CRIBARI-NETO, F. *An introduction to Bartlett correction and bias reduction*. Nova York: Springer, 2014.
- CORDEIRO, G. M.; CRIBARI-NETO, F.; AUBIN, E. C. Q.; FERRARI, S. L. P. Bartlett correction for one-parameter exponential family models. *Journal of Statistical Computation and Simulation*, v. 53, p. 211–231, 1995.
- CORDEIRO, G. M.; CYSNEIROS, A. H. M. A.; CYSNEIROS, F. J. A. Bartlett adjustments for Overdispersed Generalized linear Models. *Communications in Statistics – Theory and Methods*, v. 35, p. 937–952, 2006.
- CORDEIRO, G. M.; FERRARI, S. L. P. A modified score test statistic having chi-squared distribution to order n^{-1} . *Biometrika*, v. 78, p. 573–582, 1991.
- CORDEIRO, G. M.; FERRARI, S. L. P.; CYSNEIROS, A. H. M. A. A formula to improve score test statistics. *Journal of Statistical Computation and Simulation*, v. 62, p. 123–136, 1998.
- CORDEIRO, G. M.; FERRARI, S. L. P.; PAULA, G. A. Improved score tests for generalized linear models. *Journal of the Royal Statistical Society B*, v. 55, p. 661–674, 1993.
- CORDEIRO, G. M.; FERRARI, S. L. P.; URIBE-OPAZO, M. A.; VASCONCELLOS, K. L. P. Corrected maximum-likelihood estimation in a class of symmetric nonlinear regression models. *Statistics and Probability Letters*, v. 46, p. 317–328, 2000.
- CORDEIRO, G. M.; MCCULLAGH, P. Bias correction in generalized linear models. *Journal of the Royal Statistical Society B*, v. 53, p. 629–643, 1991.
- CORDEIRO, G. M.; PAULA, G. A. Improved likelihood ratio statistics for exponential family nonlinear models. *Biometrika*, v. 76, 93–100, 1989.
- CORDEIRO, G. M.; VASCONCELLOS, K. L. P. Bias correction for a class of multivariate nonlinear regression models. *Statistics and Probability Letters*, v. 35, p. 155–164, 1997.
- COX, D. R. Some aspects of conditional and asymptotic inference: a review. *Sankhya: The Indian Journal of Statistics, Series A (1961-2002)*, v. 50, p. 314–337, 1988.

- COX, D. R.; HINKLEY, D. V. *Theoretical Statistics*. London: Chapman and Hall, 1974.
- COX, D. R.; SNELL, E. A general definition of residuals. *Journal of the Royal Statistical Society B*, v. 30, p. 248–275, 1968.
- CRIBARI-NETO, F.; FERRARI, S. L. P. Bartlett corrected tests for heteroskedastic linear models. *Economics Letters*, v. 48, p. 113–118, 1995a.
- CRIBARI-NETO, F.; FERRARI, S. L. P. Second order asymptotics for score tests in generalized linear models. *Biometrika*, v. 82, p. 426–432, 1995b.
- CRIBARI-NETO, F.; QUEIROZ, M. P. F. On testing inference in beta regression. *Journal of Statistical Computation and Simulation*, v. 84, p. 186–203, 2014.
- CRIBARI-NETO, F.; ZARKOS, S. G. Improved test statistics for multivariate regression. *Economics Letters*, v. 49, p. 113–120, 1995.
- CRIBARI-NETO, F.; ZARKOS, S. G. R: yet another econometric programming environment. *Journal of Applied Econometrics*, v. 14, p. 319–329, 1999.
- CRIBARI-NETO, F.; ZARKOS, S. G. Econometric and statistical computing using Ox. *Computational Economics*, v. 21, p. 277–295, 2003.
- CYSNEIROS, A. H. M. A.; FERRARI, S. L. P. An improved likelihood ratio test for varying dispersion in exponential family nonlinear models. *Statistics and Probability Letters*, v. 76, p. 255–265, 2006.
- CYSNEIROS, A. H. M. A.; RODRIGUES, K. S. P.; CORDEIRO, G. M.; FERRARI, S. L. P. Three-Bartlett-type corrections for Score statistics in Symmetric nonlinear regression models. *Statistical Papers*, v. 51, p. 273–284, 2010.
- CYSNEIROS, F. J. A.; VANEGAS, L. H. Residuals and their statistical properties in symmetrical nonlinear models. *Statistics and Probability Letter*, v. 78, p. 3269–3273, 2008.
- DICICCIO, T. J. Approximate inference for the generalized gamma distribution. *Technometrics*, v. 29, p. 33–40, 1987.
- DOORNIK, J. A. *Object-Oriented Matrix Programming Using Ox*. 3th ed. Timberlake Consultants Press and Oxford, London, 2007. Disponível em: <http://www.doornik.com>. Acesso em: 30 ago. 2016.
- EFRON, B. Bootstrap methods: another look at the jackknife. *Annals of Statistics*, v. 7, p. 1–26, 1979.
- EFRON, B.; TIBSHIRANI, R. J. Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science*, v. 1, p. 54–77, 1986.
- ESCOBAR, L. A.; MEEKER, W. Q. Assessing influence in regression analysis with censored data. *Biometrics*, v. 48, p. 507–528, 1992.

FERRARI, S. L. P.; ARELLANO-VALLE, R. B. Modified likelihood ratio and score tests in regression models using the t distribution. *Brazilian Journal of Probability and Statistics*, v. 10, p. 15–33, 1996.

FERRARI, S. L. P.; CYSNEIROS, A. H. M. A.; CRIBARI-NETO, F. An improved test for heteroskedasticity using adjusted modified profile likelihood inference. *Journal of Statistical Planning and Inferences*, v. 57, p. 353–361, 2004.

FERRARI, S. L. P.; URIBE-OPAZO, M. A. Corrected likelihood ratio test in class of symmetric linear regression models. *Brazilian Journal of Probability and Statistics*, v. 15, p. 49–67, 2001.

FOX, J.; WEISBERG, S. *An R Companion to Applied Regression*. 2th ed. Thousand Oaks CA: Sage, 2011.

FREUND, R. J. Regression with SAS emphasis on PROC REG. In: EIGHTH ANUAL SAS USERS GROUP INTERNATIONAL CONFERENCE 16-19. New Orleans, Louisiana, 1983.

FUNG, W. H.; ZHU, Z. Y.; WEI, B. C.; HE, X. Influence diagnostics and outliers tests for semi-parametric mixed models. *Journal of the Royal Statistical Society B*, v. 64, p. 565–579, 2002.

GALEA, M.; RIQUELME, M.; PAULA, G. A. Diagnostics methods in elliptical linear regression models. *Brazilian Journal of Probability and Statistics*, v. 14, p. 167–184, 2000.

HALL, P. *The Bootstrap and Edgeworth Expansion*. Springer-Verlag, New York, 1992.

HARRIS, P. An asymptotic expansion for the null distribution of the efficient score statistic. *Biometrika*, v. 72, p. 653–659, 1985.

HASHIMOTO, E. M.; ORTEGA, E. M. M.; CANCHO, V. G.; CORDEIRO, G. M. On estimation and diagnostics analysis in log-generalized gamma regression model for interval-censored data. *Statistics: A Journal of Theoretical and Applied Statistics*, v. 47, p. 379–398, 2013.

HAYAKAWA, T. The likelihood ratio criterion and the asymptotic expansion of its distribution. *Annals of the Institute of Statistical Mathematics*, v. 29, p. 359–378, 1977.

IHAKA, R.; GENTLEMAN, R. R. a language for data analysis and graphics. *Journal of Computational Graphics and Statistics*, v. 5, p. 299–314, 1996.

KAKISAWA, Y. Higher order monotone Bartlett-type adjustment for some multivariate test statistics. *Biometrika*, v. 83, p. 923–927, 1996.

LAGOS, B. M.; MORETTIN, P. A.; BARROSO, L. P. Some corrections of the score test statistic for Gaussian ARMA models. *Brazilian Journal of Probability and Statistics*, v. 24, 434–456, 2010.

LAWLESS, J. F. Inference in the Generalized Gamma and Log Gamma Distributions. *Technometrics*, v. 22, p. 409–419, 1980.

- LAWLESS, J. F. *Statistical models and methods for lifetime data*. New York: Wiley, 2003.
- LAWLEY, D. A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika*, v. 43, p. 295–303, 1956.
- LEMONTE, A. J. *The gradient test: another likelihood - based test*. London: Academic Press, 2016.
- LEMONTE, A. J.; CORDEIRO, G. M. Birnbaum-Saunders nonlinear regression models. *Computational Statistics and Data Analysis*, v. 53, p. 4441–4452, 2009.
- LEMONTE, A.J.; FERRARI, S.L.P. Small-sample corrections for score tests in Birnbaum-Saunders regressions. *Communications in Statistics – Theory and Methods*, v. 40, p. 232–243, 2011.
- LEMONTE, A. J.; FERRARI, S. L. P. The local power of the gradient test. *Annals of the Institute of Statistical Mathematics*, v. 64, p. 373–381, 2012.
- LEMONTE, A. J.; FERRARI, S. L. P.; CRIBARI-NETO, F. Improved likelihood inference in Birnbaum Saunders regressions. *Computational Statistics and Data Analysis*, v. 54, p. 1307–1316, 2010.
- LESAFFRE, F.; VERBEKE, G. Local influence in linear mixed models. *Biometrics*, v. 38, p. 963–974, 1998.
- MEDEIROS, F.M.C.; FERRARI, S. L. P. Small-sample testing inference in symmetric and log-symmetric linear regression models. arXiv:1602.00769v1[stat.ME], 2016.
- MELLO, M. P.; PETERNELLI, L. A. *Conhecendo o R - Uma Visão mais que Estatística*. Viçosa: Editora UFV, 2013.
- MONTENEGRO, L. C. C.; CORDEIRO, G. M. Bartlett corrected likelihood ratio tests in location-scale nonlinear models. *Communications in Statistics, Theory and Methods*, v. 30, p. 1353–1372, 2002.
- NASCIMENTO, K. P. Correção tipo-Bartlett em modelos não lineares simétricos heteroscedásticos, 59p., Dissertação (Mestrado em Biometria e Estatística Aplicada)– Universidade Federal Rural de Pernambuco, Recife, 2010.
- ORTEGA, E. M. M.; BOLFARINE H.; PAULA G. A. Influence diagnostics in generalized log-gamma regression models. *Computational Statistics and Data Analysis*, v. 42, p. 165–186, 2003.
- ORTEGA, E. M. M.; CANCHO, V. G.; PAULA G. A. Generalized log-gamma regression models with cure fraction. *Sankhya. Series B*, v. 71, p. 1–29, 2008.
- ORTEGA, E. M. M.; CORDEIRO, G. M.; PASCOA, M. A. R.; COUTO, E. V. The log-exponentiated generalized gamma regression model for censored data. *Journal of Statistical Computation and Simulation*, v. 82, p. 1169–1189, 2012.

- OSPINA, R.; CRIBARI-NETO, F.; VASCONCELLOS, K. L. P. Improved point and interval estimation for a beta regression model. *Computational Statistics and Data Analysis*, v. 51, p. 960–981, 2006.
- PREGIBON, D. Logistic regression diagnostics. *Annals of Statistics*, v. 9, 705–724, 1981.
- PRENTICE, R. L. A log-gamma model and its maximum likelihood estimation. *Biometrika*, v. 61, p. 539–544, 1974.
- RAO, C. R. Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society*, v. 44, p. 50–57, 1948.
- RAO, C. R. *Linear Statistical Inference and its Applications*. 2 ed. Wiley: New York, 1973.
- RIECK, J. R.; NEDELMAN, J. R. A log-linear model for the Birnbaum-Saunders distribution. *Technometrics*, v. 33, p. 51–60, 1991.
- RIZZO, M. L. *Statistical computing with R*. Boca Raton: Chapman and Hall/CRC, 2008.
- ROCKE, D. M. Bootstrap Bartlett adjustment in seemingly unrelated regression. *Journal of the American Statistical Association*, v. 84, p. 598–601, 1989.
- SEBER, G. A. F.; WILD, C. J. *Nonlinear Regression*. New York: Wiley, 1989.
- SILVA, G. O.; ORTEGA, E. M. M.; PAULA, G. A. Residuals for log-Burr XII regression models in survival analysis. *Journal of Applied Statistics*, v. 38, p. 1435–1445, 2010.
- STACY, E. W. A Generalization of the Gamma Distribution. *Annals of Mathematical Statistics*, v. 33, p. 1187–1192, 1962.
- SUN, B. R.; WEI, B. C. On influence assessment for LAD regression. *Statistics and Probability Letters*, v. 67, p. 97–110, 2004.
- TANG, N. S.; WEI, B. C.; WANG, X. R. Influence diagnostics in nonlinear reproductive dispersion models. *Statistics and Probability Letters*, v. 46, p. 59–68, 2000.
- TANIGUCHI, M. Third-order asymptotic properties of a class of test statistics under a local alternative. *Journal of Multivariate Analysis*, v. 37, p. 223–238, 1991.
- TERRELL, G. R. The gradient statistic. *Computing Science and Statistics*, v. 34, p. 206–215, 2002.
- URIBE-OPAZO, M. A.; FERRARI, S. L. P.; CORDEIRO, G. M. Improved score tests in symmetric linear regression models. *Communications in Statistics – Theory and Methods*, v. 37, p. 261–276, 2008,
- VANEGAS, L. H.; CYSNEIROS, F. J. A. Assessment of diagnostic procedures in symmetrical nonlinear regression models. *Computational Statistics and Data Analysis*, v. 54, p. 1002–1016, 2010.

VANEGAS, L. H.; RONDÓN, L. M.; CYSNEIROS, F. J. A. Diagnostic procedures in Birnbaum-Saunders nonlinear regression models. *Computational Statistics and Data Analysis*, v. 56, p. 1662–1680, 2012.

VARGAS, T. M.; FERRARI, S. L. P.; LEMONTE, A. J. Gradient statistic: Higher-order asymptotics and Bartlett-type correction. *Electronic Journal of Statistics*, v. 7, p. 43–61, 2013.

VARGAS, T. M.; FERRARI, S. L. P.; LEMONTE, A. J. Improved likelihood inference in generalized linear models. *Computational Statistics e Data Analysis*, v. 74, p. 110–124, 2014.

VENABLES, W. N.; RIPLEY, B. D. *Modern Applied Statistics with S-Plus*. Springer-Verlag: New York, 2002.

WALD, A. Tests of statistical hypothesis concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, v. 54, p. 426–482, 1943.

WEI, B. C. *Exponential Family Nonlinear Models*. Springer-Verlag: Singapore, 1998.

WEI, B. C.; HU, Y. Q.; FUNG, W. K. Generalized leverage and its applications. *Scandinavian Journal of Statistics*, v. 25, p. 25–37, 1998.

WILKS, S. S. The large-sample distribution of the likelihood ratio for testing composite hypothesis. *Annals of Mathematical Statistics*, v. 9, p. 60–62, 1938.

XIE, F. C.; WEI, B. C. Diagnostics analysis for log-Birnbaum-Saunders regression models. *Computational Statistical and Data Analysis*, v. 51, p. 4692–4706, 2007.

YOUNG, D. H.; BAKIR, S. T. Bias Correction for Generalized Log-Gamma Regression Models. *Technometrics*, v. 29, p. 183–191, 1987.

Apêndice A

Cálculo dos Momentos

Considere que Y_1, \dots, Y_n sejam variáveis independentes e um MNLLGG definido por (1) e (2). Nesta seção, apresentamos a obtenção os cálculos dos cumulantes de Y e as suas derivadas em relação ao vetor de parâmetros $\xi = (\beta, \theta)^\top$. De (2), temos que $z_i = (y_i - \eta_i)/\theta$, em que $\eta_i = \eta(x_i, \beta)$, assim o logaritmo da função de verossimilhança (função de log-verossimilhança) do parâmetro ξ , dado o vetor de observações $(y_1, \dots, y_n)^\top$ e a constante λ , do MNLLGG pode ser expresso em função de log-verossimilhança da variável Z ,

$$l(\xi) = \sum_{i=1}^n \log f_z \left(\frac{y_i - \eta_i}{\theta} \right) - n \log \theta.$$

A.1 Derivadas da Função de Log-Verossimilhança

Assumindo que r, s, t , são indexadores do espaço paramétrico correspondente às componentes β e denotando $\tilde{x}_{ir} = \partial \eta_i / \partial \beta_r$, $\tilde{x}_{irs} = \partial^2 \eta_i / \partial \beta_r \partial \beta_s$ e $\tilde{x}_{irst} = \partial^3 \eta_i / \partial \beta_r \partial \beta_s \partial \beta_t$, por simples diferenciação em relação aos componentes do parâmetro ξ , temos:

$$U_r = \frac{\partial l(\xi)}{\partial \beta_r} = -\frac{1}{\theta} \sum_i \frac{\partial \log f_z(z_i)}{\partial z_i} \tilde{x}_{ir} \quad \text{e} \quad U_\theta = \frac{\partial l(\xi)}{\partial \theta} = -\frac{1}{\theta} \sum_i \frac{\partial \log f_z(z_i)}{\partial z_i} z_i - \frac{n}{\theta}.$$

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

$$\begin{aligned} U_{rs} &= \frac{\partial^2 l(\xi)}{\partial \beta_r \partial \beta_s} = \frac{1}{\theta^2} \sum_i \frac{\partial^2 \log f_z(z_i)}{\partial z_i^2} \tilde{x}_{ir} \tilde{x}_{is} - \frac{1}{\theta} \sum_i \frac{\partial \log f_z(z_i)}{\partial z_i} \tilde{x}_{irs}, \\ U_{r\theta} &= \frac{\partial^2 l(\xi)}{\partial \beta_r \partial \theta} = \frac{1}{\theta^2} \sum_i \frac{\partial^2 \log f_z(z_i)}{\partial z_i^2} z_i \tilde{x}_{ir} + \frac{1}{\theta^2} \sum_i \frac{\partial \log f_z(z_i)}{\partial z_i} \tilde{x}_{ir}, \\ U_{\theta\theta} &= \frac{\partial^2 l(\xi)}{\partial \theta^2} = \frac{1}{\theta^2} \sum_i \frac{\partial^2 \log f_z(z_i)}{\partial z_i^2} z_i^2 + \frac{2}{\theta^2} \sum_i \frac{\partial \log f_z(z_i)}{\partial z_i} z_i + \frac{n}{\theta^2}, \end{aligned}$$

$$\begin{aligned}
U_{rs\theta} &= \frac{\partial^3 l(\xi)}{\partial \beta_r \partial \beta_s \partial \theta} = -\frac{1}{\theta^3} \sum_i \frac{\partial^3 \log f_z(z_i)}{\partial z_i^3} z_i \tilde{x}_{ir} \tilde{x}_{is} - \frac{2}{\theta^3} \sum_i \frac{\partial^2 \log f_z(z_i)}{\partial z_i^2} \tilde{x}_{ir} \tilde{x}_{is} \\
&\quad + \frac{1}{\theta^2} \sum_i \frac{\partial^2 \log f_z(z_i)}{\partial z_i^2} z_i \tilde{x}_{irs} + \frac{1}{\theta^2} \sum_i \frac{\partial \log f_z(z_i)}{\partial z_i} \tilde{x}_{irs}, \\
U_{r\theta\theta} &= \frac{\partial^3 l(\xi)}{\partial \beta_r \partial \theta \partial \theta} = -\frac{3}{\theta^3} \sum_i \frac{\partial^2 \log f_z(z_i)}{\partial z_i^2} z_i \tilde{x}_{ir} - \frac{2}{\theta^3} \sum_i \frac{\partial \log f_z(z_i)}{\partial z_i} \tilde{x}_{ir} \\
&\quad - \frac{1}{\theta^3} \sum_i \frac{\partial^3 \log f_z(z_i)}{\partial z_i^3} z_i^2 \tilde{x}_{ir} - \frac{1}{\theta^3} \sum_i \frac{\partial^2 \log f_z(z_i)}{\partial z_i^2} z_i \tilde{x}_{ir}, \\
U_{\theta\theta\theta} &= \frac{\partial^3 l(\xi)}{\partial \theta \partial \theta \partial \theta} = -\frac{6}{\theta^3} \sum_i \frac{\partial^2 \log f_z(z_i)}{\partial z_i^2} z_i^2 - \frac{6}{\theta^3} \sum_i \frac{\partial \log f_z(z_i)}{\partial z_i} z_i \\
&\quad - \frac{1}{\theta^3} \sum_i \frac{\partial^3 \log f_z(z_i)}{\partial z_i^3} z_i^3 - \frac{2n}{\theta^3}.
\end{aligned}$$

A derivadas de quarta ordem apenas para o vetor β é dada por:

$$\begin{aligned}
U_{rstu} &= \frac{\partial^4 l(\xi)}{\partial \beta_r \partial \beta_s \partial \beta_t \partial \beta_u} = \frac{1}{\theta^4} \sum_i \frac{\partial^4 \log f_z(z_i)}{\partial z_i^4} \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} - \frac{1}{\theta^3} \sum_i \frac{\partial^3 \log f_z(z_i)}{\partial z_i^3} (\tilde{x}_{iru} \tilde{x}_{is} \tilde{x}_{it} \\
&\quad + \tilde{x}_{ir} \tilde{x}_{isu} \tilde{x}_{it} + \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu} + \tilde{x}_{iu} \tilde{x}_{irt} \tilde{x}_{is} + \tilde{x}_{iu} \tilde{x}_{ir} \tilde{x}_{ist} + \tilde{x}_{iu} \tilde{x}_{irs} \tilde{x}_{it}) \\
&\quad + \frac{1}{\theta^2} \sum_i \frac{\partial^2 \log f_z(z_i)}{\partial z_i^2} (\tilde{x}_{irtu} \tilde{x}_{is} + \tilde{x}_{irt} \tilde{x}_{isu} + \tilde{x}_{iru} \tilde{x}_{ist} + \tilde{x}_{ir} \tilde{x}_{istu} + \tilde{x}_{irsu} \tilde{x}_{it} + \tilde{x}_{irs} \tilde{x}_{itu} \\
&\quad + \tilde{x}_{iu} \tilde{x}_{irst}) - \frac{1}{\theta} \sum_i \frac{\partial \log f_z(z_i)}{\partial z_i} \tilde{x}_{irstu}.
\end{aligned}$$

A.2 Cálculo dos Cumulantes

Para obtenção dos cumulantes, utilizamos a seguinte notação de Young e Bakir (1987):

$$E_l^{(j)} = E\{-Z^j [\partial^l \log f(Z) / \partial Z^l]\} \text{ e } E_{l,m}^{(j)} = E\{-Z^j [\partial^l \log f(Z) / \partial Z^l] [\partial^m \log f(Z) / \partial Z^m]\}.$$

Adicionalmente, Young e Bakir (1987) mostraram alguns resultados importantes:

$$\begin{aligned}
E_1^{(0)} &= E_1^{(2)} = 0, \\
E_1^{(1)} &= E_2^{(0)} = -E_{1,1}^{(0)} = 1, \\
E_{2,1}^{(1)} &= -(1 + \lambda^{-2}), \\
E_2^{(1)} &= E_3^{(0)} = -E_{2,1}^{(0)} = -E_{1,1}^{(1)} = \lambda E_3^{(1)} = \lambda^{-1},
\end{aligned}$$

$$\begin{aligned}
E_3^{(3)} &= 3\psi^{(1)}(\lambda^2) + \lambda^2\psi^{(2)}(\lambda^2), \\
E_{1,1}^{(2)} &= \lambda E_{2,1}^{(2)} = -\left\{2 + \lambda^2\psi^{(1)}(\lambda^2)\right\}, \\
E_2^{(2)} &= \frac{1}{3}E_1^{(3)} = \lambda E_3^{(2)} = \lambda^2\psi^{(1)}(\lambda^2).
\end{aligned}$$

Tomando as esperanças das derivadas de segunda e terceira ordem, obtemos os seguintes cumulantes:

$$\begin{aligned}
\kappa_{rs} &= -\frac{1}{\theta^2} \sum_i E_2^{(0)} \tilde{x}_{ir} \tilde{x}_{is} + \frac{1}{\theta} \sum_i E_1^{(0)} \tilde{x}_{irs} = -\frac{1}{\theta^2} \sum_i \tilde{x}_{ir} \tilde{x}_{is}, \\
\kappa_{r\theta} &= -\frac{1}{\theta^2} \sum_i E_2^{(1)} \tilde{x}_{ir} - \frac{1}{\theta^2} \sum_i E_1^{(0)} \tilde{x}_{ir} = -\frac{1}{\lambda\theta^2} \sum_i \tilde{x}_{ir}, \\
\kappa_{\theta\theta} &= -\frac{1}{\theta^2} \sum_i E_2^{(2)} - \frac{2}{\theta^2} \sum_i E_1^{(1)} + \frac{n}{\theta^2} = -\frac{n}{\theta^2} \left\{1 + \lambda^2\psi^{(1)}(\lambda^2)\right\}, \\
\kappa_{rst} &= \frac{1}{\theta^3} \sum_i E_3^{(0)} \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} - \frac{1}{\theta^2} \sum_i E_2^{(0)} (\tilde{x}_{irt} \tilde{x}_{is} + \tilde{x}_{ir} \tilde{x}_{ist} + \tilde{x}_{irs} \tilde{x}_{it}) + \frac{1}{\theta} \sum_i E_1^{(0)} \tilde{x}_{irst} \\
&= \frac{1}{\lambda\theta^3} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} - \frac{1}{\theta^2} \sum_i (\tilde{x}_{irt} \tilde{x}_{is} + \tilde{x}_{ir} \tilde{x}_{ist} + \tilde{x}_{irs} \tilde{x}_{it}), \\
\kappa_{rs\theta} &= \frac{1}{\theta^3} \sum_i E_3^{(1)} \tilde{x}_{ir} \tilde{x}_{is} + \frac{2}{\theta^3} \sum_i E_2^{(0)} \tilde{x}_{ir} \tilde{x}_{is} - \frac{1}{\theta^2} \sum_i E_2^{(1)} \tilde{x}_{irs} - \frac{1}{\theta^2} \sum_i E_1^{(0)} \tilde{x}_{irs} \\
&= \frac{1}{\theta^3} \left(\frac{1}{\lambda^2} + 2\right) \sum_i \tilde{x}_{ir} \tilde{x}_{is} - \frac{1}{\lambda\theta^2} \sum_i \tilde{x}_{irs}, \\
\kappa_{r\theta\theta} &= \frac{3}{\theta^3} \sum_i E_2^{(1)} \tilde{x}_{ir} + \frac{2}{\theta^3} \sum_i E_1^{(0)} \tilde{x}_{ir} + \frac{1}{\theta^3} \sum_i E_3^{(2)} \tilde{x}_{ir} + \frac{1}{\theta^3} \sum_i E_2^{(1)} \tilde{x}_{ir} \\
&= \frac{1}{\theta^3} \left(\frac{4}{\lambda} + \lambda\psi^{(1)}(\lambda^2)\right) \sum_i \tilde{x}_{ir}, \\
\kappa_{\theta\theta\theta} &= \frac{6}{\theta^3} \sum_i E_2^{(2)} + \frac{6}{\theta^3} \sum_i E_1^{(1)} + \frac{1}{\theta^3} \sum_i E_3^{(3)} - \frac{2n}{\theta^3} \\
&= \frac{n}{\theta^3} \left\{3(2\lambda^2 + 1)\psi^{(1)}(\lambda^2) + \lambda^2\psi^{(2)}(\lambda^2) + 4\right\}.
\end{aligned}$$

Do mesmo modo, o quarto cumulante é dado por:

$$\begin{aligned} \kappa_{rstu} = & -\frac{1}{\lambda^2 \theta^4} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} + \frac{1}{\lambda \theta^3} \sum_i (\tilde{x}_{iru} \tilde{x}_{is} \tilde{x}_{it} + \tilde{x}_{ir} \tilde{x}_{isu} \tilde{x}_{it} + \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu} \\ & + \tilde{x}_{iu} \tilde{x}_{irt} \tilde{x}_{is} + \tilde{x}_{iu} \tilde{x}_{ir} \tilde{x}_{ist} + \tilde{x}_{iu} \tilde{x}_{irs} \tilde{x}_{it}) - \frac{1}{\theta^2} \sum_i (\tilde{x}_{irtu} \tilde{x}_{is} \\ & + \tilde{x}_{irt} \tilde{x}_{isu} + \tilde{x}_{iru} \tilde{x}_{ist} + \tilde{x}_{ir} \tilde{x}_{istu} + \tilde{x}_{irsu} \tilde{x}_{it} + \tilde{x}_{irs} \tilde{x}_{itu} + \tilde{x}_{iu} \tilde{x}_{irst}). \end{aligned}$$

Além disso, temos para o modelo linear log-gama generalizado os seguintes cumulantes:

$$\begin{aligned} \kappa_{rs,t} &= E(U_{rs}U_t) = -\frac{1}{\lambda \theta^3} \sum_i x_{ir} x_{is} x_{it}, \\ \kappa_{r,s,t} &= E(U_r U_s U_t) = \frac{2}{\lambda \theta^3} \sum_i x_{ir} x_{is} x_{it}, \\ \kappa_{rs,tu} &= E(U_{rs}U_{tu}) = \frac{1}{\theta^4 \lambda^2} \sum_i x_{ir} x_{is} x_{it} x_{iu}, \\ \kappa_{r,s,t,u} &= E(U_r U_s U_t U_u) = \frac{6}{\theta^4 \lambda^2} \sum_i x_{ir} x_{is} x_{it} x_{iu}, \\ \kappa_{r,s,tu} &= E(U_r U_s U_{tu}) = -\frac{2}{\theta^4 \lambda^2} \sum_i x_{ir} x_{is} x_{it} x_{iu}. \end{aligned}$$

A.3 Derivadas dos Cumulantes

Calculando as derivadas dos segundos cumulantes em relação aos componentes de ξ , obtemos:

$$\begin{aligned} \kappa_{rs}^{(t)} &= \frac{\partial \kappa_{rs}}{\partial \beta_t} = -\frac{1}{\theta^2} \sum_i (\tilde{x}_{irt} \tilde{x}_{is} + \tilde{x}_{ir} \tilde{x}_{ist}), \\ \kappa_{rs}^{(\theta)} &= \frac{\partial \kappa_{rs}}{\partial \theta} = \frac{2}{\theta^3} \sum_i \tilde{x}_{ir} \tilde{x}_{is}, \\ \kappa_{r\theta}^{(t)} &= \frac{\partial \kappa_{r\theta}}{\partial \beta_t} = -\frac{1}{\lambda \theta^2} \sum_i \tilde{x}_{irt}, \\ \kappa_{r\theta}^{(\theta)} &= \frac{\partial \kappa_{r\theta}}{\partial \theta} = \frac{2}{\lambda \theta^3} \sum_i \tilde{x}_{ir}, \\ \kappa_{\theta\theta}^{(t)} &= \frac{\partial \kappa_{\theta\theta}}{\partial \beta_t} = 0, \\ \kappa_{\theta\theta}^{(\theta)} &= \frac{\partial \kappa_{\theta\theta}}{\partial \theta} = \frac{2n}{\theta^3} \left\{ 1 + \lambda^2 \psi^{(1)}(\lambda^2) \right\}. \end{aligned}$$

Adicionalmente, temos:

$$\begin{aligned}\kappa_{rs}^{(tu)} &= \frac{\partial \kappa_{rs}^{(t)}}{\partial \beta_u} = -\frac{1}{\theta^2} \sum_i (\tilde{x}_{irtu} \tilde{x}_{is} + \tilde{x}_{irt} \tilde{x}_{isu} + \tilde{x}_{iru} \tilde{x}_{ist} + \tilde{x}_{ir} \tilde{x}_{istu}), \\ \kappa_{rst}^{(u)} &= \frac{\partial \kappa_{rst}}{\partial \beta_u} = \frac{1}{\lambda \theta^3} \sum_i (\tilde{x}_{iru} \tilde{x}_{is} \tilde{x}_{it} + \tilde{x}_{ir} \tilde{x}_{isu} \tilde{x}_{it} + \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu}) - \frac{1}{\theta^2} \sum_i (\tilde{x}_{irtu} \tilde{x}_{is} + \tilde{x}_{irt} \tilde{x}_{isu} + \tilde{x}_{iru} \tilde{x}_{ist} \\ &\quad + \tilde{x}_{ir} \tilde{x}_{istu} + \tilde{x}_{irsu} \tilde{x}_{it} + \tilde{x}_{irs} \tilde{x}_{itu}).\end{aligned}$$

Além disso,

$$\begin{aligned}\kappa_{r,s,t} &= \frac{2}{\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \\ \kappa_{r,st} &= \frac{1}{\theta^2} \sum_i \tilde{x}_{ir} \tilde{x}_{ist} - \frac{1}{\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \\ \kappa_{rs,tu} &= \frac{1}{\theta^4 \lambda^2} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} - \frac{1}{\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu} - \frac{1}{\theta^3 \lambda} \sum_i \tilde{x}_{irs} \tilde{x}_{it} \tilde{x}_{iu} + \frac{1}{\theta^2} \sum_i \tilde{x}_{irs} \tilde{x}_{itu} \\ \kappa_{r,s,t,u} &= \frac{6}{\theta^4 \lambda^2} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} \\ \kappa_{r,s,tu} &= -\frac{2}{\theta^4 \lambda^2} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} + \frac{2}{\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu}\end{aligned}$$

Apêndice B

Cálculo dos Termos de $B(\hat{\xi})$

Nesta seção, apresentamos as principais quantidades necessárias aos cálculos da correção de viés do estimador de máxima verossimilhança $\hat{\xi}$ do vetor de parâmetro ξ que indexa o MNLLGG, apresentada no Capítulo 3.

$$\begin{aligned}
\kappa_{rs}^{(t)} - \frac{1}{2}\kappa_{rst} &= -\frac{1}{2\theta^2} \sum_i (\tilde{x}_{irt}\tilde{x}_{is} + \tilde{x}_{ir}\tilde{x}_{ist} - \tilde{x}_{irs}\tilde{x}_{it}) - \frac{1}{2\lambda\theta^3} \sum_i \tilde{x}_{ir}\tilde{x}_{is}\tilde{x}_{it}, \\
\kappa_{rs}^{(\theta)} - \frac{1}{2}\kappa_{rs\theta} &= \frac{1}{\theta^3} \left(1 - \frac{1}{2\lambda^2}\right) \sum_i \tilde{x}_{ir}\tilde{x}_{is} + \frac{1}{2\lambda\theta^2} \sum_i \tilde{x}_{irs}, \\
\kappa_{r\theta}^{(t)} - \frac{1}{2}\kappa_{r\theta t} &= -\frac{1}{\theta^3} \left(1 + \frac{1}{2\lambda^2}\right) \sum_i \tilde{x}_{ir}\tilde{x}_{it} - \frac{1}{2\lambda\theta^2} \sum_i \tilde{x}_{irt}, \\
\kappa_{\theta s}^{(t)} - \frac{1}{2}\kappa_{\theta st} &= -\frac{1}{\theta^3} \left(1 + \frac{1}{2\lambda^2}\right) \sum_i \tilde{x}_{is}\tilde{x}_{it} - \frac{1}{2\lambda\theta^2} \sum_i \tilde{x}_{ist}, \\
\kappa_{r\theta}^{(\theta)} - \frac{1}{2}\kappa_{r\theta\theta} &= -\frac{\lambda\psi^{(1)}(\lambda^2)}{2\theta^3} \sum_i \tilde{x}_{ir}, \\
\kappa_{\theta s}^{(\theta)} - \frac{1}{2}\kappa_{\theta s\theta} &= -\frac{\lambda\psi^{(1)}(\lambda^2)}{2\theta^3} \sum_i \tilde{x}_{is}, \\
\kappa_{\theta\theta}^{(t)} - \frac{1}{2}\kappa_{\theta\theta t} &= -\frac{1}{2\theta^3} \left(\frac{4}{\lambda} + \lambda\psi^{(1)}(\lambda^2)\right) \sum_i \tilde{x}_{it}, \\
\kappa_{\theta\theta}^{(\theta)} - \frac{1}{2}\kappa_{\theta\theta\theta} &= -\frac{n}{2\theta^3} \left\{2\lambda^2\psi^{(1)}(\lambda^2) + 3\psi^{(1)}(\lambda^2) - \lambda^2\psi^{(2)}(\lambda^2)\right\}.
\end{aligned}$$

B.1 Cálculo dos Termos de $B(\hat{\beta}_b)$

Os termos em (10) são dados por:

$$\begin{aligned}
& \sum_{r,s,t} \kappa^{br} \kappa^{st} \left(\kappa_{rs}^{(t)} - \frac{1}{2} \kappa_{rst} \right) = \\
&= -\frac{1}{2\theta^2} \sum_{r,s,t} \kappa^{br} \kappa^{st} \sum_i \tilde{x}_{irt} \tilde{x}_{is} - \frac{1}{2\theta^2} \sum_{r,s,t} \kappa^{br} \kappa^{st} \sum_i \tilde{x}_{ir} \tilde{x}_{ist} \\
&\quad + \frac{1}{2\theta^2} \sum_{r,s,t} \kappa^{br} \kappa^{st} \sum_i \tilde{x}_{irs} \tilde{x}_{it} - \frac{1}{2\lambda\theta^3} \sum_{r,s,t} \kappa^{br} \kappa^{st} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \\
&= -\frac{1}{2\theta^2} \sum_i \left(\sum_{r,s,t} \kappa^{br} \tilde{x}_{irt} \kappa^{st} \tilde{x}_{is} \right) - \frac{1}{2\theta^2} \sum_i \left(\sum_r \kappa^{br} \tilde{x}_{ir} \right) \left(\sum_{s,t} \kappa^{st} \tilde{x}_{ist} \right) \\
&\quad + \frac{1}{2\theta^2} \sum_i \left(\sum_{r,s,t} \kappa^{br} \tilde{x}_{irs} \kappa^{st} \tilde{x}_{it} \right) - \frac{1}{2\lambda\theta^3} \sum_i \left(\sum_r \kappa^{br} \tilde{x}_{ir} \right) \left(\sum_{s,t} \tilde{x}_{is} \kappa^{st} \tilde{x}_{it} \right) \\
&= -\frac{1}{2\theta^2} \sum_i \left(\sum_r \kappa^{br} \tilde{x}_{ir} \right) \left(\sum_{s,t} \kappa^{st} \tilde{x}_{ist} \right) - \frac{1}{2\lambda\theta^3} \sum_i \left(\sum_r \kappa^{br} \tilde{x}_{ir} \right) \left(\sum_{s,t} \tilde{x}_{is} \kappa^{st} \tilde{x}_{it} \right),
\end{aligned}$$

$$\begin{aligned}
& \sum_{r,s} \kappa^{br} \kappa^{s\theta} \left(\kappa_{rs}^{(\theta)} - \frac{1}{2} \kappa_{rs\theta} \right) = \\
&= \frac{1}{\theta^3} \left(1 - \frac{1}{2\lambda^2} \right) \sum_{r,s} \kappa^{br} \kappa^{s\theta} \sum_i \tilde{x}_{ir} \tilde{x}_{is} + \frac{1}{2\lambda\theta^2} \sum_{r,s} \kappa^{br} \kappa^{s\theta} \sum_i \tilde{x}_{irs} \\
&= \frac{1}{\theta^3} \left(1 - \frac{1}{2\lambda^2} \right) \sum_i \left(\sum_r \kappa^{br} \tilde{x}_{ir} \right) \left(\sum_s \kappa^{s\theta} \tilde{x}_{is} \right) + \frac{1}{2\lambda\theta^2} \sum_i \left(\sum_{r,s} \kappa^{br} \tilde{x}_{irs} \kappa^{s\theta} \right),
\end{aligned}$$

$$\begin{aligned}
& \sum_{r,t} \kappa^{br} \kappa^{\theta t} \left(\kappa_{r\theta}^{(t)} - \frac{1}{2} \kappa_{r\theta t} \right) = \\
&= -\frac{1}{\theta^3} \left(1 + \frac{1}{2\lambda^2} \right) \sum_{r,t} \kappa^{br} \kappa^{\theta t} \sum_i \tilde{x}_{ir} \tilde{x}_{it} - \frac{1}{2\lambda\theta^2} \sum_{r,t} \kappa^{br} \kappa^{\theta t} \sum_i \tilde{x}_{irt} \\
&= -\frac{1}{\theta^3} \left(1 + \frac{1}{2\lambda^2} \right) \sum_i \left(\sum_r \kappa^{br} \tilde{x}_{ir} \right) \left(\sum_t \kappa^{\theta t} \tilde{x}_{it} \right) - \frac{1}{2\lambda\theta^2} \sum_i \left(\sum_{r,t} \kappa^{br} \tilde{x}_{irt} \kappa^{\theta t} \right),
\end{aligned}$$

$$\begin{aligned}
& \sum_{s,t} \kappa^{b\theta} \kappa^{st} \left(\kappa_{\theta s}^{(t)} - \frac{1}{2} \kappa_{\theta st} \right) = \\
&= -\frac{1}{\theta^3} \left(1 + \frac{1}{2\lambda^2} \right) \sum_{s,t} \kappa^{b\theta} \kappa^{st} \sum_i \tilde{x}_{is} \tilde{x}_{it} - \frac{1}{2\lambda\theta^2} \sum_{s,t} \kappa^{b\theta} \kappa^{st} \sum_i \tilde{x}_{ist} \\
&= -\frac{1}{\theta^3} \left(1 + \frac{1}{2\lambda^2} \right) \kappa^{b\theta} \sum_i \left(\sum_{s,t} \tilde{x}_{is} \kappa^{st} \tilde{x}_{it} \right) - \frac{1}{2\lambda\theta^2} \kappa^{b\theta} \sum_i \left(\sum_{s,t} \kappa^{st} \tilde{x}_{ist} \right),
\end{aligned}$$

$$\begin{aligned}
\sum_r \kappa^{br} \kappa^{\theta\theta} \left(\kappa_{r\theta}^{(\theta)} - \frac{1}{2} \kappa_{r\theta\theta} \right) &= -\frac{\lambda \psi^{(1)}(\lambda^2)}{2\theta^3} \kappa^{\theta\theta} \sum_i \left(\sum_r \kappa^{br} \tilde{x}_{ir} \right), \\
\sum_s \kappa^{b\theta} \kappa^{s\theta} \left(\kappa_{\theta s}^{(\theta)} - \frac{1}{2} \kappa_{\theta s\theta} \right) &= -\frac{\lambda \psi^{(1)}(\lambda^2)}{2\theta^3} \kappa^{b\theta} \sum_i \left(\sum_s \kappa^{s\theta} \tilde{x}_{is} \right), \\
\sum_t \kappa^{b\theta} \kappa^{\theta t} \left(\kappa_{\theta\theta}^{(t)} - \frac{1}{2} \kappa_{\theta\theta t} \right) &= -\frac{1}{2\theta^3} \left(\frac{4}{\lambda} + \lambda \psi^{(1)}(\lambda^2) \right) \kappa^{b\theta} \sum_i \left(\sum_t \kappa^{\theta t} \tilde{x}_{it} \right), \\
\kappa^{b\theta} \kappa^{\theta\theta} \left(\kappa_{\theta\theta}^{(\theta)} - \frac{1}{2} \kappa_{\theta\theta\theta} \right) &= -\kappa^{b\theta} \kappa^{\theta\theta} \frac{n}{2\theta^3} \left\{ 2\lambda^2 \psi^{(1)}(\lambda^2) + 3\psi^{(1)}(\lambda^2) - \lambda^2 \psi^{(2)}(\lambda^2) \right\}.
\end{aligned}$$

Os termos de (11) são dados substituindo o índice sobrescrito b pelo θ .

Apêndice C

Cálculo do Fator de Correção de Bartlett

Nesta seção, apresentamos a obtenção de expressões gerais para as quantidades necessárias ao cálculo do fator da correção de Bartlett da estatística da razão de verossimilhanças, apresentada no Capítulo 4. Seja Σ o somatório sobre o vetor ξ . Definimos a seguir quantidades, para $i, j = 1, \dots, n$, que serão de grande valia para a simplificação dos cálculos:

$$\begin{aligned}
 f_{ij} &= \left(\sum \mathbf{K}^{tu} \tilde{x}_{irtu} \mathbf{K}^{rs} \tilde{x}_{js} \right), \\
 g_{ij} &= \left(\sum \tilde{x}_{ir} \mathbf{K}^{rs} \tilde{x}_{jsw} \mathbf{K}^{vw} \tilde{x}_{itv} \mathbf{K}^{tu} \tilde{x}_{ju} \right), \\
 h_{ij} &= \left(\sum \tilde{x}_{ir} \mathbf{K}^{rs} \tilde{x}_{isu} \mathbf{K}^{tu} \tilde{x}_{jt} \right), \\
 l_{ij} &= \left(\sum \tilde{x}_{ir} \mathbf{K}^{rs} \tilde{x}_{isw} \mathbf{K}^{vw} \tilde{x}_{jtv} \mathbf{K}^{tu} \tilde{x}_{ju} \right).
 \end{aligned}$$

C.1 Cálculo da Primeira Parcela $\sum \lambda_{rstu}$

$$\begin{aligned}
\lambda_{rstu} &= \kappa^{rs} \kappa^{tu} \left(\kappa_{rstu}/4 - \kappa_{rst}^{(u)} + \kappa_{rt}^{(su)} \right) \\
&= \kappa^{rs} \kappa^{tu} \left(-\frac{1}{4\theta^4 \lambda^2} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} + \frac{1}{4\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu} + \frac{1}{4\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{isu} \tilde{x}_{it} \right. \\
&\quad + \frac{1}{4\theta^3 \lambda} \sum_i \tilde{x}_{iru} \tilde{x}_{is} \tilde{x}_{it} + \frac{1}{4\theta^3 \lambda} \sum_i \tilde{x}_{iu} \tilde{x}_{irt} \tilde{x}_{is} + \frac{1}{4\theta^3 \lambda} \sum_i \tilde{x}_{iu} \tilde{x}_{ir} \tilde{x}_{ist} + \frac{1}{4\theta^3 \lambda} \sum_i \tilde{x}_{iu} \tilde{x}_{irs} \tilde{x}_{it} \\
&\quad - \frac{1}{4\theta^2} \sum_i \tilde{x}_{irtu} \tilde{x}_{is} - \frac{1}{4\theta^2} \sum_i \tilde{x}_{istu} \tilde{x}_{ir} - \frac{1}{4\theta^2} \sum_i \tilde{x}_{irsu} \tilde{x}_{it} - \frac{1}{4\theta^2} \sum_i \tilde{x}_{irst} \tilde{x}_{iu} \\
&\quad - \frac{1}{4\theta^2} \sum_i \tilde{x}_{irt} \tilde{x}_{isu} - \frac{1}{4\theta^2} \sum_i \tilde{x}_{iru} \tilde{x}_{ist} - \frac{1}{4\theta^2} \sum_i \tilde{x}_{irs} \tilde{x}_{itu} - \frac{1}{\lambda \theta^3} \sum_i \tilde{x}_{iru} \tilde{x}_{is} \tilde{x}_{it} \\
&\quad - \frac{1}{\lambda \theta^3} \sum_i \tilde{x}_{ir} \tilde{x}_{isu} \tilde{x}_{it} - \frac{1}{\lambda \theta^3} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu} + \frac{1}{\theta^2} \sum_i \tilde{x}_{irtu} \tilde{x}_{is} + \frac{1}{\theta^2} \sum_i \tilde{x}_{irt} \tilde{x}_{isu} \\
&\quad + \frac{1}{\theta^2} \sum_i \tilde{x}_{iru} \tilde{x}_{ist} + \frac{1}{\theta^2} \sum_i \tilde{x}_{ir} \tilde{x}_{istu} + \frac{1}{\theta^2} \sum_i \tilde{x}_{irsu} \tilde{x}_{it} + \frac{1}{\theta^2} \sum_i \tilde{x}_{irs} \tilde{x}_{itu} \\
&\quad \left. - \frac{1}{\theta^2} \sum_i \tilde{x}_{irsu} \tilde{x}_{it} - \frac{1}{\theta^2} \sum_i \tilde{x}_{irs} \tilde{x}_{itu} - \frac{1}{\theta^2} \sum_i \tilde{x}_{iru} \tilde{x}_{its} - \frac{1}{\theta^2} \sum_i \tilde{x}_{ir} \tilde{x}_{istu} \right) \\
\lambda_{rstu} &= \kappa^{rs} \kappa^{tu} \left(-\frac{1}{4\theta^4 \lambda^2} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} - \frac{3}{4\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu} - \frac{3}{4\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{isu} \tilde{x}_{it} \right. \\
&\quad - \frac{3}{4\theta^3 \lambda} \sum_i \tilde{x}_{iru} \tilde{x}_{is} \tilde{x}_{it} + \frac{1}{4\theta^3 \lambda} \sum_i \tilde{x}_{iu} \tilde{x}_{irt} \tilde{x}_{is} + \frac{1}{4\theta^3 \lambda} \sum_i \tilde{x}_{iu} \tilde{x}_{ir} \tilde{x}_{ist} + \frac{1}{4\theta^3 \lambda} \sum_i \tilde{x}_{iu} \tilde{x}_{irs} \tilde{x}_{it} \\
&\quad + \frac{3}{4\theta^2} \sum_i \tilde{x}_{irtu} \tilde{x}_{is} - \frac{1}{4\theta^2} \sum_i \tilde{x}_{istu} \tilde{x}_{ir} - \frac{1}{4\theta^2} \sum_i \tilde{x}_{irsu} \tilde{x}_{it} - \frac{1}{4\theta^2} \sum_i \tilde{x}_{irst} \tilde{x}_{iu} \\
&\quad \left. + \frac{3}{4\theta^2} \sum_i \tilde{x}_{irt} \tilde{x}_{isu} - \frac{1}{4\theta^2} \sum_i \tilde{x}_{iru} \tilde{x}_{ist} - \frac{1}{4\theta^2} \sum_i \tilde{x}_{irs} \tilde{x}_{itu} \right)
\end{aligned}$$

$$\begin{aligned}
\sum \lambda_{rstu} = & -\frac{1}{4\theta^4\lambda^2} \sum_i (\sum \tilde{x}_{ir} \mathbf{K}^{rs} \tilde{x}_{is}) (\sum \tilde{x}_{it} \mathbf{K}^{tu} \tilde{x}_{iu}) - \frac{3}{4\theta^3\lambda} \sum_i (\sum \tilde{x}_{ir} \mathbf{K}^{rs} \tilde{x}_{is}) (\sum \tilde{x}_{itu} \mathbf{K}^{tu}) \\
& - \frac{3}{4\theta^3\lambda} \sum_i (\sum \tilde{x}_{ir} \mathbf{K}^{rs} \tilde{x}_{isu} \mathbf{K}^{tu} \tilde{x}_{it}) - \frac{3}{4\theta^3\lambda} \sum_i (\sum \tilde{x}_{it} \mathbf{K}^{tu} \tilde{x}_{iru} \mathbf{K}^{rs} \tilde{x}_{is}) \\
& + \frac{1}{4\theta^3\lambda} \sum_i (\sum \tilde{x}_{iu} \mathbf{K}^{tu} \tilde{x}_{irt} \mathbf{K}^{rs} \tilde{x}_{is}) + \frac{1}{4\theta^3\lambda} \sum_i (\sum \tilde{x}_{ir} \mathbf{K}^{rs} \tilde{x}_{ist} \mathbf{K}^{tu} \tilde{x}_{iu}) \\
& + \frac{1}{4\theta^3\lambda} \sum_i (\sum \tilde{x}_{irs} \mathbf{K}^{rs}) (\sum \tilde{x}_{it} \mathbf{K}^{tu} \tilde{x}_{iu}) + \frac{3}{4\theta^2} \sum_i (\sum \mathbf{K}^{tu} \tilde{x}_{irtu} \mathbf{K}^{rs} \tilde{x}_{is}) \\
& - \frac{1}{4\theta^2} \sum_i (\sum \mathbf{K}^{tu} \tilde{x}_{istu} \mathbf{K}^{rs} \tilde{x}_{ir}) - \frac{1}{4\theta^2} \sum_i (\sum \mathbf{K}^{rs} \tilde{x}_{irsu} \mathbf{K}^{tu} \tilde{x}_{it}) \\
& - \frac{1}{4\theta^2} \sum_i (\sum \mathbf{K}^{rs} \tilde{x}_{irst} \mathbf{K}^{tu} \tilde{x}_{iu}) + \frac{3}{4\theta^2} \sum_i (\sum \tilde{x}_{irt} \mathbf{K}^{rs} \tilde{x}_{isu} \mathbf{K}^{tu}) \\
& - \frac{1}{4\theta^2} \sum_i (\sum \tilde{x}_{iru} \mathbf{K}^{rs} \tilde{x}_{ist} \mathbf{K}^{tu}) - \frac{1}{4\theta^2} \sum_i (\sum \tilde{x}_{irs} \mathbf{K}^{rs}) (\sum \tilde{x}_{itu} \mathbf{K}^{tu})
\end{aligned}$$

$$\begin{aligned}
\sum \lambda_{rstu} = & -\frac{1}{4\theta^4\lambda^2} \sum_i z_{ii}^2 - \frac{3}{4\theta^3\lambda} \sum_i z_{ii} d_i - \frac{3}{4\theta^3\lambda} \sum_i c_{ii} - \frac{3}{4\theta^3\lambda} \sum_i c_{ii} + \frac{1}{4\theta^3\lambda} \sum_i c_{ii} \\
& + \frac{1}{4\theta^3\lambda} \sum_i c_{ii} + \frac{1}{4\theta^3\lambda} \sum_i d_i z_{ii} + \frac{3}{4\theta^2} \sum_i f_{ii} - \frac{1}{4\theta^2} \sum_i f_{ii} - \frac{1}{4\theta^2} \sum_i f_{ii} \\
& - \frac{1}{4\theta^2} \sum_i f_{ii} + \frac{3}{4\theta^2} \sum_i b_{ii} - \frac{1}{4\theta^2} \sum_i b_{ii} - \frac{1}{4\theta^2} \sum_i d_i^2
\end{aligned}$$

$$\sum \lambda_{rstu} = -\frac{1}{4\theta^4\lambda^2} \sum_i z_{ii}^2 - \frac{1}{2\theta^3\lambda} \sum_i z_{ii} d_i - \frac{1}{\theta^3\lambda} \sum_i c_{ii} + \frac{1}{2\theta^2} \sum_i b_{ii} - \frac{1}{4\theta^2} \sum_i d_i^2$$

C.2 Cálculo da Segunda Parcela $\sum \lambda_{rstuvw}$

$$\begin{aligned}
& \kappa_{rtv} \left(\kappa_{suw}/6 - \kappa_{sw}^{(u)} \right) = \\
& = \left\{ \frac{1}{\lambda \theta^3} \sum_i \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} - \frac{1}{\theta^2} \sum_i \tilde{x}_{irv} \tilde{x}_{it} - \frac{1}{\theta^2} \sum_i \tilde{x}_{ir} \tilde{x}_{itv} - \frac{1}{\theta^2} \sum_i \tilde{x}_{irt} \tilde{x}_{iv} \right\} \times \\
& \left\{ \frac{1}{6\lambda \theta^3} \sum_j \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} - \frac{1}{6\theta^2} \sum_j \tilde{x}_{jsw} \tilde{x}_{ju} + \frac{5}{6\theta^2} \sum_j \tilde{x}_{js} \tilde{x}_{juw} + \frac{5}{6\theta^2} \sum_j \tilde{x}_{jsu} \tilde{x}_{jw} \right\} \\
& = \frac{1}{6\lambda^2 \theta^6} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} \\
& - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{jsw} \tilde{x}_{ju} + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{jsw} \tilde{x}_{ju} \\
& + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{jsw} \tilde{x}_{ju} + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{jsw} \tilde{x}_{ju} + \frac{5}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{juw} \\
& - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{juw} - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{js} \tilde{x}_{juw} - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{juw} \\
& + \frac{5}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{jsu} \tilde{x}_{jw} - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{jsu} \tilde{x}_{jw} - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{jsu} \tilde{x}_{jw} \\
& - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{jsu} \tilde{x}_{jw}
\end{aligned}$$

$$\begin{aligned}
& \kappa_{rtu} \left(\kappa_{svw}/4 - \kappa_{sw}^{(v)} \right) = \\
& = \left\{ \frac{1}{\lambda \theta^3} \sum_i \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} - \frac{1}{\theta^2} \sum_i \tilde{x}_{iru} \tilde{x}_{it} - \frac{1}{\theta^2} \sum_i \tilde{x}_{ir} \tilde{x}_{itu} - \frac{1}{\theta^2} \sum_i \tilde{x}_{irt} \tilde{x}_{iu} \right\} \times \\
& \left\{ \frac{1}{4\lambda \theta^3} \sum_j \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} - \frac{1}{4\theta^2} \sum_j \tilde{x}_{jsw} \tilde{x}_{jv} + \frac{3}{4\theta^2} \sum_j \tilde{x}_{js} \tilde{x}_{jvw} + \frac{3}{4\theta^2} \sum_j \tilde{x}_{jsv} \tilde{x}_{jw} \right\} \\
& = \frac{1}{4\lambda^2 \theta^6} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} \\
& - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} \\
& - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{jsw} \tilde{x}_{jv} + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{jsw} \tilde{x}_{jv} + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{jsw} \tilde{x}_{jv} \\
& + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{jsw} \tilde{x}_{jv} + \frac{3}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jvw} - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{jvw} \\
& - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{js} \tilde{x}_{jvw} - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jvw} + \frac{3}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{jsv} \tilde{x}_{jw} \\
& - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{jsv} \tilde{x}_{jw} - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{jsv} \tilde{x}_{jw} - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{jsv} \tilde{x}_{jw} \\
\kappa_{rt}^{(v)} \kappa_{sw}^{(u)} & = \left\{ -\frac{1}{\theta^2} \sum_i \tilde{x}_{irv} \tilde{x}_{it} - \frac{1}{\theta^2} \sum_i \tilde{x}_{ir} \tilde{x}_{itv} \right\} \left\{ -\frac{1}{\theta^2} \sum_j \tilde{x}_{jsu} \tilde{x}_{jw} - \frac{1}{\theta^2} \sum_j \tilde{x}_{js} \tilde{x}_{jwu} \right\} \\
& = \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{jsu} \tilde{x}_{jw} + \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{jsu} \tilde{x}_{jw} + \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{jwu} \\
& + \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{js} \tilde{x}_{jwu}
\end{aligned}$$

$$\begin{aligned}
\lambda_{rstuvw} &= \kappa^{rs} \kappa^{tu} \kappa^{vw} \left\{ \kappa_{rtv} \left(\kappa_{suw}/6 - \kappa_{sw}^{(u)} \right) + \kappa_{rtu} \left(\kappa_{svw}/4 - \kappa_{sw}^{(v)} \right) + \kappa_{rt}^{(v)} \kappa_{sw}^{(u)} + \kappa_{rt}^{(u)} \kappa_{sw}^{(v)} \right\} \\
&= \kappa^{rs} \kappa^{tu} \kappa^{vw} \left\{ \frac{1}{6\lambda^2 \theta^6} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} \right. \\
&\quad - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} \\
&\quad - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{jsw} \tilde{x}_{ju} + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{jsw} \tilde{x}_{ju} + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{itv} \tilde{x}_{jsw} \tilde{x}_{ju} \\
&\quad + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{jsw} \tilde{x}_{ju} + \frac{5}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{juw} - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{juw} \\
&\quad - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{js} \tilde{x}_{juw} - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{juw} + \frac{5}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{jsu} \tilde{x}_{jw} \\
&\quad - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{jsu} \tilde{x}_{jw} - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{jsu} \tilde{x}_{jw} - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{jsu} \tilde{x}_{jw} \\
&\quad + \frac{1}{4\lambda^2 \theta^6} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} \\
&\quad - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} \\
&\quad - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{jsw} \tilde{x}_{jv} + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{jsw} \tilde{x}_{jv} + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{jsw} \tilde{x}_{jv} \\
&\quad + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{jsw} \tilde{x}_{jv} + \frac{3}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jvw} - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{jvw} \\
&\quad - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{js} \tilde{x}_{jvw} - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jvw} + \frac{3}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{jsv} \tilde{x}_{jw} \\
&\quad - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{jsv} \tilde{x}_{jw} - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{jsv} \tilde{x}_{jw} - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{jsv} \tilde{x}_{jw} \\
&\quad + \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{jsu} \tilde{x}_{jw} + \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{jsu} \tilde{x}_{jw} + \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{jwu} \\
&\quad + \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{js} \tilde{x}_{jwu} + \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{jsv} \tilde{x}_{jw} + \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{jsv} \tilde{x}_{jw} \\
&\quad \left. + \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{jvw} + \frac{1}{\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{js} \tilde{x}_{jvw} \right\}
\end{aligned}$$

$$\begin{aligned}
\lambda_{rstuvw} = & \kappa^{rs} \kappa^{tu} \kappa^{vw} \left\{ \frac{1}{6\lambda^2 \theta^6} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} \right. \\
& - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{ju} \tilde{x}_{jw} \\
& - \frac{1}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{jsw} \tilde{x}_{ju} + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{jsw} \tilde{x}_{ju} + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{jsw} \tilde{x}_{ju} \\
& + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{jsw} \tilde{x}_{ju} + \frac{5}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{juw} + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{juw} \\
& + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{juw} - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{js} \tilde{x}_{juw} + \frac{5}{6\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iv} \tilde{x}_{jsw} \tilde{x}_{juw} \\
& + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{irv} \tilde{x}_{it} \tilde{x}_{jsw} \tilde{x}_{juw} + \frac{1}{6\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itv} \tilde{x}_{jsw} \tilde{x}_{juw} - \frac{5}{6\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iv} \tilde{x}_{jsw} \tilde{x}_{juw} \\
& + \frac{1}{4\lambda^2 \theta^6} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} \\
& - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jv} \tilde{x}_{jw} \\
& - \frac{1}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{jsw} \tilde{x}_{jv} + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{jsw} \tilde{x}_{jv} + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{jsw} \tilde{x}_{jv} \\
& + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{jsw} \tilde{x}_{jv} + \frac{3}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jvw} + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{js} \tilde{x}_{jvw} \\
& + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{js} \tilde{x}_{jvw} - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{js} \tilde{x}_{jvw} + \frac{3}{4\lambda \theta^5} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{it} \tilde{x}_{iu} \tilde{x}_{jsw} \tilde{x}_{jw} \\
& \left. + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{iru} \tilde{x}_{it} \tilde{x}_{jsw} \tilde{x}_{jw} + \frac{1}{4\theta^4} \sum_{i,j} \tilde{x}_{ir} \tilde{x}_{itu} \tilde{x}_{jsw} \tilde{x}_{jw} - \frac{3}{4\theta^4} \sum_{i,j} \tilde{x}_{irt} \tilde{x}_{iu} \tilde{x}_{jsw} \tilde{x}_{jw} \right\}
\end{aligned}$$

$$\begin{aligned}
\sum \lambda_{rstuvw} = & -\frac{1}{6\lambda^2\theta^6} \sum_{i,j} z_{ij}^3 + \frac{1}{6\lambda\theta^5} \sum_{i,j} z_{ij}c_{ji} + \frac{1}{6\lambda\theta^5} \sum_{i,j} z_{ij}c_{ji} + \frac{1}{6\lambda\theta^5} \sum_{i,j} c_{ji}z_{ij} \\
& + \frac{1}{6\lambda\theta^5} \sum_{i,j} c_{ij}z_{ij} - \frac{1}{6\theta^4} \sum_{i,j} z_{ij}b_{ji} + \frac{1}{6\theta^4} \sum_{i,j} g_{ij} + \frac{1}{6\theta^4} \sum_{i,j} g_{ij} \\
& - \frac{5}{6\lambda\theta^5} \sum_{i,j} z_{ij}c_{ij} + \frac{1}{6\theta^4} \sum_{i,j} g_{ij} - \frac{1}{6\theta^4} \sum_{i,j} z_{ij}b_{ji} - \frac{5}{6\theta^4} \sum_{i,j} g_{ij} \\
& - \frac{5}{6\lambda\theta^5} \sum_{i,j} c_{ij}z_{ij} + \frac{1}{6\theta^4} \sum_{i,j} g_{ij} + \frac{1}{6\theta^4} \sum_{i,j} g_{ij} + \frac{5}{6\theta^4} \sum_{i,j} b_{ji}z_{ij} \\
& - \frac{1}{4\lambda^2\theta^6} \sum_{i,j} z_{ij}z_{ii}z_{jj} + \frac{1}{4\lambda\theta^5} \sum_{i,j} h_{ij}z_{jj} + \frac{1}{4\lambda\theta^5} \sum_{i,j} z_{ij}d_i z_{jj} \\
& + \frac{1}{4\lambda\theta^5} \sum_{i,j} h_{ij}z_{jj} + \frac{1}{4\lambda\theta^5} \sum_{i,j} h_{ji}z_{ii} + \frac{1}{4\theta^4} \sum_{i,j} l_{ij} - \frac{1}{4\theta^4} \sum_{i,j} h_{ji}d_i \\
& + \frac{1}{4\theta^4} \sum_{i,j} l_{ij} - \frac{3}{4\lambda\theta^5} \sum_{i,j} z_{ij}z_{ii}d_j - \frac{1}{4\theta^4} \sum_{i,j} h_{ij}d_j - \frac{1}{4\theta^4} \sum_{i,j} z_{ij}d_i d_j \\
& + \frac{3}{4\theta^4} \sum_{i,j} h_{ij}d_j - \frac{3}{4\lambda\theta^5} \sum_{i,j} h_{ji}z_{ii} + \frac{1}{4\theta^4} \sum_{i,j} l_{ij} - \frac{1}{4\theta^4} \sum_{i,j} h_{ji}d_i \\
& - \frac{3}{4\theta^4} \sum_{i,j} l_{ij}
\end{aligned}$$

$$\begin{aligned}
\sum \lambda_{rstuvw} = & -\frac{1}{6\lambda^2\theta^6} \sum_{i,j} z_{ij}^3 - \frac{1}{\lambda\theta^5} \sum_{i,j} z_{ij}c_{ji} + \frac{1}{2\theta^4} \sum_{i,j} z_{ij}b_{ji} - \frac{1}{4\lambda^2\theta^6} \sum_{i,j} z_{ij}z_{ii}z_{jj} \\
& - \frac{1}{2\lambda\theta^5} \sum_{i,j} z_{ij}d_i z_{jj} - \frac{1}{4\theta^4} \sum_{i,j} z_{ij}d_i d_j
\end{aligned}$$

Apêndice D

Cálculo do Fator de Correção Tipo-Bartlett

Neste apêndice, apresentamos a obtenção de expressões gerais para as quantidades necessárias ao cálculo dos fatores da correção tipo-Bartlett da estatística escore para a classe dos MNLLGG e da estatística gradiente para a classe dos MLLGG, com os parâmetros λ e θ considerados fixos, apresentados na Seção 4.5. Mais precisamente, vamos descrever a obtenção das quantidades A 's, dadas em (26)–(28) e (30)–(32). Sejam $-\kappa^{su} = \kappa^{s,u}$, o elemento (s, u) da inversa da matriz de informação de Fisher K_β e, a^{rs} e m^{rs} , os elementos (r, s) das matrizes

$$\mathbf{A} = \begin{bmatrix} 0 & 0 \\ 0 & \kappa_{22}^{-1} \end{bmatrix} \text{ e } M = K_\beta^{-1} - A,$$

respectivamente. Para a simplificação dos cálculos, considere as seguintes definições: $z_{ij} = -\sum x_{is} \kappa^{su} x_{ju}$, $z_{aij} = \sum x_{is} a^{su} x_{ju}$, $z_{mij} = \sum x_{is} m^{su} x_{ju}$, $b_{aij} = \sum a^{tw} \tilde{x}_{irt} a^{ru} \tilde{x}_{juw}$, $d_{ai} = \sum a^{rs} \tilde{x}_{irs}$ e $c_{aij} = \sum \tilde{x}_{is} a^{sv} \tilde{x}_{jvw} a^{tw} \tilde{x}_{it}$, em que $i, j = 1, \dots, n$ e \sum o somatório sobre o vetor β . Pela definições das matrizes A e M , temos a seguinte relação $z_{mij} = z_{ij} - z_{aij}$. Todos os cumulantes necessários para os cálculos das quantidades A 's são apresentados no Apêndice A.

D.1 Correção Tipo-Bartlett à Estatística Escore

Nesta seção detalhamos a obtenção das quantidades (26)–(28) que compõem o fator da correção tipo-Bartlett para a estatística escore definido na Subseção 4.5.1 no MNLLGG. Seja $A_1 = A_{11} + A_{12} + A_{13} + A_{14}$, em que

$$\begin{aligned} A_{11} &= 3 \sum (\kappa_{rst} + 2\kappa_{r,st}) (\kappa_{uvw} + 2\kappa_{uv,w}) a^{rs} a^{vw} m^{tu} \\ A_{12} &= -6 \sum (\kappa_{rst} + 2\kappa_{r,st}) \kappa_{u,v,w} a^{rs} a^{tu} m^{vw} \\ A_{13} &= 6 \sum (\kappa_{r,st} - \kappa_{r,s,t}) (\kappa_{uvw} + 2\kappa_{uv,w}) a^{sv} a^{tw} m^{ru} \\ A_{14} &= -6 \sum (\kappa_{r,s,t,u} + \kappa_{r,s,tu}) a^{tu} m^{rs}. \end{aligned}$$

Dessa forma, temos que

$$\begin{aligned}
A_{11} &= 3 \sum \left\{ -\frac{1}{\lambda \theta^3} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} - \frac{1}{\theta^2} \sum_i \tilde{x}_{irt} \tilde{x}_{is} + \frac{1}{\theta^2} \sum_i \tilde{x}_{ir} \tilde{x}_{ist} - \frac{1}{\theta^2} \sum_i \tilde{x}_{irs} \tilde{x}_{it} \right\} \times \\
&\quad \left\{ -\frac{1}{\lambda \theta^3} \sum_j \tilde{x}_{jv} \tilde{x}_{ju} \tilde{x}_{jv} - \frac{1}{\theta^2} \sum_j \tilde{x}_{jvw} \tilde{x}_{ju} + \frac{1}{\theta^2} \sum_j \tilde{x}_{jv} \tilde{x}_{juv} - \frac{1}{\theta^2} \sum_j \tilde{x}_{jwu} \tilde{x}_{jv} \right\} a^{rs} a^{vw} m^{tu} \\
&= \frac{3}{\lambda^2 \theta^6} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{is}) (\sum \tilde{x}_{it} m^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} a^{vw} \tilde{x}_{jw}) \\
&\quad + \frac{3}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{is}) (\sum \tilde{x}_{it} m^{tu} \tilde{x}_{ju}) (\sum a^{vw} \tilde{x}_{jvw}) - \frac{3}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{is}) (\sum \tilde{x}_{it} m^{tu} \tilde{x}_{juv} a^{vw} \tilde{x}_{jw}) \\
&\quad + \frac{3}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{is}) (\sum \tilde{x}_{it} m^{tu} \tilde{x}_{jwu} a^{vw} \tilde{x}_{jv}) + \frac{3}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{is} a^{rs} \tilde{x}_{irt} m^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} a^{vw} \tilde{x}_{jw}) \\
&\quad + \frac{3}{\theta^4} \sum_{i,j} (\sum \tilde{x}_{is} a^{rs} \tilde{x}_{irt} m^{tu} \tilde{x}_{ju}) (\sum a^{vw} \tilde{x}_{jvw}) - \frac{3}{\theta^4} \sum_{i,j} (\sum \tilde{x}_{is} a^{rs} \tilde{x}_{irt} m^{tu} \tilde{x}_{juv} a^{vw} \tilde{x}_{jw}) \\
&\quad + \frac{3}{\theta^4} \sum_{i,j} (\sum \tilde{x}_{is} a^{rs} \tilde{x}_{irt} m^{tu} \tilde{x}_{jwu} a^{vw} \tilde{x}_{jv}) - \frac{3}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{ist} m^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} a^{vw} \tilde{x}_{jw}) \\
&\quad - \frac{3}{\theta^4} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{ist} m^{tu} \tilde{x}_{ju}) (\sum a^{vw} \tilde{x}_{jvw}) + \frac{3}{\theta^4} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{ist} m^{tu} \tilde{x}_{juv} a^{vw} \tilde{x}_{jw}) \\
&\quad - \frac{3}{\theta^4} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{ist} m^{tu} \tilde{x}_{jwu} a^{vw} \tilde{x}_{jv}) + \frac{3}{\lambda \theta^5} \sum_{i,j} (\sum a^{rs} \tilde{x}_{irs}) (\sum \tilde{x}_{it} m^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} a^{vw} \tilde{x}_{jw}) \\
&\quad + \frac{3}{\theta^4} \sum_{i,j} (\sum a^{rs} \tilde{x}_{irs}) (\sum \tilde{x}_{it} m^{tu} \tilde{x}_{ju}) (\sum a^{vw} \tilde{x}_{jvw}) - \frac{3}{\theta^4} \sum_{i,j} (\sum a^{rs} \tilde{x}_{irs}) (\sum \tilde{x}_{it} m^{tu} \tilde{x}_{juv} a^{vw} \tilde{x}_{jw}) \\
&\quad + \frac{3}{\theta^4} \sum_{i,j} (\sum a^{rs} \tilde{x}_{irs}) (\sum \tilde{x}_{it} m^{tu} \tilde{x}_{jwu} a^{vw} \tilde{x}_{jv})
\end{aligned}$$

$$\begin{aligned}
A_{12} &= -6 \sum \left\{ -\frac{1}{\lambda \theta^3} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} - \frac{1}{\theta^2} \sum_i \tilde{x}_{irt} \tilde{x}_{is} + \frac{1}{\theta^2} \sum_i \tilde{x}_{ir} \tilde{x}_{ist} - \frac{1}{\theta^2} \sum_i \tilde{x}_{irs} \tilde{x}_{it} \right\} \times \\
&\quad \left\{ \frac{2}{\theta^3 \lambda} \sum_j \tilde{x}_{ju} \tilde{x}_{jv} \tilde{x}_{jw} \right\} a^{rs} a^{tu} m^{vw} \\
&= \frac{12}{\lambda^2 \theta^6} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{is}) (\sum \tilde{x}_{it} a^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} m^{vw} \tilde{x}_{jw}) \\
&\quad + \frac{12}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{is} a^{rs} \tilde{x}_{irt} a^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} m^{vw} \tilde{x}_{jw}) \\
&\quad - \frac{12}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{ist} a^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} m^{vw} \tilde{x}_{jw}) \\
&\quad + \frac{12}{\lambda \theta^5} \sum_{i,j} (\sum a^{rs} \tilde{x}_{irs}) (\sum \tilde{x}_{it} a^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} m^{vw} \tilde{x}_{jw})
\end{aligned}$$

$$\begin{aligned}
A_{13} &= 6 \sum \left\{ \frac{1}{\theta^2} \sum_i \tilde{x}_{ir} \tilde{x}_{ist} - \frac{3}{\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \right\} \times \\
&\quad \left\{ -\frac{1}{\lambda \theta^3} \sum_j \tilde{x}_{jw} \tilde{x}_{ju} \tilde{x}_{jv} - \frac{1}{\theta^2} \sum_j \tilde{x}_{jwv} \tilde{x}_{ju} + \frac{1}{\theta^2} \sum_j \tilde{x}_{jw} \tilde{x}_{juv} - \frac{1}{\theta^2} \sum_j \tilde{x}_{jwu} \tilde{x}_{jv} \right\} a^{sv} a^{tw} m^{ru} \\
&= -\frac{6}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{ir} m^{ru} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} a^{sv} \tilde{x}_{ist} a^{tw} \tilde{x}_{jw}) - \frac{6}{\theta^4} \sum_{i,j} (\sum \tilde{x}_{ir} m^{ru} \tilde{x}_{ju}) (\sum \tilde{x}_{jwv} a^{sv} \tilde{x}_{ist} a^{tw}) \\
&\quad + \frac{6}{\theta^4} \sum_{i,j} (\sum \tilde{x}_{ir} m^{ru} \tilde{x}_{juv} a^{sv} \tilde{x}_{ist} a^{tw} \tilde{x}_{jw}) - \frac{6}{\theta^4} \sum_{i,j} (\sum \tilde{x}_{ir} m^{ru} \tilde{x}_{jwu} a^{tw} \tilde{x}_{ist} a^{sv} \tilde{x}_{jv}) \\
&\quad + \frac{18}{\lambda^2 \theta^6} \sum_{i,j} (\sum \tilde{x}_{ir} m^{ru} \tilde{x}_{ju}) (\sum \tilde{x}_{is} a^{sv} \tilde{x}_{jv}) (\sum \tilde{x}_{it} a^{tw} \tilde{x}_{jw}) \\
&\quad + \frac{18}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{ir} m^{ru} \tilde{x}_{ju}) (\sum \tilde{x}_{is} a^{sv} \tilde{x}_{jwv} a^{tw} \tilde{x}_{it}) \\
&\quad - \frac{18}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{ir} m^{ru} \tilde{x}_{juv} a^{sv} \tilde{x}_{is}) (\sum \tilde{x}_{it} a^{tw} \tilde{x}_{jw}) \\
&\quad + \frac{18}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{ir} m^{ru} \tilde{x}_{jwu} a^{tw} \tilde{x}_{it}) (\sum \tilde{x}_{is} a^{sv} \tilde{x}_{jv})
\end{aligned}$$

e

$$\begin{aligned}
A_{14} &= -6 \sum \left\{ \frac{6}{\theta^4 \lambda^2} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} - \frac{2}{\theta^4 \lambda^2} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} + \frac{2}{\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{itu} \right\} a^{tu} m^{rs} \\
&= -\frac{24}{\theta^4 \lambda^2} \sum_i (\sum \tilde{x}_{ir} m^{rs} \tilde{x}_{is}) (\sum \tilde{x}_{it} a^{tu} \tilde{x}_{iu}) - \frac{12}{\theta^3 \lambda} \sum_i (\sum \tilde{x}_{ir} m^{rs} \tilde{x}_{is}) (\sum a^{tu} \tilde{x}_{itu})
\end{aligned}$$

Simplificando e utilizando as definições z_{ij} , z_{aj} , z_{mij} , b_{aj} , c_{aj} e d_{ai} , temos que:

$$\begin{aligned}
A_1 &= A_{11} + A_{12} + A_{13} + A_{14} \\
&= \frac{3}{\lambda^2 \theta^6} \sum_{i,j} z_{aii} z_{mij} z_{ajj} + \frac{6}{\lambda \theta^5} \sum_{i,j} z_{aii} z_{mij} d_{aj} + \frac{3}{\theta^4} \sum_{i,j} d_{ai} z_{mij} d_{aj} + \frac{12}{\lambda^2 \theta^6} \sum_{i,j} z_{aii} z_{aij} z_{mjj} \\
&\quad + \frac{12}{\lambda \theta^5} \sum_{i,j} d_{ai} z_{aij} z_{mjj} + \frac{12}{\lambda \theta^5} \sum_{i,j} z_{mij} c_{aji} - \frac{6}{\theta^4} \sum_{i,j} z_{mij} b_{aji} + \frac{18}{\lambda^2 \theta^6} \sum_{i,j} z_{mij} z_{aij} z_{aij} \\
&\quad - \frac{24}{\theta^4 \lambda^2} \sum_i z_{mii} z_{aii} - \frac{12}{\theta^3 \lambda} \sum_i z_{mii} d_{ai} \\
&= \frac{3}{\lambda^2 \theta^6} \sum_{i,j} z_{aii} (z_{ij} - z_{aij}) z_{ajj} + \frac{6}{\lambda \theta^5} \sum_{i,j} z_{aii} (z_{ij} - z_{aij}) d_{aj} + \frac{3}{\theta^4} \sum_{i,j} d_{ai} (z_{ij} - z_{aij}) d_{aj} \\
&\quad + \frac{12}{\lambda^2 \theta^6} \sum_{i,j} z_{aii} z_{aij} (z_{jj} - z_{ajj}) + \frac{12}{\lambda \theta^5} \sum_{i,j} d_{ai} z_{aij} (z_{jj} - z_{ajj}) + \frac{12}{\lambda \theta^5} \sum_{i,j} (z_{ij} - z_{aij}) c_{aji} \\
&\quad - \frac{6}{\theta^4} \sum_{i,j} (z_{ij} - z_{aij}) b_{aji} + \frac{18}{\lambda^2 \theta^6} \sum_{i,j} (z_{ij} - z_{aij}) z_{aij} z_{aij} - \frac{24}{\theta^4 \lambda^2} \sum_i (z_{ii} - z_{aii}) z_{aii} \\
&\quad - \frac{12}{\theta^3 \lambda} \sum_i (z_{ii} - z_{aii}) d_{ai} \\
&= \frac{3}{\lambda^2 \theta^6} \sum_{i,j} z_{aii} (z_{ij} - z_{aij}) [z_{ajj} + 2\lambda \theta d_{aj}] + \frac{3}{\theta^4} \sum_{i,j} d_{ai} (z_{ij} - z_{aij}) d_{aj} \\
&\quad + \frac{12}{\lambda^2 \theta^6} \sum_{i,j} [z_{aii} + \lambda \theta d_{ai}] z_{aij} (z_{jj} - z_{ajj}) + \frac{6}{\lambda \theta^5} \sum_{i,j} (z_{ij} - z_{aij}) [2c_{aji} - \lambda \theta b_{aji}] \\
&\quad + \frac{18}{\lambda^2 \theta^6} \sum_{i,j} (z_{ij} - z_{aij}) z_{aij} z_{aij} - \frac{12}{\theta^4 \lambda^2} \sum_i (z_{ii} - z_{aii}) [2z_{aii} + \lambda \theta d_{ai}]
\end{aligned}$$

Para a expressão A_2 , temos que

$$\begin{aligned}
A_2 &= -3 \sum \kappa_{r,s,t} \kappa_{u,v,w} a^{tu} m^{rs} m^{vw} + 6 \sum (\kappa_{rst} + 2\kappa_{r,st}) \kappa_{u,v,w} a^{rs} m^{tu} m^{vw} \\
&\quad - 6 \sum \kappa_{r,s,t} \kappa_{u,v,w} a^{tw} m^{ru} m^{sv} + 3 \sum \kappa_{r,s,t,u} m^{rs} m^{tu}
\end{aligned}$$

Logo,

$$\begin{aligned}
A_2 &= -3 \sum \left\{ \frac{2}{\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \right\} \left\{ \frac{2}{\theta^3 \lambda} \sum_j \tilde{x}_{ju} \tilde{x}_{jv} \tilde{x}_{jw} \right\} a^{tu} m^{rs} m^{vw} + 6 \sum \left\{ -\frac{1}{\lambda \theta^3} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} - \frac{1}{\theta^2} \sum_i \tilde{x}_{irt} \tilde{x}_{is} \right. \\
&\quad \left. + \frac{1}{\theta^2} \sum_i \tilde{x}_{ir} \tilde{x}_{ist} - \frac{1}{\theta^2} \sum_i \tilde{x}_{irs} \tilde{x}_{it} \right\} \left\{ \frac{2}{\theta^3 \lambda} \sum_j \tilde{x}_{ju} \tilde{x}_{jv} \tilde{x}_{jw} \right\} a^{rs} m^{tu} m^{vw} - 6 \sum \left\{ \frac{2}{\theta^3 \lambda} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \right\} \times \\
&\quad \left\{ \frac{2}{\theta^3 \lambda} \sum_j \tilde{x}_{ju} \tilde{x}_{jv} \tilde{x}_{jw} \right\} a^{wv} m^{ru} m^{sv} + 3 \sum \left\{ \frac{6}{\theta^4 \lambda^2} \sum_i \tilde{x}_{ir} \tilde{x}_{is} \tilde{x}_{it} \tilde{x}_{iu} \right\} m^{rs} m^{tu} \\
&= -\frac{12}{\theta^6 \lambda^2} \sum_{i,j} (\sum \tilde{x}_{ir} m^{rs} \tilde{x}_{is}) (\sum \tilde{x}_{it} a^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} m^{vw} \tilde{x}_{jw}) \\
&\quad - \frac{12}{\lambda^2 \theta^6} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{is}) (\sum \tilde{x}_{it} m^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} m^{vw} \tilde{x}_{jw}) - \frac{12}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{is} a^{rs} \tilde{x}_{irt} m^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} m^{vw} \tilde{x}_{jw}) \\
&\quad + \frac{12}{\lambda \theta^5} \sum_{i,j} (\sum \tilde{x}_{ir} a^{rs} \tilde{x}_{ist} m^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} m^{vw} \tilde{x}_{jw}) - \frac{12}{\lambda \theta^5} \sum_{i,j} (\sum a^{rs} \tilde{x}_{irs}) (\sum \tilde{x}_{it} m^{tu} \tilde{x}_{ju}) (\sum \tilde{x}_{jv} m^{vw} \tilde{x}_{jw}) \\
&\quad - \frac{24}{\theta^6 \lambda^2} \sum_{i,j} (\sum \tilde{x}_{ir} m^{ru} \tilde{x}_{ju}) (\sum \tilde{x}_{is} m^{sv} \tilde{x}_{jv}) (\sum \tilde{x}_{it} a^{tw} \tilde{x}_{jw}) + \frac{18}{\theta^4 \lambda^2} \sum_i (\sum \tilde{x}_{ir} m^{rs} \tilde{x}_{is}) (\sum \tilde{x}_{it} m^{tu} \tilde{x}_{iu})
\end{aligned}$$

Simplificando e utilizando as definições z_{ij} , z_{aij} , z_{mij} , b_{aij} , c_{aij} e d_{ai} , temos que:

$$\begin{aligned}
A_2 &= -\frac{12}{\theta^6 \lambda^2} \sum_{i,j} z_{mii} z_{aij} z_{mjj} - \frac{12}{\lambda^2 \theta^6} \sum_{i,j} z_{aii} z_{mij} z_{mjj} - \frac{12}{\lambda \theta^5} \sum_{i,j} d_{ai} z_{mij} z_{mjj} \\
&\quad - \frac{24}{\theta^6 \lambda^2} \sum_{i,j} z_{mij} z_{mij} z_{aij} + \frac{18}{\theta^4 \lambda^2} \sum_i z_{mii} z_{mii} \\
&= -\frac{12}{\theta^6 \lambda^2} \sum_{i,j} (z_{ii} - z_{aii}) z_{aij} (z_{jj} - z_{ajj}) - \frac{12}{\lambda^2 \theta^6} \sum_{i,j} z_{aii} (z_{ij} - z_{aij}) (z_{jj} - z_{ajj}) \\
&\quad - \frac{12}{\lambda \theta^5} \sum_{i,j} d_{ai} (z_{ij} - z_{aij}) (z_{jj} - z_{ajj}) - \frac{24}{\theta^6 \lambda^2} \sum_{i,j} (z_{ij} - z_{aij})^2 z_{aij} \\
&\quad + \frac{18}{\theta^4 \lambda^2} \sum_i (z_{ii} - z_{aii})^2
\end{aligned}$$

Por último, temos que

$$\begin{aligned}
A_3 &= 3 \sum \kappa_{r,s,t} \kappa_{u,v,w} m^{rs} m^{tu} m^{vw} + 2 \sum \kappa_{r,s,t} \kappa_{u,v,w} m^{ru} m^{sv} m^{tw} \\
&= 3 \sum \left\{ \frac{2}{\lambda \theta^3} \sum_i x_{ir} x_{is} x_{it} \right\} \left\{ \frac{2}{\lambda \theta^3} \sum_j x_{ju} x_{jv} x_{jw} \right\} m^{rs} m^{tu} m^{vw} \\
&\quad + 2 \sum \left\{ \frac{2}{\lambda \theta^3} \sum_i x_{ir} x_{is} x_{it} \right\} \left\{ \frac{2}{\lambda \theta^3} \sum_j x_{ju} x_{jv} x_{jw} \right\} m^{ru} m^{sv} m^{tw} \\
&= \frac{12}{\lambda^2 \theta^6} \sum_{i,j} (\sum x_{ir} m^{rs} x_{is}) (\sum x_{it} m^{tu} x_{ju}) (\sum x_{jv} m^{vw} x_{jw}) \\
&\quad + \frac{8}{\lambda^2 \theta^6} \sum_{i,j} (\sum x_{ir} m^{ru} x_{ju}) (\sum x_{is} m^{sv} x_{jv}) (\sum x_{it} m^{tw} x_{jw})
\end{aligned}$$

Logo,

$$\begin{aligned}
A_3 &= \frac{12}{\lambda^2 \theta^6} \sum_{i,j} z_{mii} z_{mij} z_{mjj} + \frac{8}{\lambda^2 \theta^6} \sum_{i,j} z_{mij} z_{mij} z_{mij} \\
&= \frac{12}{\lambda^2 \theta^6} \sum_{i,j} (z_{ii} - z_{a ii}) (z_{ij} - z_{a ij}) (z_{jj} - z_{a jj}) + \frac{8}{\lambda^2 \theta^6} \sum_{i,j} (z_{ij} - z_{a ij})^3
\end{aligned}$$

D.2 Correção Tipo-Bartlett à Estatística Gradiente

Apresentamos, nesta seção, os cálculos das quantidades (30)–(32) que compõem o fator da correção tipo-Bartlett para a estatística gradiente definido na Subseção 4.5.2 no modelo MLLGG. Seja $A_1^g = A_{11} + A_{12} + A_{13} + A_{14} + A_{15} + A_{16}$, em que

$$\begin{aligned}
A_{11} &= 3 \sum \kappa_{rst} \kappa_{uvw} [m^{rs} a^{vw} (m^{tu} + 2a^{tu}) + a^{rs} m^{tu} a^{vw} + 2m^{ru} a^{sv} a^{tw}], \\
A_{12} &= -12 \sum \kappa_{rs}^{(t)} \kappa_{uv}^{(w)} (\kappa^{tr} \kappa^{su} \kappa^{vw} + a^{tr} a^{su} a^{vw} + \kappa^{tu} \kappa^{vr} \kappa^{sw} + a^{tu} a^{vr} a^{sw}), \\
A_{13} &= -6 \sum \kappa_{rst} \kappa_{uv}^{(w)} \left[(a^{tw} - \kappa^{tw}) (\kappa^{ru} \kappa^{vs} - a^{ru} a^{vs}) + m^{rs} (a^{tu} a^{vw} + \kappa^{tu} \kappa^{vw}) \right. \\
&\quad \left. + 2a^{st} (\kappa^{ru} \kappa^{vw} - a^{ru} a^{vw}) + 2a^{su} a^{vt} m^{rw} \right], \\
A_{14} &= 6 \sum \kappa_{rstw} m^{rs} a^{tw}, \\
A_{15} &= -6 \sum \kappa_{rst}^{(w)} [m^{rs} (a^{tw} - \kappa^{tw}) + 2m^{rw} a^{st}] \\
&\quad \text{e} \\
A_{16} &= 12 \sum \kappa_{st}^{(rw)} (\kappa^{rs} \kappa^{tw} - a^{rs} a^{tw}).
\end{aligned}$$

Uma vez que no modelo linear log-gama generalizado $\kappa_{rs}^{(t)} = \kappa_{rst}^{(w)} = \kappa_{st}^{(rw)} = 0$, temos que $A_{12} = A_{13} = A_{15} = A_{16} = 0$. Conseqüentemente, a expressão do A_1^g é reduzida à $A_1^g = A_{11} + A_{14}$,

sendo calculado como segue:

$$\begin{aligned}
A_1^g &= 3 \sum \left\{ \frac{1}{\lambda^2 \theta^6} \sum_{i,j} x_{ir} x_{is} x_{it} x_{ju} x_{jv} x_{jw} \right\} [m^{rs} a^{vw} (m^{tu} + 2a^{tu}) + a^{rs} m^{tu} a^{vw} + 2m^{ru} a^{sv} a^{tw}] \\
&\quad + 6 \sum \left\{ -\frac{1}{\theta^4 \lambda^2} \sum_i x_{ir} x_{is} x_{it} x_{iw} \right\} m^{rs} a^{tw} \\
&= \frac{3}{\lambda^2 \theta^6} \sum_{i,j} (\sum x_{ir} m^{rs} x_{is}) (\sum x_{it} m^{tu} x_{ju}) (\sum x_{jv} a^{vw} x_{jw}) \\
&\quad + \frac{6}{\lambda^2 \theta^6} \sum_{i,j} (\sum x_{ir} m^{rs} x_{is}) (\sum x_{it} a^{tu} x_{ju}) (\sum x_{jv} a^{vw} x_{jw}) \\
&\quad + \frac{3}{\lambda^2 \theta^6} \sum_{i,j} (\sum x_{ir} a^{rs} x_{is}) (\sum x_{it} m^{tu} x_{ju}) (\sum x_{jv} a^{vw} x_{jw}) \\
&\quad + \frac{6}{\lambda^2 \theta^6} \sum_{i,j} (\sum x_{ir} m^{ru} x_{ju}) (\sum x_{is} a^{sv} x_{jv}) (\sum x_{it} a^{tw} x_{jw}) \\
&\quad - \frac{6}{\theta^4 \lambda^2} \sum_i (\sum x_{ir} m^{rs} x_{is}) (\sum x_{it} a^{tw} x_{iw}).
\end{aligned}$$

Ou seja,

$$\begin{aligned}
A_1^g &= \frac{3}{\lambda^2 \theta^6} \sum_{i,j} z_{mii} z_{mij} z_{ajj} + \frac{6}{\lambda^2 \theta^6} \sum_{i,j} z_{mii} z_{aij} z_{ajj} + \frac{3}{\lambda^2 \theta^6} \sum_{i,j} z_{aii} z_{mij} z_{ajj} + \frac{6}{\lambda^2 \theta^6} \sum_{i,j} z_{mij} z_{aij} z_{ajj} \\
&\quad - \frac{6}{\theta^4 \lambda^2} \sum_i z_{mii} z_{aii} \\
&= \frac{3}{\lambda^2 \theta^6} \sum_{i,j} z_{mii} (z_{mij} + 2z_{aij}) z_{ajj} + \frac{3}{\lambda^2 \theta^6} \sum_{i,j} z_{aii} z_{mij} z_{ajj} + \frac{6}{\lambda^2 \theta^6} \sum_{i,j} z_{mij} z_{aij} z_{ajj} \\
&\quad - \frac{6}{\theta^4 \lambda^2} \sum_i z_{mii} z_{aii} \\
&= \frac{3}{\lambda^2 \theta^6} \sum_{i,j} (z_{ii} - z_{aii}) (z_{ij} - z_{aij} + 2z_{ajj}) z_{ajj} + \frac{3}{\lambda^2 \theta^6} \sum_{i,j} z_{aii} (z_{ij} - z_{aij}) z_{ajj} \\
&\quad + \frac{6}{\lambda^2 \theta^6} \sum_{i,j} (z_{ij} - z_{aij}) z_{aij} z_{ajj} - \frac{6}{\theta^4 \lambda^2} \sum_i (z_{ii} - z_{aii}) z_{aii} \\
&= \frac{3}{\lambda^2 \theta^6} \sum_{i,j} \{ z_{ii} z_{ij} z_{ajj} + z_{ii} z_{aij} z_{ajj} - 2z_{aii} z_{aij} z_{ajj} \} + \frac{6}{\lambda^2 \theta^6} \sum_{i,j} (z_{ij} - z_{aij}) z_{aij}^2 \\
&\quad - \frac{6}{\theta^4 \lambda^2} \sum_i (z_{ii} - z_{aii}) z_{aii}.
\end{aligned}$$

O A_2^g também pode ser expresso como uma soma de quatro termos, ou seja, $A_2^g = A_{21} + A_{22} + A_{23} + A_{24}$, em que:

$$\begin{aligned}
 A_{21} &= -3 \sum \kappa_{rst} \kappa_{uvw} \left[m^{rs} m^{tu} a^{vw} + m^{rs} a^{tu} m^{vw} + 2m^{ru} m^{sv} a^{tw} \right. \\
 &\quad \left. + \frac{1}{4} (3m^{rs} m^{tu} m^{vw} + 2m^{ru} m^{sv} m^{tw}) \right], \\
 A_{22} &= 6 \sum \kappa_{rst} \kappa_{uv}^{(w)} \left[m^{tw} (\kappa^{ru} \kappa^{vs} - a^{ru} a^{vs}) + m^{rs} (\kappa^{tu} \kappa^{vw} - a^{tu} a^{vw}) \right], \\
 A_{23} &= 6 \sum \kappa_{rst}^{(w)} m^{rs} m^{tw} \\
 &\quad e \\
 A_{24} &= -3 \sum \kappa_{rstw} m^{rs} m^{tw}
 \end{aligned}$$

No modelo linear log-gama generalizado os termos A_{22} e A_{23} também são nulos, fazendo com que expressão do A_2^g se reduza à $A_2^g = A_{21} + A_{24}$, sendo calculado como segue:

$$\begin{aligned}
 A_2^g &= -3 \sum \left\{ \frac{1}{\lambda^2 \theta^6} \sum_{i,j} x_{ir} x_{is} x_{it} x_{ju} x_{jv} x_{jw} \right\} \left[m^{rs} m^{tu} a^{vw} + m^{rs} a^{tu} m^{vw} + 2m^{ru} m^{sv} a^{tw} \right. \\
 &\quad \left. + \frac{1}{4} (3m^{rs} m^{tu} m^{vw} + 2m^{ru} m^{sv} m^{tw}) \right] - 3 \sum \left\{ -\frac{1}{\lambda^2 \theta^4} \sum_i x_{ir} x_{is} x_{it} x_{iw} \right\} m^{rs} m^{tw} \\
 &= -\frac{3}{\lambda^2 \theta^6} \sum_{i,j} (\sum x_{ir} m^{rs} x_{is}) (\sum x_{it} m^{tu} x_{ju}) (\sum x_{jv} a^{vw} x_{jw}) \\
 &\quad - \frac{3}{\lambda^2 \theta^6} \sum_{i,j} (\sum x_{ir} m^{rs} x_{is}) (\sum x_{it} a^{tu} x_{ju}) (\sum x_{jv} m^{vw} x_{jw}) \\
 &\quad - \frac{6}{\lambda^2 \theta^6} \sum_{i,j} (\sum x_{ir} m^{ru} x_{ju}) (\sum x_{is} m^{sv} x_{jv}) (\sum x_{it} a^{tw} x_{jw}) \\
 &\quad - \frac{9}{4\lambda^2 \theta^6} \sum_{i,j} (\sum x_{ir} m^{rs} x_{is}) (\sum x_{it} m^{tu} x_{ju}) (\sum x_{jv} m^{vw} x_{jw}) \\
 &\quad - \frac{6}{4\lambda^2 \theta^6} \sum_{i,j} (\sum x_{ir} m^{ru} x_{ju}) (\sum x_{is} m^{sv} x_{jv}) (\sum x_{it} m^{tw} x_{jw}) \\
 &\quad + \frac{3}{\theta^4 \lambda^2} \sum_i (\sum x_{ir} m^{rs} x_{is}) (\sum x_{it} m^{tw} x_{iw}).
 \end{aligned}$$

Simplificando,

$$\begin{aligned}
A_2^g &= -\frac{3}{\lambda^2\theta^6} \sum_{i,j} z_{mii}z_{mij}z_{ajj} - \frac{3}{\lambda^2\theta^6} \sum_{i,j} z_{mii}z_{aij}z_{mjj} - \frac{6}{\lambda^2\theta^6} \sum_{i,j} z_{mij}z_{mij}z_{aij} \\
&\quad - \frac{9}{4\lambda^2\theta^6} \sum_{i,j} z_{mii}z_{mij}z_{mjj} - \frac{6}{4\lambda^2\theta^6} \sum_{i,j} z_{mij}^3 + \frac{3}{\theta^4\lambda^2} \sum_i z_{mii}^2 \\
&= -\frac{3}{\lambda^2\theta^6} \sum_{i,j} (z_{ii} - z_{a ii}) (z_{ij} - z_{aij}) z_{ajj} - \frac{3}{\lambda^2\theta^6} \sum_{i,j} (z_{ii} - z_{a ii}) z_{aij} (z_{jj} - z_{ajj}) \\
&\quad - \frac{6}{\lambda^2\theta^6} \sum_{i,j} (z_{ij} - z_{aij}) (z_{ij} - z_{aij}) z_{aij} - \frac{9}{4\lambda^2\theta^6} \sum_{i,j} (z_{ii} - z_{a ii}) (z_{ij} - z_{aij}) (z_{jj} - z_{ajj}) \\
&\quad - \frac{6}{4\lambda^2\theta^6} \sum_{i,j} (z_{ij} - z_{aij})^3 + \frac{3}{\theta^4\lambda^2} \sum_i (z_{ii} - z_{a ii})^2 \\
&= -\frac{3}{4\lambda^2\theta^6} \sum_{i,j} (z_{ii} - z_{a ii}) (z_{ij} - z_{aij}) (3z_{jj} + z_{ajj}) - \frac{3}{\lambda^2\theta^6} \sum_{i,j} (z_{ii} - z_{a ii}) z_{aij} (z_{jj} - z_{ajj}) \\
&\quad - \frac{6}{\lambda^2\theta^6} \sum_{i,j} (z_{ij} - z_{aij})^2 z_{aij} - \frac{6}{4\lambda^2\theta^6} \sum_{i,j} (z_{ij} - z_{aij})^3 + \frac{3}{\theta^4\lambda^2} \sum_i (z_{ii} - z_{a ii})^2.
\end{aligned}$$

Por fim, temos que

$$\begin{aligned}
A_3^g &= \frac{1}{4} \sum \kappa_{rst} \kappa_{uvw} (3m^{rs}m^{tu}m^{vw} + 2m^{ru}m^{sv}m^{tw}) \\
&= \frac{1}{4} \sum \left\{ \frac{1}{\lambda^2\theta^6} \sum_{i,j} x_{ir}x_{is}x_{it}x_{ju}x_{jv}x_{jw} \right\} (3m^{rs}m^{tu}m^{vw} + 2m^{ru}m^{sv}m^{tw}) \\
&= \frac{3}{4\lambda^2\theta^6} \sum_{i,j} (\sum x_{ir}m^{rs}x_{is}) (\sum x_{it}m^{tu}x_{ju}) (\sum x_{jv}m^{vw}x_{jw}) \\
&\quad + \frac{1}{2\lambda^2\theta^6} \sum_{i,j} (\sum x_{ir}m^{ru}x_{ju}) (\sum x_{is}m^{sv}x_{jv}) (\sum x_{it}m^{tw}x_{jw}),
\end{aligned}$$

o qual pode ser reescrito como:

$$\begin{aligned}
A_3^g &= \frac{3}{4\lambda^2\theta^6} \sum_{i,j} z_{mii}z_{mij}z_{mjj} + \frac{1}{2\lambda^2\theta^6} \sum_{i,j} z_{mij}^3 \\
&= \frac{3}{4\lambda^2\theta^6} \sum_{i,j} (z_{ii} - z_{a ii}) (z_{ij} - z_{aij}) (z_{jj} - z_{ajj}) + \frac{1}{2\lambda^2\theta^6} \sum_{i,j} (z_{ij} - z_{aij})^3.
\end{aligned}$$

Apêndice E

Influência Global

Suponha que a função de log-verossimilhança para o parâmetro β seja agora expresso na forma:

$$l_{\delta}(\beta) = \sum_{i=1}^n \delta_i l(\beta; y_i), \quad (44)$$

em que $l(\beta; y_i)$ denota a função de log-verossimilhança correspondente à i -ésima observação e δ_i é um tipo de perturbação, tal que $0 \leq \delta_i \leq 1$. Quando $\delta_i = 1, \forall i$, significa que não há perturbação no modelo e quando $\delta_i = 0$ significa que a i -ésima observação foi excluída.

A estimativa de β , supondo a estrutura (44), é dada por $\hat{\beta}_{\delta} = (\tilde{X}^{\top} \Delta \tilde{X})^{-1} \tilde{X}^{\top} \Delta \hat{\delta}_1$, em que $\hat{\delta}_1 = \tilde{X} \hat{\beta} - \lambda \theta \hat{M}_1 \mathbf{1}$ e Δ é uma matriz diagonal de uns com δ na i -ésima posição. Ou seja, $\Delta = \text{diag}\{1, \dots, \delta, \dots, 1\}$. Assim,

$$\begin{aligned} \tilde{X}^{\top} \Delta \tilde{X} &= \tilde{X}^{\top} [I - (I - \Delta)] \tilde{X} \\ &= \tilde{X}^{\top} \tilde{X} - \tilde{X}^{\top} (I - \Delta) \tilde{X} \\ &= \tilde{X}^{\top} \tilde{X} - (1 - \delta) \tilde{x}_i \tilde{x}_i^{\top}, \end{aligned}$$

em que \tilde{x}_i^{\top} é a i -ésima linha da matriz \tilde{X} .

Analogamente, temos que $\tilde{X}^{\top} \Delta \hat{\delta}_1 = \tilde{X}^{\top} \hat{\delta}_1 - (1 - \delta) \tilde{x}_i \hat{\delta}_{1i}$. Para calcular $(\tilde{X}^{\top} \Delta \tilde{X})^{-1}$, considere o seguinte resultado de álgebra linear:

$$(\mathbf{A} + \mathbf{U}\mathbf{V}^{\top})^{-1} = \mathbf{A}^{-1} - \frac{(\mathbf{A}^{-1}\mathbf{U})(\mathbf{V}^{\top}\mathbf{A}^{-1})}{1 + \mathbf{V}^{\top}\mathbf{A}^{-1}\mathbf{U}},$$

com $\mathbf{A} = \tilde{X}^{\top} \tilde{X}$, $\mathbf{U} = -(1 - \delta) \tilde{x}_i$ e $\mathbf{V}^{\top} = \tilde{x}_i^{\top}$. Logo,

$$\begin{aligned} (\tilde{X}^{\top} \Delta \tilde{X})^{-1} &= (\tilde{X}^{\top} \tilde{X})^{-1} + \frac{(\tilde{X}^{\top} \tilde{X})^{-1} (1 - \delta) \tilde{x}_i \tilde{x}_i^{\top} (\tilde{X}^{\top} \tilde{X})^{-1}}{1 - \tilde{x}_i^{\top} (\tilde{X}^{\top} \tilde{X})^{-1} (1 - \delta) \tilde{x}_i} \\ &= (\tilde{X}^{\top} \tilde{X})^{-1} + \frac{(1 - \delta) (\tilde{X}^{\top} \tilde{X})^{-1} \tilde{x}_i \tilde{x}_i^{\top} (\tilde{X}^{\top} \tilde{X})^{-1}}{1 - (1 - \delta) h_{ii}} \end{aligned}$$

em que $h_{ii} = \tilde{x}_i^\top (\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i$. Portanto,

$$\begin{aligned}
\hat{\beta}_\delta &= (\tilde{X}^\top \Delta \tilde{X})^{-1} \tilde{X}^\top \Delta \hat{\delta}_1 \\
&= \left\{ (\tilde{X}^\top \tilde{X})^{-1} + \frac{(1-\delta)(\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \tilde{x}_i^\top (\tilde{X}^\top \tilde{X})^{-1}}{1 - (1-\delta)h_{ii}} \right\} \left\{ \tilde{X}^\top \hat{\delta}_1 - (1-\delta)\tilde{x}_i \hat{\delta}_{1i} \right\} \\
&= (\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top \hat{\delta}_1 - (1-\delta)(\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \hat{\delta}_{1i} + \frac{(1-\delta)(\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \tilde{x}_i^\top (\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top \hat{\delta}_1}{1 - (1-\delta)h_{ii}} \\
&\quad - \frac{(1-\delta)^2 (\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \tilde{x}_i^\top (\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \hat{\delta}_{1i}}{1 - (1-\delta)h_{ii}} \\
&= (\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top \hat{\delta}_1 - (1-\delta)(\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \hat{\delta}_{1i} + \frac{(1-\delta)(\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \tilde{x}_i^\top \hat{\beta}}{1 - (1-\delta)h_{ii}} \\
&\quad - \frac{(1-\delta)^2 (\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i h_{ii} \hat{\delta}_{1i}}{1 - (1-\delta)h_{ii}} \\
&= \hat{\beta} - (1-\delta)(\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \hat{\delta}_{1i} \left(1 + \frac{(1-\delta)h_{ii}}{1 - (1-\delta)h_{ii}} \right) + \frac{(1-\delta)(\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \tilde{x}_i^\top \hat{\beta}}{1 - (1-\delta)h_{ii}} \\
&= \hat{\beta} - (1-\delta)(\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \hat{\delta}_{1i} \left(\frac{1}{1 - (1-\delta)h_{ii}} \right) + \frac{(1-\delta)(\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \tilde{x}_i^\top \hat{\beta}}{1 - (1-\delta)h_{ii}} \\
&= \hat{\beta} - (1-\delta)(\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \left(\frac{\hat{\delta}_{1i} - \tilde{x}_i^\top \hat{\beta}}{1 - (1-\delta)h_{ii}} \right) \\
&= \hat{\beta} - (1-\delta)(\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \left(\frac{r_i^*}{1 - (1-\delta)h_{ii}} \right),
\end{aligned}$$

em que $r_i^* = \hat{\delta}_{1i} - \tilde{x}_i^\top \hat{\beta}$.

No caso em que $\delta = 0$ significa que o i -ésimo ponto foi excluído, ou seja, $\beta_{(i)}$ fica dado por

$$\hat{\beta}_{(i)} = \hat{\beta} - (\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_i \left(\frac{r_i^*}{1 - h_{ii}} \right),$$

em que h_{ii} corresponde ao i -ésimo elemento da diagonal principal da matriz $H = \tilde{X}(\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top$.

Apêndice F

Material Suplementar

F.1 Fator de Correção de Bartlett em MNLLGG com θ não fixo

Nesta seção, apresentamos as expressões gerais para as quantidades necessárias ao cálculo do fator da correção de Bartlett da estatística da razão de verossimilhanças, definidas no Capítulo 4, no MNLLGG no caso em que o parâmetro escala θ é considerado não fixo. Seja Σ o somatório sobre o vetor ξ . Adotamos as letras maiúsculas R, S, T, U, V, W como indexadores do espaço paramétrico do vetor ξ e as letras minúsculas r, s, t, u, v, w como indexadores do espaço paramétrico do vetor β . Considere também as seguintes quantidades, para $i, j = 1, \dots, n$: $z_{ij} = -\sum \tilde{x}_{ir} \kappa^{rs} \tilde{x}_{js}$, $d_i = -\sum \tilde{x}_{itu} \kappa^{tu}$, $c_{ij} = \sum \tilde{x}_{ir} \kappa^{rs} \tilde{x}_{jsu} \kappa^{tu} \tilde{x}_{it}$, $b_{ij} = \sum \tilde{x}_{irt} \kappa^{rs} \tilde{x}_{jsu} \kappa^{tu}$, $f_{1i} = -\sum \tilde{x}_{is} \kappa^{\theta s}$, $f_{2ij} = \sum \kappa^{\theta s} \tilde{x}_{ist} \kappa^{tu} \tilde{x}_{ju}$, $f_{3i} = \sum \kappa^{\theta s} \tilde{x}_{istu} \kappa^{tu}$ e $f_{4i} = \sum \kappa^{\theta s} \tilde{x}_{isu} \kappa^{\theta u}$.

A expressão para $\sum \lambda_{RSTU}$ será dada por:

$$\begin{aligned} \sum \lambda_{RSTU} &= -\frac{1}{4\theta^4 \lambda^2} \sum_i z_{ii}^2 - \frac{1}{2\theta^3 \lambda} \sum_i z_{ii} d_i - \frac{1}{\theta^3 \lambda} \sum_i c_{ii} + \frac{1}{2\theta^2} \sum_i b_{ii} - \frac{1}{4\theta^2} \sum_i d_i^2 - \frac{1}{\theta^4 \lambda^3} \sum_i f_{1i} z_{ii} \\ &\quad - \frac{2}{\theta^3} \left(\frac{1}{\lambda^2} + 2 \right) \sum_i f_{2ii} - \frac{1}{\theta^3 \lambda^2} \sum_i f_{1i} d_i - \frac{1}{4\lambda \theta^2} \sum_i f_{3i} - \frac{1}{2\theta^4} \left(\psi^{(1)}(\lambda^2) - 6 \right) \kappa^{\theta\theta} \sum_i z_{ii} \\ &\quad - \frac{1}{2\theta^3} \lambda \psi^{(1)}(\lambda^2) \kappa^{\theta\theta} \sum_i d_i - \frac{1}{\theta^4} \psi^{(1)}(\lambda^2) \sum_i f_{1i}^2 - \frac{1}{\theta^3} \left(\frac{4}{\lambda} + \lambda \psi^{(1)}(\lambda^2) \right) \sum_i f_{4i} \\ &\quad + \frac{1}{4\theta^4 \lambda} \left[24 + (3\lambda^2 - 12) \psi^{(1)}(\lambda^2) - 4\lambda^2 \psi^{(2)}(\lambda^2) \right] \kappa^{\theta\theta} \sum_i f_{1i} \\ &\quad + \frac{n}{4\theta^4} \left[12\lambda^2 \psi^{(1)}(\lambda^2) - 4\psi^{(2)}(\lambda^2) - \lambda^2 \psi^{(3)}(\lambda^2) + 6 \right] \kappa^{\theta\theta} \kappa^{\theta\theta}. \end{aligned}$$

Já a expressão $\sum \lambda_{RSTUVW}$ pode ser escrito como

$$\begin{aligned}
& \sum \lambda_{RSTUVW} \\
&= \frac{1}{6\lambda^2\theta^6} \sum_{i,j} z_{ij}^3 + \frac{1}{\lambda\theta^5} \sum_{i,j} z_{ij}c_{ij} - \frac{1}{2\theta^4} \sum_{i,j} z_{ij}b_{ji} + \frac{1}{4\lambda^2\theta^6} \sum_{i,j} z_{ij}z_{ii}z_{jj} + \frac{1}{2\lambda\theta^5} \sum_{i,j} z_{ij}d_{iz_{jj}} + \frac{1}{4\theta^4} \sum_{i,j} z_{ij}d_id_j \\
&+ \frac{1}{\lambda^3\theta^6} \sum_{i,j} f_{1j}z_{ij}^2 + \frac{1}{2\theta^5} \left(\frac{5}{\lambda^2} + 8 \right) \sum_{i,j} f_{2ji}z_{ij} + \frac{1}{\lambda^2\theta^5} \sum_{i,j} f_{1j}c_{ji} + \frac{1}{\lambda^2\theta^5} \sum_{i,j} f_{1j}c_{ij} - \frac{1}{\lambda\theta^4} \sum_{i,j} f_{1j}b_{ij} \\
&+ \frac{1}{2\lambda\theta^6} \left(\frac{1}{\lambda^2} + 2 \right) \sum_{i,j} f_{1j}z_{ii}z_{jj} + \frac{1}{2\lambda^2\theta^5} \sum_{i,j} f_{1j}d_{iz_{jj}} + \frac{1}{2\theta^5} \left(\frac{1}{\lambda^2} + 2 \right) \sum_{i,j} f_{1j}z_{ii}d_j + \frac{1}{2\lambda\theta^4} \sum_{i,j} f_{1j}d_id_j \\
&+ \frac{1}{\lambda\theta^6} \left(\frac{1}{\lambda^2} - 1 \right) \sum_{i,j} z_{ij}f_{1i}z_{jj} + \frac{1}{\lambda^2\theta^5} \sum_{i,j} z_{ij}f_{1i}d_j + \frac{1}{2\theta^6} \left(\frac{1}{\lambda^4} - 4 \right) \kappa^{\theta\theta} \sum_{i,j} z_{ij}^2 + \frac{1}{\lambda^3\theta^5} \kappa^{\theta\theta} \sum_{i,j} c_{ji} \\
&- \frac{1}{2\lambda^2\theta^4} \kappa^{\theta\theta} \sum_{i,j} b_{ij} + \frac{1}{4\theta^6} \left(\frac{1}{\lambda^2} + 2 \right)^2 \kappa^{\theta\theta} \sum_{i,j} z_{ii}z_{jj} + \frac{1}{2\lambda\theta^5} \left(\frac{1}{\lambda^2} + 2 \right) \kappa^{\theta\theta} \sum_{i,j} d_{iz_{jj}} + \frac{1}{4\lambda^2\theta^4} \kappa^{\theta\theta} \sum_{i,j} d_id_j \\
&+ \frac{1}{4\theta^6} \psi^{(1)}(\lambda^2) \kappa^{\theta\theta} \sum_{i,j} z_{ij}z_{jj} + \frac{1}{2\theta^5} \lambda \psi^{(1)}(\lambda^2) \kappa^{\theta\theta} \sum_{i,j} z_{ij}d_j + \frac{1}{\theta^6} \psi^{(1)}(\lambda^2) \sum_{i,j} f_{1j}^2 z_{ij} \\
&+ \frac{2}{\theta^5} \left(\frac{2}{\lambda} + \lambda \psi^{(1)}(\lambda^2) \right) \sum_{i,j} f_{2ji}f_{1j} + \frac{1}{\theta^5} \left(\frac{4}{\lambda} + \lambda \psi^{(1)}(\lambda^2) \right) \sum_{i,j} f_{4j}z_{ij} + \frac{1}{\lambda\theta^6} \left(\frac{2}{\lambda} + \lambda \psi^{(1)}(\lambda^2) \right) \sum_{i,j} f_{1j}f_{1i}z_{jj} \\
&+ \frac{1}{\theta^5} \left(\frac{2}{\lambda} + \lambda \psi^{(1)}(\lambda^2) \right) \sum_{i,j} f_{1j}f_{1i}d_j + \frac{2}{\lambda^4\theta^6} \sum_{i,j} f_{1j}f_{1i}z_{ij} + \frac{1}{\lambda\theta^5} \left(\frac{2}{\lambda^2} + 3 \right) \sum_{i,j} f_{1j}f_{2ij} \\
&+ \frac{1}{\lambda^2\theta^6} \left(\frac{1}{\lambda^2} + 2 \right) \sum_{i,j} f_{1j}^2 z_{ii} + \frac{1}{\lambda^3\theta^5} \sum_{i,j} f_{1j}^2 d_i + \frac{1}{\lambda\theta^6} \left(3\psi^{(1)}(\lambda^2) - 8 \right) \kappa^{\theta\theta} \sum_{i,j} f_{1j}z_{ij} \\
&+ \frac{1}{\lambda\theta^5} \left(\frac{4}{\lambda} + 2\lambda \psi^{(1)}(\lambda^2) \right) \kappa^{\theta\theta} \sum_{i,j} f_{2ji} + \frac{1}{2\theta^6} \left(\frac{1}{\lambda^2} + 2 \right) \left(\frac{4}{\lambda} + 3\lambda \psi^{(1)}(\lambda^2) \right) \kappa^{\theta\theta} \sum_{i,j} f_{1i}z_{jj} \\
&+ \frac{1}{2\lambda\theta^5} \left(\frac{4}{\lambda} + 3\lambda \psi^{(1)}(\lambda^2) \right) \kappa^{\theta\theta} \sum_{i,j} f_{1i}d_j + \frac{n}{2\lambda\theta^6} \left\{ (2\lambda^2 + 3)\psi^{(1)}(\lambda^2) + \lambda^2\psi^{(2)}(\lambda^2) \right\} \kappa^{\theta\theta} \sum_{i,j} f_{1j}z_{jj} \\
&+ \frac{n}{2\theta^5} \left\{ (2\lambda^2 + 3)\psi^{(1)}(\lambda^2) + \lambda^2\psi^{(2)}(\lambda^2) \right\} \kappa^{\theta\theta} \sum_{i,j} f_{1j}d_j + \frac{n}{3\lambda\theta^6} \left\{ 3\psi^{(1)}(\lambda^2) + \lambda^2\psi^{(2)}(\lambda^2) - 2 \right\} \sum_i f_{1i}^3 \\
&+ \frac{n}{\theta^5} \left\{ (4\lambda^2 + 3)\psi^{(1)}(\lambda^2) + \lambda^2\psi^{(2)}(\lambda^2) + 2 \right\} \sum_i f_{4i}f_{1i} + \frac{1}{\lambda\theta^6} \left(\frac{4}{\lambda^2} + 3\psi^{(1)}(\lambda^2) \right) \sum_{i,j} f_{1j}^2 f_{1i} \\
&+ \frac{1}{3\lambda\theta^5} \left(\frac{10}{\lambda} + 3\lambda \psi^{(1)}(\lambda^2) \right) \sum_{i,j} f_{4j}f_{1i} - \frac{1}{4\theta^6} \left\{ \frac{16}{\lambda^2} - 3\lambda^2(\psi^{(1)}(\lambda^2))^2 \right\} \kappa^{\theta\theta} \kappa^{\theta\theta} \sum_{i,j} z_{ij} \\
&+ \frac{n}{2\theta^6} \left(\frac{1}{\lambda^2} + 2 \right) \left\{ (2\lambda^2 + 3)\psi^{(1)}(\lambda^2) + \lambda^2\psi^{(2)}(\lambda^2) \right\} \kappa^{\theta\theta} \kappa^{\theta\theta} \sum_j z_{jj} \\
&+ \frac{n}{2\lambda\theta^5} \left\{ (2\lambda^2 + 3)\psi^{(1)}(\lambda^2) + \lambda^2\psi^{(2)}(\lambda^2) \right\} \kappa^{\theta\theta} \kappa^{\theta\theta} \sum_j d_j \\
&+ \frac{2n}{\theta^6} \left\{ \left(\frac{3}{\lambda^2} - 2\lambda^2 + 1 \right) \psi^{(1)}(\lambda^2) + \psi^{(2)}(\lambda^2) - \left(\frac{1}{\lambda^2} + 2 \right) \right\} \kappa^{\theta\theta} \sum_j f_{1j}^2 \\
&+ \frac{n}{\lambda\theta^5} \left\{ (4\lambda^2 + 3)\psi^{(1)}(\lambda^2) + \lambda^2\psi^{(2)}(\lambda^2) + 2 \right\} \kappa^{\theta\theta} \sum_j f_{4j} - \frac{3}{\theta^6} \left(\lambda^2\psi^{(1)}(\lambda^2) + 2 \right) \psi^{(1)}(\lambda^2) \kappa^{\theta\theta} \sum_{i,j} f_{1i}f_{1j} \\
&+ \frac{n}{2\theta^6} \left\{ 6\lambda \psi^{(1)}(\lambda^2) \left(\lambda^2\psi^{(1)}(\lambda^2) - 2 \right) + \left(5\lambda \psi^{(1)}(\lambda^2) + \frac{4}{\lambda} \right) \left(\lambda^2\psi^{(2)}(\lambda^2) + 3\psi^{(1)}(\lambda^2) \right) - \frac{16}{\lambda} \right\} \kappa^{\theta\theta} \kappa^{\theta\theta} \sum_i f_{1i} \\
&+ \left\{ \frac{n^2}{12\theta^6} \left(3(2\lambda^2 + 1)\psi^{(1)}(\lambda^2) + \lambda^2\psi^{(2)}(\lambda^2) + 4 \right) \left(3(5 - 6\lambda^2)\psi^{(1)}(\lambda^2) + 5\lambda^2\psi^{(2)}(\lambda^2) - 28 \right) \right. \\
&+ \left. \frac{8n^2}{\theta^6} \left\{ 1 + \lambda^2\psi^{(1)}(\lambda^2) \right\}^2 \right\} \kappa^{\theta\theta} \kappa^{\theta\theta} \kappa^{\theta\theta}
\end{aligned}$$

F.2 Fator de Correção tipo-Bartlett à Estatística Gradiente em MNLLGG

Apresentamos, nesta seção, as quantidades (30)–(32) que compõem o fator da correção de tipo-Bartlett para a estatística gradiente definido na Subseção 4.5.2 para o modelo MNLLGG. Para a simplificação dos cálculos, consideremos as seguintes definições:

$$\begin{aligned} z_{ij} &= \sum \tilde{x}_{is} \mathbf{K}^{su} \tilde{x}_{ju} & z_{mij} &= \sum \tilde{x}_{is} m^{su} \tilde{x}_{ju} & z_{aij} &= \sum \tilde{x}_{is} a^{su} \tilde{x}_{ju} \\ d_i &= \sum \mathbf{K}^{rs} \tilde{x}_{irs} & d_{mi} &= \sum m^{rs} \tilde{x}_{irs} & d_{ai} &= \sum a^{rs} \tilde{x}_{irs} \end{aligned}$$

$$\begin{aligned} c_{ij} &= \sum \tilde{x}_{is} \mathbf{K}^{su} \tilde{x}_{iuv} \mathbf{K}^{vw} \tilde{x}_{jv} & l_{ij} &= \sum \tilde{x}_{is} \mathbf{K}^{sv} \tilde{x}_{jvw} \mathbf{K}^{tw} \tilde{x}_{it} \\ c_{mmij} &= \sum \tilde{x}_{is} m^{su} \tilde{x}_{iuv} m^{vw} \tilde{x}_{jv} & l_{mmij} &= \sum \tilde{x}_{is} m^{sv} \tilde{x}_{jvw} m^{tw} \tilde{x}_{it} \\ c_{aaij} &= \sum \tilde{x}_{is} a^{su} \tilde{x}_{iuv} a^{vw} \tilde{x}_{jv} & l_{aaij} &= \sum \tilde{x}_{is} a^{sv} \tilde{x}_{jvw} a^{tw} \tilde{x}_{it} \\ c_{amij} &= \sum \tilde{x}_{is} a^{su} \tilde{x}_{iuv} m^{vw} \tilde{x}_{jv} & l_{maj} &= \sum \tilde{x}_{is} m^{sv} \tilde{x}_{jvw} a^{tw} \tilde{x}_{it} \\ c_{mai} &= \sum \tilde{x}_{is} m^{su} \tilde{x}_{iuv} a^{vw} \tilde{x}_{jv} & l_{\kappa aij} &= \sum \tilde{x}_{is} \mathbf{K}^{sv} \tilde{x}_{jvw} a^{tw} \tilde{x}_{it} \\ c_{mkij} &= \sum \tilde{x}_{is} m^{su} \tilde{x}_{iuv} \mathbf{K}^{vw} \tilde{x}_{jv} & l_{\kappa mij} &= \sum \tilde{x}_{is} \mathbf{K}^{sv} \tilde{x}_{jvw} m^{tw} \tilde{x}_{it} \\ c_{akij} &= \sum \tilde{x}_{is} a^{su} \tilde{x}_{iuv} \mathbf{K}^{vw} \tilde{x}_{jv} \end{aligned}$$

$$\begin{aligned} b_{ij} &= \sum \mathbf{K}^{tw} \tilde{x}_{irt} \mathbf{K}^{ru} \tilde{x}_{juw} & o_{ij} &= \sum \mathbf{K}^{rs} \tilde{x}_{irsw} \mathbf{K}^{tw} \tilde{x}_{jt} \\ b_{mmij} &= \sum m^{tw} \tilde{x}_{irt} m^{ru} \tilde{x}_{juw} & o_{mmij} &= \sum m^{rs} \tilde{x}_{irsw} m^{tw} \tilde{x}_{jt} \\ b_{aaij} &= \sum a^{tw} \tilde{x}_{irt} a^{ru} \tilde{x}_{juw} & o_{aaij} &= \sum a^{rs} \tilde{x}_{irsw} a^{tw} \tilde{x}_{jt} \\ b_{mai} &= \sum m^{tw} \tilde{x}_{irt} a^{ru} \tilde{x}_{juw} & o_{maj} &= \sum m^{rs} \tilde{x}_{irsw} a^{tw} \tilde{x}_{jt} \\ b_{\kappa aij} &= \sum \mathbf{K}^{tw} \tilde{x}_{irt} a^{ru} \tilde{x}_{juw} & o_{\kappa aij} &= \sum \mathbf{K}^{tw} \tilde{x}_{irtw} a^{rs} \tilde{x}_{js} \\ b_{\kappa mij} &= \sum \mathbf{K}^{tw} \tilde{x}_{irt} m^{ru} \tilde{x}_{juw} & o_{\kappa mij} &= \sum \mathbf{K}^{tw} \tilde{x}_{irtw} m^{rs} \tilde{x}_{js} \\ & & o_{akij} &= \sum a^{rs} \tilde{x}_{irsw} \mathbf{K}^{tw} \tilde{x}_{jt} \\ & & o_{mkij} &= \sum m^{rs} \tilde{x}_{irsw} \mathbf{K}^{tw} \tilde{x}_{jt} \\ & & o_{amij} &= \sum a^{tw} \tilde{x}_{irtw} m^{rs} \tilde{x}_{js} \end{aligned}$$

$$\begin{aligned} g_{amij} &= \sum \tilde{x}_{is} a^{rs} \tilde{x}_{irt} m^{tu} \tilde{x}_{juw} m^{vw} \tilde{x}_{jv} & h_{amij} &= \sum \tilde{x}_{is} m^{rs} \tilde{x}_{irt} a^{tu} \tilde{x}_{juw} m^{vw} \tilde{x}_{jv} \\ g_{akij} &= \sum \tilde{x}_{is} a^{rs} \tilde{x}_{irt} \mathbf{K}^{tu} \tilde{x}_{juw} \mathbf{K}^{vw} \tilde{x}_{jv} & h_{akij} &= \sum \tilde{x}_{is} \mathbf{K}^{rs} \tilde{x}_{irt} a^{tu} \tilde{x}_{juw} \mathbf{K}^{vw} \tilde{x}_{jv} \\ g_{maj} &= \sum \tilde{x}_{is} m^{rs} \tilde{x}_{irt} a^{tu} \tilde{x}_{juw} a^{vw} \tilde{x}_{jv} & h_{mai} &= \sum \tilde{x}_{is} a^{rs} \tilde{x}_{irt} m^{tu} \tilde{x}_{juw} a^{vw} \tilde{x}_{jv} \\ g_{mkij} &= \sum \tilde{x}_{is} m^{rs} \tilde{x}_{irt} \mathbf{K}^{tu} \tilde{x}_{juw} \mathbf{K}^{vw} \tilde{x}_{jv} & h_{mkij} &= \sum \tilde{x}_{is} \mathbf{K}^{rs} \tilde{x}_{irt} m^{tu} \tilde{x}_{juw} \mathbf{K}^{vw} \tilde{x}_{jv} \\ g_{\kappa mij} &= \sum \tilde{x}_{is} \mathbf{K}^{rs} \tilde{x}_{irt} m^{tu} \tilde{x}_{juw} m^{vw} \tilde{x}_{jv} & h_{\kappa mij} &= \sum \tilde{x}_{is} m^{rs} \tilde{x}_{irt} \mathbf{K}^{tu} \tilde{x}_{juw} m^{vw} \tilde{x}_{jv} \\ g_{\kappa aij} &= \sum \tilde{x}_{is} \mathbf{K}^{rs} \tilde{x}_{irt} a^{tu} \tilde{x}_{juw} a^{vw} \tilde{x}_{jv} & h_{\kappa aij} &= \sum \tilde{x}_{is} a^{rs} \tilde{x}_{irt} \mathbf{K}^{tu} \tilde{x}_{juw} a^{vw} \tilde{x}_{jv} \end{aligned}$$

$$\begin{aligned} q_{ij} &= \sum \tilde{x}_{is} \mathbf{K}^{sw} \tilde{x}_{juw} \mathbf{K}^{tu} \tilde{x}_{irt} \mathbf{K}^{vr} \tilde{x}_{jv} & f_{ij} &= \sum \tilde{x}_{ir} \mathbf{K}^{tr} \tilde{x}_{ist} \mathbf{K}^{su} \tilde{x}_{juw} \mathbf{K}^{vw} \tilde{x}_{jv} \\ q_{aaij} &= \sum \tilde{x}_{is} a^{sw} \tilde{x}_{juw} a^{tu} \tilde{x}_{irt} a^{vr} \tilde{x}_{jv} & f_{mmij} &= \sum \tilde{x}_{ir} m^{tr} \tilde{x}_{ist} m^{su} \tilde{x}_{juw} m^{vw} \tilde{x}_{jv} \\ q_{mmij} &= \sum \tilde{x}_{is} m^{sw} \tilde{x}_{juw} m^{tu} \tilde{x}_{irt} m^{vr} \tilde{x}_{jv} & f_{aaij} &= \sum \tilde{x}_{ir} a^{tr} \tilde{x}_{ist} a^{su} \tilde{x}_{juw} a^{vw} \tilde{x}_{jv} \end{aligned}$$

$$\begin{aligned} p_{mai} &= \sum \tilde{x}_{ir} m^{ru} \tilde{x}_{juw} a^{tw} \tilde{x}_{ist} a^{sv} \tilde{x}_{jv} & e_{mai} &= \sum \tilde{x}_{jv} a^{sv} \tilde{x}_{irs} m^{ru} \tilde{x}_{juw} a^{tw} \tilde{x}_{it} \\ p_{mij} &= \sum \tilde{x}_{ir} a^{ru} \tilde{x}_{juw} m^{tw} \tilde{x}_{ist} m^{sv} \tilde{x}_{jv} & e_{amij} &= \sum \tilde{x}_{jv} m^{sv} \tilde{x}_{irs} a^{ru} \tilde{x}_{juw} m^{tw} \tilde{x}_{it} \\ p_{mkij} &= \sum \tilde{x}_{ir} m^{ru} \tilde{x}_{juw} \mathbf{K}^{tw} \tilde{x}_{ist} \mathbf{K}^{sv} \tilde{x}_{jv} & e_{mkij} &= \sum \tilde{x}_{jv} \mathbf{K}^{sv} \tilde{x}_{irs} m^{ru} \tilde{x}_{juw} \mathbf{K}^{tw} \tilde{x}_{it} \\ p_{\kappa mij} &= \sum \tilde{x}_{ir} \mathbf{K}^{ru} \tilde{x}_{juw} m^{tw} \tilde{x}_{ist} m^{sv} \tilde{x}_{jv} & e_{\kappa mij} &= \sum \tilde{x}_{jv} m^{sv} \tilde{x}_{irs} \mathbf{K}^{ru} \tilde{x}_{juw} m^{tw} \tilde{x}_{it} \\ p_{\kappa aij} &= \sum \tilde{x}_{ir} \mathbf{K}^{ru} \tilde{x}_{juw} a^{tw} \tilde{x}_{ist} a^{sv} \tilde{x}_{jv} & e_{\kappa aij} &= \sum \tilde{x}_{jv} a^{sv} \tilde{x}_{irs} \mathbf{K}^{ru} \tilde{x}_{juw} a^{tw} \tilde{x}_{it} \\ p_{akij} &= \sum \tilde{x}_{ir} a^{ru} \tilde{x}_{juw} \mathbf{K}^{tw} \tilde{x}_{ist} \mathbf{K}^{sv} \tilde{x}_{jv} & e_{akij} &= \sum \tilde{x}_{jv} \mathbf{K}^{sv} \tilde{x}_{irs} a^{ru} \tilde{x}_{juw} \mathbf{K}^{tw} \tilde{x}_{it} \end{aligned}$$

Assim, as quantidades A_1 , A_2 e A_3 são dadas por

$$\begin{aligned}
A_1 = & \frac{3}{\lambda^2 \theta^6} \sum_{i,j} z_{mii} (z_{mij} + 2z_{aij}) z_{ajj} - \frac{6}{\lambda \theta^5} \sum_{i,j} z_{mii} (c_{amji} + 2c_{aaji}) - \frac{3}{\lambda \theta^5} \sum_{i,j} z_{mii} (z_{mij} + 2z_{aij}) d_{aj} \\
& - \frac{6}{\lambda \theta^5} \sum_{i,j} (c_{mmij} + 2c_{maij}) z_{ajj} - \frac{3}{\lambda \theta^5} \sum_{i,j} d_{mi} (z_{mij} + 2z_{aij}) z_{ajj} + \frac{12}{\theta^4} \sum_{i,j} g_{amji} \\
& + \frac{6}{\theta^4} \sum_{i,j} d_{mi} (c_{amji} + 2c_{aaji}) + \frac{6}{\theta^4} \sum_{i,j} (c_{mmij} + 2c_{maij}) d_{aj} + \frac{3}{\theta^4} \sum_{i,j} d_{mi} (z_{mij} + 2z_{aij}) d_{aj} \\
& + \frac{24}{\theta^4} \sum_{i,j} g_{maj} + \frac{3}{\lambda^2 \theta^6} \sum_{i,j} z_{aii} z_{mij} z_{ajj} - \frac{12}{\lambda \theta^5} \sum_{i,j} z_{aii} c_{amji} - \frac{6}{\lambda \theta^5} \sum_{i,j} z_{aii} z_{mij} d_{aj} + \frac{12}{\theta^4} \sum_{i,j} h_{maj} \\
& + \frac{12}{\theta^4} \sum_{i,j} d_{ai} c_{amji} + \frac{3}{\theta^4} \sum_{i,j} d_{ai} z_{mij} d_{aj} + \frac{6}{\lambda^2 \theta^6} \sum_{i,j} z_{mij} z_{aij} z_{aij} - \frac{24}{\lambda \theta^5} \sum_{i,j} l_{maj} z_{aij} \\
& - \frac{12}{\lambda \theta^5} \sum_{i,j} z_{mij} l_{aaij} + \frac{12}{\theta^4} \sum_{i,j} b_{maj} z_{aij} + \frac{24}{\theta^4} \sum_{i,j} p_{maj} + \frac{12}{\theta^4} \sum_{i,j} e_{maj} + \frac{6}{\theta^4} \sum_{i,j} z_{mij} b_{aaij} \\
& - \frac{24}{\theta^4} \sum_{i,j} d_i c_{ji} - \frac{12}{\theta^4} \sum_{i,j} d_i z_{ij} d_j - \frac{12}{\theta^4} \sum_{i,j} f_{ij} - \frac{24}{\theta^4} \sum_{i,j} d_{ai} c_{aaji} - \frac{12}{\theta^4} \sum_{i,j} d_{ai} z_{aij} d_{aj} - \frac{12}{\theta^4} \sum_{i,j} f_{aaij} \\
& - \frac{36}{\theta^4} \sum_{i,j} q_{ij} - \frac{12}{\theta^4} \sum_{i,j} z_{ij} b_{ij} - \frac{36}{\theta^4} \sum_{i,j} q_{aaij} - \frac{12}{\theta^4} \sum_{i,j} z_{aij} b_{aaij} + \frac{24}{\lambda \theta^5} \sum_{i,j} l_{kaij} (z_{ij} + z_{aij}) \\
& - \frac{24}{\lambda \theta^5} \sum_{i,j} (l_{ij} z_{ij} + l_{aaij} z_{aij} - 2l_{maj} z_{aij}) - \frac{24}{\theta^4} \sum_{i,j} (b_{ij} - b_{kaij}) z_{ij} - \frac{24}{\theta^4} \sum_{i,j} (e_{akij} + e_{kaij} + 2e_{maj}) \\
& + \frac{48}{\theta^4} \sum_{i,j} (q_{ij} + q_{aaij}) - \frac{24}{\theta^4} \sum_{i,j} (p_{kaij} + p_{akij} + 2p_{maj}) - \frac{24}{\theta^4} \sum_{i,j} (b_{kaij} - b_{aaij} + 2b_{maj}) z_{aij} \\
& + \frac{12}{\lambda \theta^5} \sum_{i,j} (z_{mii} - 2z_{aii}) c_{aaji} + \frac{12}{\lambda \theta^5} \sum_{i,j} z_{mii} (z_{aij} d_{aj} + z_{ij} d_j) - \frac{24}{\theta^4} \sum_{i,j} (g_{mkij} + g_{maj} + 2g_{akij}) \\
& - \frac{24}{\theta^4} \sum_{i,j} (c_{maj} + c_{ij} - 3c_{aaij}) d_{aj} - \frac{12}{\theta^4} \sum_{i,j} d_{mi} (c_{aaji} + c_{ji}) - \frac{12}{\theta^4} \sum_{i,j} (d_{mi} - 2d_{ai}) z_{aij} d_{aj} \\
& + \frac{12}{\lambda \theta^5} \sum_{i,j} (z_{mii} + 2z_{aii}) c_{ji} - \frac{24}{\theta^4} \sum_{i,j} (c_{mkij} + 2c_{akij}) d_j - \frac{12}{\theta^4} \sum_{i,j} (d_{mi} + 2d_{ai}) z_{ij} d_j \\
& + \frac{24}{\lambda \theta^5} \sum_{i,j} (z_{ij} d_j - z_{aij} d_{aj}) z_{aii} + \frac{48}{\theta^4} \sum_{i,j} f_{aaij} - \frac{6}{\theta^4 \lambda^2} \sum_i z_{mii} z_{aii} + \frac{6}{\theta^3 \lambda} \sum_i z_{mii} d_{ai} + \frac{24}{\theta^3 \lambda} \sum_i c_{mai} \\
& + \frac{6}{\theta^3 \lambda} \sum_i d_{mi} z_{aii} - \frac{24}{\theta^2} \sum_i o_{amii} - \frac{18}{\theta^2} \sum_i b_{mai} - \frac{36}{\lambda \theta^3} \sum_i c_{mai} + \frac{12}{\lambda \theta^3} \sum_i c_{mkii} \\
& + \frac{6}{\lambda \theta^3} \sum_i z_{mii} (d_i - d_{ai}) + \frac{24}{\theta^2} \sum_i o_{amii} - \frac{12}{\theta^2} \sum_i o_{kmii} + \frac{36}{\theta^2} \sum_i b_{mai} - \frac{12}{\theta^2} \sum_i b_{kmii} + \frac{30}{\theta^2} \sum_i o_{mai} \\
& - \frac{6}{\theta^2} \sum_i o_{mkii} - \frac{6}{\theta^2} \sum_i (d_i - 3d_{ai}) d_{mi} - \frac{12}{\lambda \theta^3} \sum_i d_{mi} z_{aii} - \frac{24}{\theta^2} \sum_i o_{ii} - \frac{12}{\theta^2} \sum_i d_i^2 - \frac{12}{\theta^2} \sum_i b_{ii} \\
& + \frac{24}{\theta^2} \sum_i o_{aaii} + \frac{12}{\theta^2} \sum_i d_{ai}^2 + \frac{12}{\theta^2} \sum_i b_{aaii}
\end{aligned}$$

$$\begin{aligned}
A_2 = & -\frac{3}{\lambda^2\theta^6} \sum_{i,j} z_{mii}z_{mij}z_{ajj} + \frac{6}{\lambda\theta^5} \sum_{i,j} z_{mii}c_{amji} + \frac{3}{\lambda\theta^5} \sum_{i,j} z_{mii}z_{mij}d_{aj} + \frac{6}{\lambda\theta^5} \sum_{i,j} c_{mmij}z_{ajj} \\
& + \frac{3}{\lambda\theta^5} \sum_{i,j} d_{mi}z_{mij}z_{ajj} - \frac{12}{\theta^4} \sum_{i,j} g_{amji} - \frac{6}{\theta^4} \sum_{i,j} d_{mi}c_{amji} - \frac{6}{\theta^4} \sum_{i,j} c_{mmij}d_{aj} \\
& - \frac{3}{\theta^4} \sum_{i,j} d_{mi}z_{mij}d_{aj} - \frac{3}{\lambda^2\theta^6} \sum_{i,j} z_{mii}z_{aij}z_{mjj} + \frac{12}{\lambda\theta^5} \sum_{i,j} z_{mii}c_{maji} + \frac{6}{\lambda\theta^5} \sum_{i,j} z_{mii}z_{aij}d_{mj} \\
& - \frac{12}{\theta^4} \sum_{i,j} h_{amij} - \frac{12}{\theta^4} \sum_{i,j} d_{mi}c_{maji} - \frac{3}{\theta^4} \sum_{i,j} d_{mi}z_{aij}d_{mj} - \frac{6}{\lambda^2\theta^6} \sum_{i,j} z_{mij}z_{mij}z_{aij} \\
& + \frac{24}{\lambda\theta^5} \sum_{i,j} l_{maj}z_{mij} + \frac{12}{\lambda\theta^5} \sum_{i,j} l_{mmij}z_{aij} - \frac{12}{\theta^4} \sum_{i,j} b_{maj}z_{mij} - \frac{12}{\theta^4} \sum_{i,j} e_{amij} - \frac{24}{\theta^4} \sum_{i,j} p_{amij} \\
& - \frac{6}{\theta^4} \sum_{i,j} b_{mmij}z_{aij} - \frac{9}{4\lambda^2\theta^6} \sum_{i,j} z_{mii}z_{mij}z_{mjj} + \frac{9}{\lambda\theta^5} \sum_{i,j} z_{mii}c_{mmji} + \frac{9}{2\lambda\theta^5} \sum_{i,j} z_{mii}z_{mij}d_{mj} \\
& - \frac{9}{\theta^4} \sum_{i,j} f_{mmij} - \frac{9}{\theta^4} \sum_{i,j} d_{mi}c_{mmji} - \frac{9}{4\theta^4} \sum_{i,j} d_{mi}z_{mij}d_{mj} - \frac{6}{4\lambda^2\theta^6} \sum_{i,j} z_{mij}^3 \\
& + \frac{9}{\lambda\theta^5} \sum_{i,j} l_{mmij}z_{mij} - \frac{18}{4\theta^4} \sum_{i,j} b_{mmij}z_{mij} - \frac{9}{\theta^4} \sum_{i,j} q_{mmij} - \frac{12}{\lambda\theta^5} \sum_{i,j} l_{\kappa m ij}z_{ij} \\
& + \frac{12}{\lambda\theta^5} \sum_{i,j} l_{amij}z_{ij} + \frac{12}{\theta^4} \sum_{i,j} b_{\kappa m ij}z_{ij} - \frac{12}{\theta^4} \sum_{i,j} b_{maj}z_{aij} + \frac{12}{\theta^4} \sum_{i,j} e_{\kappa m ij} \\
& - \frac{12}{\theta^4} \sum_{i,j} e_{maj} + \frac{12}{\theta^4} \sum_{i,j} p_{\kappa m ij} - \frac{12}{\theta^4} \sum_{i,j} p_{maj} - \frac{6}{\lambda\theta^5} \sum_{i,j} z_{mii}c_{ji} + \frac{6}{\lambda\theta^5} \sum_{i,j} z_{mii}c_{aaji} \\
& - \frac{6}{\lambda\theta^5} \sum_{i,j} z_{mii}z_{ij}d_j + \frac{6}{\lambda\theta^5} \sum_{i,j} z_{mii}z_{aij}d_{aj} + \frac{12}{\theta^4} \sum_{i,j} g_{\kappa m ij} - \frac{12}{\theta^4} \sum_{i,j} g_{maj} + \frac{12}{\theta^4} \sum_{i,j} c_{\kappa m ij}d_j \\
& - \frac{12}{\theta^4} \sum_{i,j} c_{maj}d_{aj} + \frac{6}{\theta^4} \sum_{i,j} d_{mi}c_{ji} - \frac{6}{\theta^4} \sum_{i,j} d_{mi}c_{aaji} + \frac{6}{\theta^4} \sum_{i,j} d_{mi}z_{ij}d_j - \frac{6}{\theta^4} \sum_{i,j} d_{mi}z_{aij}d_{aj} \\
& + \frac{12}{\lambda\theta^3} \sum_i c_{mmii} + \frac{6}{\lambda\theta^3} \sum_i z_{mii}d_{mi} - \frac{18}{\theta^2} \sum_i o_{mmii} - \frac{12}{\theta^2} \sum_i b_{mmii} - \frac{6}{\theta^2} \sum_i d_{mi}^2 \\
& + \frac{3}{\theta^4\lambda^2} \sum_i z_{mii}^2 - \frac{6}{\theta^3\lambda} \sum_i z_{mii}d_{mi} - \frac{12}{\theta^3\lambda} \sum_i c_{mmii} + \frac{12}{\theta^2} \sum_i o_{mmii} + \frac{9}{\theta^2} \sum_i b_{mmii}
\end{aligned}$$

e

$$\begin{aligned}
A_3 = & \frac{3}{4\lambda^2\theta^6} \sum_{i,j} z_{mii}z_{mij}z_{mjj} - \frac{3}{\lambda\theta^5} \sum_{i,j} z_{mii}c_{mmji} - \frac{3}{2\lambda\theta^5} \sum_{i,j} z_{mii}z_{mij}d_{mj} + \frac{3}{\theta^4} \sum_{i,j} f_{mmij} \\
& + \frac{3}{\theta^4} \sum_{i,j} d_{mi}c_{mmji} + \frac{3}{4\theta^4} \sum_{i,j} d_{mi}z_{mij}d_{mj} + \frac{1}{2\lambda^2\theta^6} \sum_{i,j} z_{mij}^3 - \frac{3}{\lambda\theta^5} \sum_{i,j} l_{mmij}z_{mij} \\
& + \frac{3}{2\theta^4} \sum_{i,j} b_{mmij}z_{mij} + \frac{3}{\theta^4} \sum_{i,j} q_{mmij}
\end{aligned}$$