

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE ESTATÍSTICA

PÓS-GRADUAÇÃO EM ESTATÍSTICA

TESTE DE RAZÃO DE VEROSSIMILHANÇAS BOOTSTRAP E
TÉCNICAS DE DIAGNÓSTICO EM MODELOS EM SÉRIES DE
POTÊNCIAS NÃO-LINEARES GENERALIZADOS

SÉBASTIEN LOZANO FORERO

DISSERTAÇÃO DE MESTRADO



Recife
2015

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE ESTATÍSTICA

SÉBASTIEN LOZANO FORERO

TESTE DE RAZÃO DE VEROSSIMILHANÇAS BOOTSTRAP E
TÉCNICAS DE DIAGNÓSTICO EM MODELOS EM SÉRIES DE
POTÊNCIAS NÃO-LINEARES GENERALIZADOS

ORIENTADORA: PROFA. DRA. AUDREY HELEN MARIZ DE AQUINO CYSNEIROS

Área de Concentração: Estatística Aplicada

Trabalho apresentado ao Programa de Pós-Graduação em Estatística do Departamento de Estatística da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de **Mestre em Estatística**.

Recife
2015

Catálogo na fonte
Bibliotecária Joana D'Arc Leão Salvador CRB4-532

L925t Lozano Forero, Sébastien.
 Teste de razão de verossimilhanças bootstrap e técnicas de diagnóstico em modelos em séries de potências não lineares generalizada / Sébastien Lozano Forero. – Recife: O Autor, 2015.
 96 f.: fig., tab.

 Orientador: Audrey Helen Mariz de Aquino Cysneiros.
 Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Estatística, 2015.
 Inclui referências e apêndices.

 1. Estatística matemática. 2. Análise de regressão. I. Cysneiros, Audrey Helen Mariz de Aquino (Orientadora). II. Título.

519.536 CDD (22. ed.) UFPE-MEI 2015-086

SÉBASTIEN LOZANO FORERO

**TESTE DE RAZÃO DE VEROSSIMILHANÇAS BOOTSTRAP E TÉCNICAS DE
DIAGNÓSTICO EM MODELOS EM SÉRIES DE POTÊNCIAS NÃO-LINEARES
GENERALIZADOS**

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

Aprovada em: 25 de junho de 2015.

BANCA EXAMINADORA

Prof.^a Dra. Audrey Helen Mariz de Aquino Cysneiros
UFPE

Prof. PhD. Getúlio José Amorim do Amaral (Examinador Interno)
UFPE

Prof.^a Dra. Dione Maria Valença (Examinador Externo)
UFRN

*Dedicado a mi mamá,
Nolba Forero.*

Dedicado a mi familia.

*Dedicado a la memoria de
Alicia Bello de Lozano*

Agradecimientos

Estudiar en Brasil fue una oportunidad increíble, además de única. Más que el inconmensurable crecimiento profesional, no puedo dejar de citar el gran crecimiento personal que tuve al haber vivido por casi dos años en este colosal país. Son muchas las personas e instituciones a las que debo agradecer por haber podido culminar con éxito esta etapa. Las más destacadas:

- A mis padres, Nolba Forero y Hector Lozano, por la vida.
- A mi mamá, Nolba Forero, por tantos y tantos años de amor y soporte incondicional. Sus esmeros y esfuerzos comienzan a rendir frutos. Hoy soy quién soy, gracias y sólo gracias a ella.
- A Julián Quintero que, mismo sin llevar su sangre por mis venas, siempre me trató y me trata como a un hijo.
- A mi familia, por el inmenso cariño, la convivencia y por las lecciones que día a día me dan sobre la vida.
- A Dorita, por su compañía, apoyo moral y por impregnarme con su emoción y amor por la vida y sus curiosidades.
- A William Rodríguez, por haber sido un gran amigo, compañero y colega durante esta maestría.
- A la Sra. Dora Villagrán e hijos, Campo E., Martha, Claudia y Beto por tratarme, desde hace varios años, como un miembro más de la familia.

- A la profesora Audrey H.M. de A. Cysneiros, por aceptar dirigir este trabajo, por confiar en mi capacidades y por introducirme en el mundo académico. Fue una experiencia muy constructiva trabajar con una persona tan instruida, pedagógica y amigable.
- A Priscila Gonçalves por haber avanzado tanto los estudios de perfeccionamiento de pruebas de hipótesis en los modelos de regresión que se estudian en esta disertación.
- A los profesores del Departamento de Estadística de la Universidade Federal de Pernambuco, Gauss Cordeiro, Francisco Cribari, Renato Cintra, Francisco Cysneiros, Leandro Chaves, Getúlio Amoral, Raydonal Opsina, entre otros, por dar esa personalidad distintiva al departamento, además de la formación recibida a lo largo de la maestría.
- A Valéria Bittencourt por su ayuda, por su agilidad y sobre todo, por su cariño con los estudiantes. Por ser esa persona diligente y cariñosa que uno quisiera encontrar en todas partes.
- A los profesores Getúlio José Amorim do Amaral y Dione Maria Valença por haber aceptado ser jueces para este trabajo y por las sugerencias dadas.
- A China por ser nuestro amigo, y casi padre, brasileño incondicional. Por enseñarnos portugués y cuidarnos tanto.
- A Alejandro, Olivia, Stephania y Alejandra por su compañía y amistad en la parte final de la maestría.
- A Lucecita y Jimmy por su amistad valiosa al comienzo de la maestría.
- A Raphael, Elvio, Saulinho, Diego, macaco, perigo, gago, Sra. Avani, Paula, Johannes, Christian y tantos otros grandes amigos que tuve el gusto de conocer en Brasil.
- A Laura, Telma, Rodrigo, Raphaela, Luana, Alisson, Jenny, Diego, Yury, Vinicius, Jose-lino y los demás compañeros de estudio por la compañía y la convivencia.
- A mis amigos y compañeros de estudio en Colombia por el apoyo y la amistad.
- A la Universidad Nacional de Colombia, sede Bogotá, por la formación.
- A FACEPE por el apoyo financiero y permitirme estudiar sin las odiosas preocupaciones financieras.

Os Modelos em Série de Potência não-Lineares Generalizados (MSPNLGs), introduzidos em [Cordeiro et al. \(2009\)](#), são apresentados como uma alternativa para a análise de regressão de dados de contagem, generalizando os modelos tradicionais, como o Consul, Poisson generalizada, binomial negativa generalizada, entre outros. Nesta dissertação, versões modificadas (via [Bartlett \(1937\)](#) e [Rocke \(1989\)](#)) das estatísticas da razão de verossimilhanças e razão de verossimilhanças bootstrap, respectivamente, são apresentadas para testes de hipóteses nesta classe de modelos. Simulações de Monte Carlo mostram que os testes baseados nas versões modificadas das estatísticas da razão de verossimilhanças e razão de verossimilhanças bootstrap exibem melhores desempenhos em amostras finitas, superando os testes originais (sem correção). Além disso, técnicas de diagnóstico são propostas para os MSPNLGs, a saber: resíduos, alavancagem generalizada, influência global, e influência local. Finalmente, um conjunto de dados reais é utilizado para avaliar as metodologias desenvolvidas tanto em técnicas de diagnóstico como em aperfeiçoamento de testes.

Palavras-chave: Alavancagem Generalizada. Bootstrap. Correção de Bartlett. Estatística da razão de verossimilhanças. Influência global. Influência local. Modelos em série de potência não-lineares generalizados.

Abstract

The Power Series Generalized nonlinear Models (PSGNLMs), introduced in [Cordeiro et al. \(2009\)](#), are presented as a prominent option to deal with counting regression models, generalizing well known models as the Consul, the generalized Poisson and the generalized negative binomial, among others. In this work, modified versions (via [Bartlett \(1937\)](#) and [Rocke \(1989\)](#)) of the likelihood ratio and bootstrap likelihood ratio test statistics, respectively, are presented for hypothesis testing in this class of models. Monte Carlo simulations show that tests based on modified versions of the likelihood test statistic exhibit better performance in finite samples, exceeding the original tests (uncorrected). Furthermore, diagnostic techniques are proposed for the MSPNLGs, namely: residuals, generalized leverage, global influence, and local influence. Finally, a real data set is used in order to assess the developed techniques in both areas, influence diagnostic and hypothesis testing.

Key Words: Bootstrap. Bartlett correction. Generalized leverage. Global influence. Likelihood ratio test local influence. Power Series Generalized Nonlinear Models.

Los Modelos en Series de Potencias no Lineales Generalizados (MSPNLGs), introducidos recientemente en [Cordeiro et al. \(2009\)](#), son presentados como una alternativa destacada para el análisis de regresión en datos de conteo, generalizando varios modelos tradicionales como el Consul, Poisson generalizada y binomial negativa generalizada, entre otros. En esta disertación, versiones modificadas (via [Bartlett \(1937\)](#) y [Rocke \(1989\)](#)) de las estadísticas de razón de verosimilitud y razón de verosimilitud bootstrap, respectivamente, son presentadas para pruebas de hipótesis en esta clase de modelos. Simulaciones de Monte Carlo muestran que los testes basados en las versiones modificadas de las estadísticas de razón de verosimilitud y razón de verosimilitud bootstrap tienen mejores desempeños en muestras finitas, superando los testes originales (sin corrección). Adicionalmente, técnicas de diagnóstico son propuestas para los MSPNLGs, a saber: residuos, Leverage generalizado, influencia global, e influencia local. Finalmente, un conjunto de datos reales es utilizando para evaluar las metodologías desarrolladas, tanto en técnicas de diagnóstico como en mejoramiento de pruebas de hipótesis.

Palabras claves: Bootstrap. Corrección de Bartlett. Estadística de razón de verosimilitud. Influencia global. Influencia local. Modelos en serie de potencias no lineales generalizados. Leverage generalizado.

Lista de Figuras

3.1	Distorção de tamanho dos testes no modelo linear BNG, com $n = 30$	39
3.2	Distorção de tamanho dos testes no modelo linear Consul, com $n = 30$	40
3.3	Distorção de tamanho dos testes no modelo linear GPO, com $n = 30$	41
3.4	Distorção de tamanho dos testes no modelo não linear BNG, com $n = 30$	47
3.5	Distorção de tamanho dos testes no modelo não linear Consul, com $n = 30$	48
3.6	Distorção de tamanho dos testes no modelo não linear GPO, com $n = 30$	49
5.1	Gráficos dos resíduos contra os índices para o modelo GPO ajustado	65
5.2	Gráficos dos resíduos contra os valores ajustados para o modelo GPO ajustado	65
5.3	Gráfico normal de probabilidade para o caso da distribuição GPO	66
5.4	Gráficos de alavancagem generalizada e da Influência local para o caso da distribuição GPO	66
5.5	Gráficos da Influência local para o caso da distribuição GPO	67
5.6	Gráficos dos resíduos contra os índices para o modelo Consul ajustado	68
5.7	Gráficos dos resíduos contra os valores ajustados para o modelo Consul ajustado	68
5.8	Gráfico normal de probabilidade para o caso da distribuição Consul	69
5.9	Gráficos de alavancagem generalizada e da Influência local para o caso da distribuição Consul	69
5.10	Gráficos da Influência local para o caso da distribuição Consul	70
6.1	Gráfico de dispersão para o banco de dados considerado	73
6.2	Gráficos dos resíduos contra os índices para o modelo Delta Binomial ajustado aos dados dos peixes da Tabela C.1	76
6.3	Gráficos dos resíduos contra os valores ajustados para o modelo Delta Binomial ajustado aos dados dos peixes da Tabela C.1	76

6.4	Gráfico normal de probabilidades para o resíduo de Comp. de desvio	77
6.5	Gráficos da Influência global para o caso do banco de dados	77
6.6	Gráficos dos índices de C_i sob o modelo Delta Binomial ajustado aos dados dos peixes da Tabela C.1	78

Lista de Tabelas

2.1	Algumas funções da família MSPNLGs	22
2.2	Funções q, t para algumas funções da família MSPNLGs	24
3.1	Preditores lineares propostos para avaliação numérica	33
3.2	Preditores não lineares propostos para avaliação numérica	34
3.3	Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ para as estatísticas $LR, LR^*, \widetilde{LR}, LR_{boot}$ e LR_{boot}^* para modelos lineares com $p = 8$ e vários valores de n	36
3.4	Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ para os testes $LR, LR^*, \widetilde{LR}, LR_{boot}$ e LR_{boot}^* em modelos lineares com $n = 30$ e vários valores de p	37
3.5	Poder dos testes $LR, LR^*, \widetilde{LR}, LR_{boot}$ e LR_{boot}^* com $n = 30, p = 4$ e $\alpha = 5\%$ no caso linear	38
3.6	Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ para as estatísticas $LR, LR^*, \widetilde{LR}, LR_{boot}$ e LR_{boot}^* para modelos não lineares com $p = 8$ e vários valores de n	44
3.7	Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ para os testes $LR, LR^*, \widetilde{LR}, LR_{boot}$ e LR_{boot}^* em modelos não lineares com $n = 30$ e vários valores de p	45
3.8	Poder dos testes $LR, LR^*, \widetilde{LR}, LR_{boot}$ e LR_{boot}^* com $n = 30, p = 4$ e $\alpha = 10\%$ no caso não linear	46
6.1	Valores do ajuste dos modelos (6.1) e (6.2)	72
6.2	Estimativas de máxima verossimilhança dos parâmetros do modelo DB ajustado aos dados dos peixes da Tabela C.1	74
6.3	Razão de mudanças para os pontos sinalizados	75
C.1	Número de espécies de peixe em um lago (y) e o logaritmo da área do lago, em $km^2, (x)$	96

1	Introdução	15
1.1	Objetivos	16
1.2	Organização da Dissertação	17
1.3	Suporte Computacional	18
2	Modelos em Séries de Potência Não Lineares Generalizados (MSPNLGs)	19
2.1	Modelos em Séries de Potência não Lineares Generalizados	20
2.2	Estimação dos Parâmetros de Regressão	23
2.3	Testes de Hipótese em MSPNLGs com ϕ Conhecido	26
2.4	Correção de Bartlett	27
2.5	Teste da Razão de Verossimilhanças Bootstrap	30
3	Avaliação Numérica - Aperfeiçoamento de testes de hipóteses	32
3.1	Cenário Linear	34
3.2	Cenário Não linear	42
4	Técnicas de diagnóstico nos MSPNLGs	50
4.1	Análise de Resíduos	52
4.1.1	Resíduo Componente de Desvio	53
4.1.2	Resíduo Componente de Desvio Padronizado	54
4.2	Alavancagem Generalizada	54
4.3	Influência Global	55
4.4	Influência Local	57
4.4.1	Perturbação Aditiva na Resposta	59

4.4.2	Perturbação de Casos Ponderados	59
4.4.3	Perturbação Aditiva no Preditor	60
5	Avaliação Numérica - Técnicas de Diagnóstico	62
6	Aplicação dos Métodos de Diagnóstico	71
7	Conclusões	79
	Referências	82
A	Obtenção dos cumulantes	88
A.1	Derivadas de $\ell(\beta; y)$	90
A.2	Cálculo dos cumulantes	92
A.3	Derivadas dos cumulantes	92
B	Influência Global	94
C	Banco de dados	96

CAPÍTULO 1

Introdução

Em [Cordeiro et al. \(2009\)](#) foram introduzidos os *Modelos em Séries de Potência Não Lineares Generalizados* (MSPNLGs) com o intuito de unificar em uma só estrutura conceitual vários dos principais modelos de regressão discreta. Os MSPNLGs abrangem modelos tradicionais tais como Poisson generalizada, binomial negativa generalizada, entre outros. Esta família de distribuições discretas tem uma estrutura bastante flexível para a modelagem de dados discretos.

Para realizar testes de hipóteses nos MSPNLGs, podem ser usadas as tradicionais estatísticas da razão de verossimilhanças (LR), escore (SR), Wald (W) e a recentemente explorada estatística gradiente (G), introduzida em [Terrell \(2002\)](#). Neste trabalho será dada ênfase a estatística da LR . Tal estatística, sob a hipótese nula tem uma distribuição aproximada qui-quadrado, com erro de ordem até n^{-1} onde n é o tamanho de amostra ¹. Desta forma, torna-se importante obter ajustes para a estatística LR , quando o tamanho de amostra for pequeno pois testes assintóticos baseados em pequenas amostras podem conduzir a resultados distorcidos. A ideia é modificar essas estatísticas por um fator de correção, visando produzir novas esta-

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 de taxa de convergência de a_n para zero.

tísticas com o primeiro momento igual ao da distribuição qui-quadrado de referência, ou seja, obtendo estatísticas com distribuição aproximada qui-quadrado com erro de ordem até n^{-2} .

Na literatura podem-se encontrar trabalhos como os de [Cordeiro & Paula \(1989\)](#) em que emprega-se um fator de correção de Bartlett para modificar a estatística da razão de verossimilhança nos modelos não lineares da família exponencial ou [Cordeiro et al. \(1993\)](#), baseado em [Cordeiro & Ferrari \(1991\)](#), em que emprega-se a correção tipo-Bartlett para modificar a estatística escore nos modelos lineares generalizados, entre muitas outras. Uma revisão bibliográfica no tema pode consultada em [Cordeiro & Cribari-Neto \(2014\)](#).

Uma etapa importante na análise de um ajuste de regressão é o estudo da robustez dos resultados obtidos com relação à presença de pontos extremos ou, mais genericamente, a perturbação no modelo adotado. Detectar observações aberrantes e/ou influentes constitui um passo importante na análise do conjunto de dados

Na literatura recente, podem-se encontrar trabalhos como os de [De Bastiani et al. \(2014\)](#), em que emprega-se as técnicas de influência e diagnóstico nos modelos espaciais lineares com distribuição da família de contornos elípticos, e em [Villegas et al. \(2013\)](#) em que são estudados os modelos lineares generalizados simétricos.

1.1 Objetivos

Esta dissertação tem três objetivos. O primeiro é reunir resultados importantes sobre fatores de correção de Bartlett para as estatísticas da razão de verossimilhanças e razão de verossimilhanças bootstrap, respectivamente, nos MSPNLGs e apresentar algumas aplicações a dados reais. O segundo é apresentar os resultados de simulação para averiguar o efeito das correções nesta classe de modelos discretos no que tange ao tamanho e poder em amostras finitas. O terceiro é apresentar técnicas de diagnóstico em modelos em séries de potência não lineares generalizados.

Especificamente:

1. Para os MSPNLGs, baseados em [Efron \(1979\)](#) e [Rocke \(1989\)](#), propor duas versões modificadas da estatística da razão de verossimilhanças nesta classe de modelos. Tais estatística corrigidas são obtidas pela metodologia de bootstrap e consideram-se uma contribuição original do presente trabalho na área de aperfeiçoamento de testes de hipótese nesta classe de modelos.

2. Comparar os desempenhos dos testes baseados nas estatísticas razão de verossimilhanças bootstrap e razão de verossimilhanças bootstrap corrigida, via Rocke com os demais testes propostos em [Silva \(2010\)](#), a saber: os testes baseados nas estatísticas da razão de verossimilhanças, razão de verossimilhanças corrigida, via Bartlett (e sua versão equivalente). Avaliaremos o impacto da quantidade de parâmetros de perturbação, o poder e o tamanho dos testes supracitados.
3. Propor resíduos e medidas de diagnóstico e influência nos MSPNLGs baseados nas metodologias de alavancagem generalizada, influência global e influência local sob vários cenários de perturbação. Vale ressaltar que este quesito constitui uma contribuição completamente nova devido ao fato de não existir estudos na literatura neste sentido nesta classe de modelos.

1.2 Organização da Dissertação

Esta dissertação está organizada da seguinte forma:

No Capítulo 2 procuramos revisar os principais resultados teóricos relacionados com os MSPNLG. Em particular, mostramos como ficam os testes da razão de verossimilhança, escore, Wald e gradiente, nesta classe de modelos. Além disso, revisamos os principais resultados que tratam da obtenção do fator de correção de Bartlett nos MSPNLGs. Obtemos o fator de correção de Bartlett bootstrap nesta classe de modelos.

No Capítulo 3, apresentamos resultados de simulação para avaliar o efeito das correções para testes em MSPNLGs discretos e, ainda, analisamos a influência do número de parâmetros de perturbação no desempenho dos testes.

No Capítulo 4 apresentamos os métodos de diagnóstico nos MSPNLG, tais como: resíduos, envelope, alavancagem generalizada, influência global e local. Sendo assim, serão apresentadas formas matriciais que foram obtidas para as metodologias de alavancagem generalizada, influência global e influência local, sendo a última realizada sob vários cenários de perturbação.

O Capítulo 5 apresenta os resultados das simulações referentes às técnicas de diagnóstico nos MSPNLG. Foram simulados dados e posteriormente aplicadas as metodologias de alavancagem generalizada, estudos dos resíduos, influência global e influência local sob vários esquemas de perturbação, com o intuito de detectar pontos que sejam influentes na hora da

obter as estimativas dos parâmetros pelo método de máxima verossimilhança.

O Capítulo 6 apresenta uma aplicação das metodologias apresentadas na dissertação em um conjunto de dados reais, objetivando mostrar a utilidade dos resultados.

Finalmente, as principais conclusões do presente trabalho são apresentadas no Capítulo 7. No Apêndice A é apresentada a derivação das quantidades necessárias para o fator de correção de Bartlett da estatística da razão de verossimilhança. No Apêndice B são encontradas quantidades necessárias para o estudo de influência global. Finalmente, no Apêndice C são apresentados os dados utilizados para avaliar as técnicas desenvolvidas nesta dissertação.

1.3 Suporte Computacional

Todos os resultados das simulações apresentadas na presente dissertação foram obtidas usando a linguagem livre de programação matricial `0x`, na sua versão 7.0. Esta linguagem foi criada por Jurgen Doornik em 1994 na Universidade de Oxford, Inglaterra, (Doornik, 2006). As maximizações não lineares, necessárias para o cálculo das quantidades relacionadas com a função de máxima verossimilhança, foram feitas através do algoritmo quasi-Newton BFGS, disponível nas funções pre-definidas da linguagem `0x`, (Nocedal & Wright, 1999, Cap. 8) e (Luenberger, 1973, Cap. 12)

Os gráficos apresentados foram realizados no ambiente de programação livre R na sua versão 0.98.501 para plataforma Windows, com ajuda do editor Rstudio. Esta linguagem foi criada por Ross Ihaka e Robert Gentleman na Universidade de Auckland (Ihaka & Gentleman, 1996).

A elaboração desta dissertação foi feita integralmente no sistema tipográfico $\text{\LaTeX}$ que, basicamente, é uma série de macros e rotinas baseadas no sistema $\text{\TeX}$. Esta linguagem foi criada por Donald Knuth em 1984 na Universidade de Stanford, USA (Knuth, 1984). Como plataforma foi usada a versão 2.02 de TeXnicCenter.

Modelos em Séries de Potência Não Lineares Generalizados (MSPNLGs)

Com o intuito de unificar vários dos modelos de regressão utilizados para dados de contagem, em [Cordeiro et al. \(2009\)](#) foram propostos os modelos em séries de potências não lineares generalizados (MSPNLGs). Uma grande vantagem na hora de utilizar os MSPNLGs é que são uma unificação de vários dos modelos mais utilizados para modelar este tipo de dados como o Poisson e binomial negativa ([Cameron & Trivedi, 1998](#)). Assim, tem-se uma estrutura conceitual que pode modelar fenômenos como sobre-dispersão e sub-dispersão, ou seja, quando a variação observada excede aquela assumida pelo modelo e quando a variação observada é menor do que aquela assumida pelo modelo, respectivamente.

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

Neste capítulo apresentaremos a definição dos MSPNLGs, iremos mostrar o processo iterativo para a estimação de máxima verossimilhança dos parâmetros do preditor que pode ser

visto como um método de mínimos quadrados ponderados, junto com o vetor escore e a matriz de informação de Fisher nesta classe de modelos. Serão também apresentadas quatro estatísticas de teste para esses modelos, a saber: razão de verossimilhança, escore, Wald e gradiente. Além disso, é feita uma revisão do fator de correção de Bartlett para a estatística da razão de verossimilhança, inicialmente obtido em [Silva \(2010\)](#), vale ressaltar que existem estudos em correção tipo-Bartlett para a estatística escore nesta classe de modelos, vide [Lima \(2012\)](#). Finalmente, propomos uma correção de Bartlett para a estatística da razão de verossimilhança bootstrap, baseada nas ideias apresentadas em [Rocke \(1989\)](#).

2.1 Modelos em Séries de Potência não Lineares Generalizados

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

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

em que $y_i \in A_p = \{p, p+1, \dots\}$ com $p \geq 0$ inteiro positivo, $a(y, \phi)$ é positiva e as funções analíticas $g(\mu_i, \phi)$ e $f(\mu_i, \phi)$ dos parâmetros μ_i e ϕ são positivas, finitas e duas vezes diferenciáveis. Assumindo que o parâmetro $\phi > 0$ é conhecido, temos que a variável aleatória Y tem uma distribuição de probabilidade que fica completamente determinada pela sua função de variância.

Para a família de distribuições dada em (2.1), tem-se que:

$$E(Y_i) = \mu_i = \frac{f'_i g_i}{f_i g'_i} \quad \text{e} \quad \text{Var}(Y_i) = V(\mu_i, \phi) = \frac{g_i}{g'_i},$$

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

$$h_i = h(\mu_i) = \eta_i = \eta(x_i; \beta), \quad i = 1, \dots, n, \quad (2.2)$$

em que $h(\cdot)$ é uma função de ligação conhecida e duplamente diferenciável, $\beta = (\beta_1, \dots, \beta_p)^\top$, é um vetor de k ($k < n$) parâmetros desconhecidos a serem estimados, $x_i = (x_{i1}, \dots, x_{ik})^\top$

representa os valores de k variáveis explicativas e $\eta(\cdot; \cdot)$ é uma função possivelmente não linear no segundo argumento, contínua e diferenciável com respeito aos componentes de β tal que a matriz de derivadas

$$\tilde{X} = \tilde{X}(\beta) = \partial\eta/\partial\beta^\top, \quad (2.3)$$

com $\eta = (\eta_1, \dots, \eta_n)^\top$, tem posto p para todo β . A matriz $\tilde{X}$ tem elementos que são, em geral, funções do vetor de parâmetros β desconhecidos. Os MSPNLGs estão definidos por (2.1) e (2.2).

Características das distribuições pertencentes a família (2.1), são apresentadas na Tabela 2.1.

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

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

2.2 Estimação dos Parâmetros de Regressão

O logaritmo da função verossimilhança para os modelos MSPNLGs, dado o vetor de observações $\mathbf{y} = (y_1, \dots, y_n)^\top$ e definida como sendo

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

O r -ésimo elemento do vetor escore para o parâmetro β_r , é dado por

$$\begin{aligned} U_{\beta_r} = \frac{\partial \ell(\boldsymbol{\beta}; \mathbf{y})}{\partial \beta_r} &= \sum_i y_i \frac{1}{g(\mu_i, \phi)} \frac{\partial g(\mu_i, \phi)}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_r} - \sum_i \frac{1}{f(\mu_i, \phi)} \frac{\partial f(\mu_i, \phi)}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_r} \\ &= \sum_i y_i \frac{1}{g_i} g_i^{(1)} \frac{1}{h_i^{(1)}} \tilde{x}_{ir} - \sum_i \frac{1}{f_i} f_i^{(1)} \frac{1}{h_i^{(1)}} \tilde{x}_{ir} = \sum_i (y_i t_i - q_i) \tilde{x}_{ir}, \end{aligned}$$

que, em notação matricial, fica dado por $U_{\boldsymbol{\beta}} = \tilde{\mathbf{X}}^\top (T\mathbf{y} - Q)$, $\tilde{\mathbf{X}}$ é dado em (2.3), $T = \text{diag}\{t_1, \dots, t_n\}$ é uma matriz diagonal de dimensão $n \times n$ e $Q = (q_1, \dots, q_n)^\top$ é um vetor $n \times 1$. As respectivas i -ésimas entradas de T e Q são dadas por

$$t_i = \frac{g'_i}{g_i h'_i} \quad \text{e} \quad q_i = \frac{f'_i}{f_i h'_i},$$

com $h'_i = h'(\mu_i)$ e $i = 1, \dots, n$. A tabela 2.2 apresenta as funções t e q para várias distribuições pertencentes na classe MSPNLGs.

A matriz de informação de Fisher de $\boldsymbol{\beta}$, dado ϕ , é dado por $K_{\boldsymbol{\beta}\boldsymbol{\beta}} = \tilde{\mathbf{X}}^\top W \tilde{\mathbf{X}}$, em que $W = \text{diag}\{w_1, \dots, w_n\}$ é uma matriz diagonal $n \times n$ com

$$w_i = \left(q'_i - \frac{f'_i g_i}{f_i g'_i} t'_i \right) \frac{1}{h'_i} \quad i = 1, \dots, n. \quad (2.5)$$

No Apêndice A encontra-se a segunda derivada do logaritmo da função de verossimilhança. Substituindo as quantidades $f_i, g_i, f'_i, g'_i, h'_i, t'_i$ e q'_i na expressão (2.5), Silva (2010) mostra que W pode ser rescrito como $W = (LVL)^{-1}$ em que $L = \text{diag}\{h'_1, \dots, h'_n\}$ e $V = \text{diag}\{V_1, \dots, V_n\}$ sendo $V_i = V(\mu_i, \phi)$.

Distribuição	$t(\mu_i, \phi)$	$q(\mu_i, \phi)$	Suporte (A_p)
1. Poisson	$\frac{1}{\mu_i! h_i}$	$\frac{1}{h_i}$	$\{0, 1, 2, \dots\}$
2. Binomial	$\frac{m}{\mu_i(m-\mu_i)h_i}$	$\frac{m}{(m-\mu_i)h_i}$	$\{0, 1, 2, \dots, m\}$
3. Binomial negativa	$\frac{\phi}{\mu_i(\mu_i+\phi)h_i}$	$\frac{\phi}{(\mu_i+\phi)h_i}$	$\{0, 1, 2, \dots\}$
4. Poisson generalizada	$\frac{1}{\mu_i(1+\phi\mu_i)^2 h_i}$	$\frac{1}{(1+\phi\mu_i)^2 h_i}$	$\{0, 1, 2, \dots\}$
5. Borel	$\frac{1}{\mu_i^2(1-\mu_i)h_i}$	$\frac{1}{\mu_i(1-\mu_i)h_i}$	$\{1, 2, \dots\}$
6. Consul	$\frac{\phi}{\mu_i(\mu_i-1)\{\mu_i(\phi-1)+1\}h_i}$	$\frac{\phi}{(\mu_i-1)\{\mu_i(\phi-1)+1\}h_i}$	$\{1, 2, \dots\}$
7. Binomial negativa generalizada	$\frac{v^2}{\mu_i(v+\mu_i\phi)\{v+\mu_i(\phi-1)\}h_i}$	$\frac{v^2}{(v+\mu_i\phi)\{v+\mu_i(\phi-1)\}h_i}$	$\{0, 1, 2, \dots\}$
8. Borel-Tanner	$\frac{m^2}{\mu_i(\mu_i-m)h_i}$	$\frac{m^2}{(\mu_i-m)h_i}$	$\{m, m+1, \dots\}$
9. Delta binomial	$\frac{m^2\phi}{\mu_i\{\mu_i(\phi-1)+m\}\{(\mu_i-m)h_i\}}$	$\frac{m^2\phi}{\{\mu_i(\phi-1)+m\}(\mu_i-m)h_i}$	$\{m, m+1, \dots\}$
10. Geeta	$\frac{\mu_i(\mu_i-1)(\phi\mu_i-1)h_i}{m^2(\phi-1)}$	$\frac{(\mu_i-1)(\phi\mu_i-1)h_i}{m^2(\phi-1)}$	$\{1, 2, \dots\}$
11. Geeta-m	$\frac{\mu_i(\mu_i-m)(\phi\mu_i-m)h_i}{1}$	$\frac{(\mu_i-m)(\phi\mu_i-m)h_i}{1}$	$\{m, m+1, \dots\}$
12. Haigh	$\frac{\mu_i(2\mu_i-1)(\mu_i-1)h_i}{1}$	$\frac{1}{(2\mu_i-1)(\mu_i-1)h_i}$	$\{1, 2, \dots\}$

Tabela 2.2: Funções q , t e o suporte de algumas distribuições da família (2.1).

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

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

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

$$\begin{aligned} \beta^{(\phi)(k+1)} &= \beta^{(\phi)(k)} + (\tilde{X}^{(\phi)(k)\top} W^{(\phi)(k)} \tilde{X}^{(\phi)(k)})^{-1} \tilde{X}^{(\phi)(k)\top} (T^{(\phi)(k)} y - Q^{(\phi)(k)}) \\ &= (\tilde{X}^{(\phi)(k)\top} W^{(\phi)(k)} \tilde{X}^{(\phi)(k)})^{-1} \tilde{X}^{(\phi)(k)\top} W^{(\phi)(k)} \delta^{(\phi)(k)}, \quad k = 0, 1, \dots, \end{aligned} \quad (2.6)$$

em que

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

Em (2.6), $\delta^{(\phi)(k)}$ desempenha o papel de uma variável dependente modificada, enquanto $W^{(\phi)(k)}$ é uma matriz de pesos que muda a cada passo do processo iterativo. O valor inicial $\beta^{(\phi)(0)}$ pode ser obtido, por exemplo, ajustando um modelo log-não linear. O estimador restrito $\hat{\beta}^{(\phi)}$ tem uma distribuição assintoticamente normal com média β e matriz de covariâncias $(K_\beta)^{-1}$ consistentemente estimada por $\left(\hat{\tilde{X}}^{(\phi)\top} \hat{W}^{(\phi)} \hat{\tilde{X}}^{(\phi)} \right)^{-1}$.

A estimação do parâmetro ϕ , quando o mesmo é desconhecido, pode ser feita diretamente pelo método da máxima verossimilhança. No entanto, esse método torna-se bastante complexo para algumas distribuições pertencentes à família (2.1), visto que a função $a(y, \phi)$ usualmente envolve razão de funções gama com argumentos que dependem de ambos y e ϕ e possíveis constantes. Portanto, a equação não linear para ϕ contém soma de funções digama. Esta dificuldade pode ser evitada estimando o parâmetro ϕ por métodos indiretos. A seguir, é descrito um desses métodos.

De posse de $\hat{\beta}^{(\phi)}$, a estimativa de ϕ pode ser obtida inserindo $\hat{\mu}_i^{(\phi)} = h^{-1}(\eta(x_i; \hat{\beta}^{(\phi)}))$ em

(2.4) e maximizando o logaritmo da função de verossimilhança perfilada $\ell_p(\phi) = \ell(\hat{\beta}^{(\phi)}, \phi)$ de ϕ dada por $\ell_p(\phi) = n \log \{ \tilde{J}(\hat{\mu}^{(\phi)}, \phi; \mathbf{y}) \}$, onde $\tilde{J}(\hat{\mu}^{(\phi)}, \phi; \mathbf{y})$ é média geométrica das quantidades

$$\tilde{J}(\hat{\mu}_i^{(\phi)}, \phi; y_i) = \frac{a(y_i, \phi) g(\hat{\mu}_i^{(\phi)}, \phi)^{y_i}}{f(\hat{\mu}_i^{(\phi)}, \phi)}$$

para $i = 1, \dots, n$. $\tilde{J}(\hat{\mu}_i^{(\phi)}, \phi; y_i)$ é a função da observação y_i , do i -ésimo valor ajustado $\hat{\mu}_i^{(\phi)}$ e do parâmetro de dispersão ϕ . Assim, para obter o valor estimado de ϕ , é preciso achar o valor

$$\hat{\phi} = \underset{\phi}{\operatorname{argmax}} \left\{ \tilde{J}(\hat{\mu}^{(\phi)}, \phi; \mathbf{y}) \right\}.$$

Uma vez que o estimador de máxima verossimilhança de ϕ , ou seja $\hat{\phi}$, é obtido, pode se trocar na expressão (2.6) para obter estimativas não restritas de $\hat{\beta} = \hat{\beta}^{(\hat{\phi})}$.

2.3 Testes de Hipótese em MSPNLGs com ϕ Conhecido

Seja $\mathbf{y} = (y_1, \dots, y_n)^\top$ uma amostra aleatória de tamanho n e cujo logaritmo da função de verossimilhança $\ell(\beta; \mathbf{y})$, dado por (2.4), dependendo do parâmetro desconhecido $\beta = (\beta_1, \dots, \beta_p)^\top$. Assumimos que $\ell(\beta; \mathbf{y})$ seja regular com respeito aos componentes de β até quarta ordem.

Considere que o vetor de parâmetros β pode ser decomposto como $\beta = (\beta_1^\top, \beta_2^\top)^\top$, sendo $\beta_1 = (\beta_1, \dots, \beta_q)^\top$ o vetor de parâmetros de interesse e $\beta_2 = (\beta_{q+1}, \dots, \beta_p)^\top$ o vetor de parâmetros de perturbação. Nosso interesse é testar as seguintes hipóteses, a saber: $H_0 : \beta_1 = \beta_1^{(0)}$, contra $H_1 : \beta_1 \neq \beta_1^{(0)}$, em que $\beta_1^{(0)}$ é um vetor especificado de dimensão q ($q \leq p$). Podemos definir a estatística da razão de verossimilhanças como

$$LR = 2 \left\{ \ell(\hat{\beta}; \mathbf{y}) - \ell(\tilde{\beta}; \mathbf{y}) \right\}, \quad (2.7)$$

em que $\hat{\beta}$ é o estimador de máxima verossimilhança irrestrito de β e $\tilde{\beta} = (\beta_1^{(0)\top}, \tilde{\beta}_2^\top)^\top$ é o estimador de máxima verossimilhança restrito de β .

A estatística escore apresentada em Rao (1948), para testar a hipótese estabelecida é expressa como sendo

$$SR = U_{\tilde{\beta}_1}^\top K^{\beta_1 \beta_1} U_{\tilde{\beta}_1}, \quad (2.8)$$

em que $U_{\tilde{\beta}_1}$ é a parte do vetor escore correspondente ao vetor de parâmetros de interesse,

avaliada em $\tilde{\beta}_1$, $K_{\beta\beta}$ é a matriz de informação de Fisher, ou seja

$$K_{\beta\beta} = \begin{bmatrix} K_{\beta_1\beta_1} & K_{\beta_1\beta_2} \\ K_{\beta_2\beta_1} & K_{\beta_2\beta_2} \end{bmatrix}.$$

Note-se que o termo $K^{\beta_1\beta_1}$ na equação (2.8) corresponde ao primeiro bloco da matriz inversa $K_{\beta\beta}^{-1}$. Assim, a matriz inversa de $K_{\beta\beta}$ é dada por

$$K_{\beta\beta}^{-1} = \begin{bmatrix} K^{\beta_1\beta_1} & K^{\beta_1\beta_2} \\ K^{\beta_2\beta_1} & K^{\beta_2\beta_2} \end{bmatrix},$$

onde $K^{\beta_1\beta_1} = K_{\beta_1\beta_1}^{-1} + (FE^{-1}F^\top)$, $K^{\beta_2\beta_1} = -E^{-1}F^\top$, $K^{\beta_1\beta_2} = FE^{-1}$ e $K^{\beta_2\beta_2} = E^{-1}$, onde $E = K_{\beta_2\beta_2} - (K_{\beta_2\beta_1}K_{\beta_1\beta_1}^{-1}K_{\beta_1\beta_2})$ e $F = K_{\beta_1\beta_1}^{-1}K_{\beta_1\beta_2}$.

A estatística Wald, para o teste de hipótese considerado é definida como sendo

$$W = (\hat{\beta}_1 - \beta_1^{(0)})^\top K^{\beta_1\beta_1} (\hat{\beta}_1 - \beta_1^{(0)}). \quad (2.9)$$

A estatística gradiente apresentada em Terrell (2002), para o teste de hipótese é definida como

$$G = U_{\hat{\beta}_1}^\top (\hat{\beta}_1 - \beta_1^{(0)}). \quad (2.10)$$

As estatísticas da razão de verossimilhanças (2.7), escore (2.8), Wald (2.9) e gradiente (2.10) têm, sob a hipótese nula, distribuição assintótica χ_q^2 . Rejeite-se a hipótese nula H_0 , ao nível de significância α , se $LR, SR, W, G > \chi_{(\alpha;q)}^2$ em que $\chi_{(\alpha;q)}^2$ é o percentil $(1 - \alpha)$ da distribuição χ_q^2 e q é o número de restrições impostas sob H_0 .

2.4 Correção de Bartlett

Testes estatísticos clássicos são usualmente baseados em aproximações assintóticas de primeira ordem, que podem ser pouco acuradas quando a amostra é pequena ou mesmo de tamanho moderado. Em geral, esse é o caso da teste da razão de verossimilhanças (LR) cuja estatística tem distribuição assintótica χ^2 sob a hipótese nula. Uma alternativa para resolver este problema é usar a teoria assintótica de alta ordem, que pode tornar métodos inferenciais tradicionais mais precisos em pequenas amostras. Em problemas regulares, a estatística LR tem distribuição χ^2 a menos de um erro de ordem n^{-1} sob hipótese nula, enquanto que a distribuição da estatística de sua versão corrigida é χ^2 com erro de ordem n^{-2} onde n é o tamanho de amostra.

Correção de Bartlett para a estatística da razão de verossimilhanças em diversos modelos podem ser encontradas na literatura; ver por exemplo, [Cordeiro \(1983\)](#), que obteve um fator de correção de Bartlett para a estatística da razão de verossimilhanças nos MLGs considerando o fator de escala conhecido. [Cordeiro \(1987\)](#) generalizou os resultados de [Cordeiro \(1983\)](#) em que obteve um fator de correção de Bartlett nos MLGs quando o fator de escala é desconhecido. [Cordeiro & Paula \(1989\)](#), obtiveram o fator de correção de Bartlett para a estatística LR nos modelos não lineares da família exponencial com parâmetro de dispersão conhecido. Fatores de correção de Bartlett para a estatística da razão de verossimilhanças foram obtidos nos seguintes modelos: Em [Cribari-Neto & Ferrari \(1995\)](#) para os modelos lineares normais heteroscedásticos; e [Cribari-Neto & Zarkos \(1995\)](#) para os modelos de regressão multivariada; [Cordeiro et al. \(1995\)](#) para a família exponencial uniparamétrica; [Ferrari & Arellano \(1996\)](#) para os modelos de regressão com erros t de student; [Ferrari & Uribe \(2001\)](#) para os modelos lineares simétricos, que após foi generalizado por [Cordeiro \(2004\)](#) para os modelos não lineares simétricos; [Cordeiro & Montenegro \(2001\)](#) para os modelos não lineares de locação e escala supondo que o parâmetro de escala é conhecido. O fator de correção de Bartlett para a estatística da razão de verossimilhança perfilada foi obtido em [Ferrari et al. \(2004\)](#) para o modelo de regressão normal linear heteroscedástico e por [Cysneiros & Ferrari \(2006\)](#) para os modelos de regressão não lineares da família exponencial. [Lemonte et al. \(2010\)](#) propôs a correção de Bartlett nos modelos de regressão Birnbaum–Saunders, estudos que foram generalizados para o caso não linear em [Lemonte et al. \(2012\)](#). Uma aplicação da estatística na medicina pode-se encontrar em [Noma \(2011\)](#) onde é abordado um problema em meta análise com correção de Bartlett. Finalmente, em [Vargas et al. \(2014\)](#) foram obtidas correções tipo-Bartlett para a estatística gradiente nos modelos lineares generalizados.

Foi mostrado em [Frydenberg & Jensen \(1989\)](#) por meio de simulação, que os resultados teóricos que garantem que a distribuição da estatística corrigida tenha uma boa aproximação com a distribuição χ^2 são válidos apenas para os modelos contínuos. No entanto, [Cordeiro \(1982\)](#) fez vários estudos de simulação envolvendo distribuições multinomial e de Poisson que mostraram que os testes modificados, baseados em LR^* , apresentam taxas de rejeição da hipótese nula bem mais próximas dos respectivos níveis nominais do que o teste não modificado. As correções dos modelos discretos também se mostraram eficazes nos estudos de simulação realizados por [Cysneiros \(1997\)](#) que obteve correções de Bartlett e tipo-Bartlett nos modelos lineares generalizados contínuos e discretos. Finalmente, tais correções também se mostraram competentes nos resultados desenvolvidos nesta dissertação apresentados no Capítulo 3.

Em problemas regulares, a estatística LR converge para a distribuição χ^2 . Entretanto,

em pequenas amostras, a aproximação de primeira ordem pode não ser satisfatória. Com o objetivo de melhorar esta aproximação [Bartlett \(1937\)](#) propôs multiplicar a estatística LR por uma constante $(1 + d)^{-1}$ ou $(1 - d)$ obtendo-se assim novas estatísticas LR^* e $\widetilde{LR}$, conhecidas como estatísticas da razão de verossimilhanças corrigida por Bartlett, definidas por

$$LR^* = LR/(1 + d) \quad \text{ou} \quad \widetilde{LR} = LR(1 - d),$$

onde os fatores de correção de Bartlett $1/(1 + d)$ e $(1 - d)$ são obtidos mediante

$$d = \frac{\epsilon_p - \epsilon_{p-q}}{q}, \quad (2.11)$$

em que os termos ϵ_p e ϵ_{p-q} são expressões que envolvem cumulantes e derivadas de até quarta ordem do logaritmo da função de verossimilhanças.

Em (2.11), d é uma constante de ordem de ordem n^{-1} escolhida apropriadamente tal que, sob H_0 , $E(LR^*) = q + O(n^{-3/2})$. É possível mostrar que $P(LR^* > x_\alpha) = \alpha + O(n^{-2})$, enquanto que $P(LR > x_\alpha) = \alpha + O(n^{-1})$, em que x_α é o quantil de ordem $1 - \alpha$ da distribuição χ_q^2 ([Barndorff-Nielsen & Hall, 1988](#)). Uma expressão geral para d em termos de cumulantes do logaritmo da função de verossimilhanças e suas derivadas foi desenvolvida por [Lawley \(1956\)](#).

[Silva \(2010\)](#) obteve o termo ϵ_p para os MSPNLGs como sendo

$$\epsilon_p = \epsilon_p^{(L)} + \epsilon_p^{(NL)}, \quad (2.12)$$

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

$$\epsilon_p^{(L)} = \mathbf{1}^\top Z_d Q_4 Z_d \mathbf{1} + \mathbf{1}^\top Q_1 Z^{(3)} Q_2 \mathbf{1} + \mathbf{1}^\top \widetilde{W}_1 Z^{(3)} \widetilde{W}_1 \mathbf{1} + \mathbf{1}^\top Q_1 Z_d Z Z_d Q_3 \mathbf{1} + \mathbf{1}^\top \widetilde{W}_1 Z_d Z Z_d \widetilde{W}_1 \mathbf{1},$$

e

$$\begin{aligned} \epsilon_p^{(NL)} = & \mathbf{1}^\top \left(\widetilde{W}_1 - \frac{1}{2} Q_1 \right) D Z_d \mathbf{1} + \mathbf{1}^\top (\widetilde{W}_1 - Q_1) C_d \mathbf{1} - \frac{1}{4} \mathbf{1}^\top W_1 (2B - D^2) \mathbf{1} \\ & + \frac{1}{4} \mathbf{1}^\top D W_1 Z \left[W_1 D + 4 Z_d \left(\widetilde{W}_1 - \frac{1}{2} Q_1 \right) \right] \mathbf{1} + \text{tr} \left\{ \left[\left(\widetilde{W}_1 - Q_1 \right) C - \frac{1}{2} W_1 B \right] W_1 Z \right\}, \end{aligned} \quad (2.13)$$

respectivamente, sendo $\mathbf{1}$ é um vetor de $n \times 1$ um, $Z = \widetilde{X}(\widetilde{X}^\top W \widetilde{X})^{-1} \widetilde{X}^\top$ é uma $n \times n$ matriz semi-definida positiva com elementos $z_{ij} = \widetilde{x}_i^\top K_\beta^{-1} \widetilde{x}_j$, $Z^{(3)} = Z^{(2)} \odot Z$ e $Z^{(2)} = Z \odot Z$ como sendo o produto direto, ou seja, o elemento (i, j) de $Z^{(3)}$ é z_{ij}^3 . Os elementos das matrizes quadradas B e C , de tamanho $n \times n$, são definidos como $b_{ij} = \text{tr}(K_{\beta\beta}^{-1} \widetilde{X}_i K_{\beta\beta}^{-1} \widetilde{X}_j)$ e $c_{ij} = \widetilde{x}_i^\top K_{\beta\beta}^{-1} \widetilde{X}_j K_{\beta\beta}^{-1} \widetilde{x}_i$, respectivamente, onde $\widetilde{x}_i^\top$ é a i -ésima linha de $\widetilde{X}$ e $\widetilde{X}_j = \partial^2 \eta_j / \partial \beta \partial \beta^\top$.

Também são definidas as matrizes diagonais $D = \text{diag}\{d_{1i}, \dots, d_{1n}\}$ com $d_{1i} = \text{tr}(K_{\beta} \tilde{\tilde{X}}_i)$. Notaremos como Z_d, B_d and C_d as matrizes diagonais formadas pelos elementos de Z, B e C , respectivamente. As matrizes $\tilde{W}_1, W_1, Q_1 = \text{diag}\{q_{11}, \dots, q_{1n}\}, Q_2 = \text{diag}\{q_{21}, \dots, q_{2n}\}, Q_3 = \text{diag}\{q_{31}, \dots, q_{3n}\}$ e $Q_4 = \text{diag}\{q_{41}, \dots, q_{4n}\}$ são diagonais de tamanho $n \times n$, cujos elementos são, $\tilde{w}_{1i}, w_{1i}, q_{1i}, q_{2i}, q_{3i}$ e q_{4i} são dados por

$$\begin{aligned} q_{1i} &= w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2}, \\ q_{2i} &= \frac{1}{6} \left(w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right) - \tilde{w}_{1i}, \\ q_{3i} &= \frac{1}{4} \left(w_{2i} - \frac{w_{1i} h_i^{(2)}}{(h_i^{(1)})^2} \right) - \tilde{w}_{1i} \quad \text{e} \\ q_{4i} &= \frac{1}{4} w_{3i} - \frac{3}{4} \frac{w_{2i} h_i^{(2)}}{(h_i^{(1)})^2} + \frac{3}{4} \frac{w_{1i} h_i^{(3)}}{(h_i^{(1)})^3} - \frac{5}{4} \frac{w_{1i} (h_i^{(2)})^2}{(h_i^{(1)})^4} + \frac{\tilde{w}_{1i} h_i^{(2)}}{(h_i^{(1)})^2} - \tilde{w}_{2i} + \tilde{w}_{1i}^*. \quad i = 1, \dots, n. \end{aligned}$$

A equação (2.13) poder ser interpretada como o termo devido a não linearidade na parte sistemática do modelo. Se $\eta(\cdot, \cdot)$, definida em (2.2), é linear então as quantidades d_{1j}, c_{ij} e b_{ij} cancelam-se e portanto $\epsilon_p^{(NL)} = 0$. Finalmente, considere as expressões $\tilde{\tilde{X}}_{2i} = \partial^2 \eta_i / \partial \beta_2 \partial \beta_2^\top$, para $i = 1, 2, \dots, n$ e $K_{\beta_2 \beta_2} = \tilde{X}_2^\top W \tilde{X}_2$. O cálculo do termo ϵ_{p-q} na equação (2.11) é completamente análogo a (2.12) só precisando trocar $\tilde{X}, \tilde{\tilde{X}}$ e $K_{\beta\beta}$ pelas expressões $\tilde{X}_2, \tilde{\tilde{X}}_{2i}$ e $K_{\beta_2 \beta_2}$, respectivamente. Detalhes adicionais da obtenção de quantidades envolvidas podem ser consultados no [Apêndice A](#).

2.5 Teste da Razão de Verossimilhanças Bootstrap

A técnica de bootstrap foi originalmente proposta em [Efron \(1979\)](#) num contexto de estimação em populações infinitas. O objetivo é encontrar a distribuição da estatística diretamente da amostra aleatória observada. Para a presente dissertação, serão consideradas versões bootstrap da estatística LR . O esquema bootstrap pode ser realizado através dos seguintes passos:

1. Construa B pseudo-amostras aleatórias $(\mathbf{y}^{\mathbf{A}^1}, \dots, \mathbf{y}^{\mathbf{A}^B})$, sob H_0 , da amostra original $\mathbf{y}$.
2. Calcule o valor da estatística LR para cada amostra aleatória. Ou seja, para cada pseudo amostra $\mathbf{y}^{\mathbf{A}^b}$, calcule o respectivo valor da estatística, chamada $LR^{\mathbf{A}^b}$, para $b = 1, 2, \dots, B$ como

$$LR^{\mathbf{A}^b} = 2 \left\{ \ell \left(\hat{\beta}^{\mathbf{A}^b}; \mathbf{y}^{\mathbf{A}^b} \right) - \ell \left(\tilde{\beta}^{\mathbf{A}^b}; \mathbf{y}^{\mathbf{A}^b} \right) \right\}, \quad (2.14)$$

onde $\hat{\beta}^{\mathbf{A}^b}$ e $\tilde{\beta}^{\mathbf{A}^b}$ são os estimadores de máxima verossimilhança de β , sob H_0 e H_1 respectivamente, que são obtidos da maximização de $\ell(\beta; \mathbf{y}^{\mathbf{A}^b})$. O percentil $1 - \alpha$ de $LR^{\mathbf{A}^b}$ é estimado como o valor $\hat{q}_{(1-\alpha)}$, tal que

$$\frac{1}{B} \sum_{b=1}^B I(\{LR^{\mathbf{A}^b} \leq \hat{q}_{(1-\alpha)}\}) = 1 - \alpha$$

onde $I(\cdot)$ é a função indicador. Ou seja, $I(\{LR^{\mathbf{A}^b} \leq \hat{q}_{(1-\alpha)}\}) = 1$ se $LR^{\mathbf{A}^b} \leq \hat{q}_{(1-\alpha)}$ e zero em outro caso.

3. Fixe o valor $\hat{q}_{(1-\alpha)}$ como segue: Ordene, do menor ao maior, as B estatísticas de teste bootstrap $LR^{\mathbf{A}^b}$, $b = 1, \dots, B$ previamente definidas. A replica $B \times (1 - \alpha)$ é o percentil estimado, z_α , assumindo que $B \times (1 - \alpha)$ é um inteiro. Se não for, use o seguinte procedimento: Seja $k = \lfloor (B + 1) \times \alpha \rfloor$ o maior inteiro $\leq (B + 1) \times \alpha$, assim, a quantidade $\hat{q}_{(1-\alpha)}$ é dada pelo $(B + 1 - k)$ elemento ordenado de $LR^{\mathbf{A}^b}$.

O teste da razão de verossimilhança bootstrap consiste em rejeitar H_0 se $LR > \hat{q}_{(1-\alpha)}$ para a estatística da razão de verossimilhança.

Rocke (1989) apresenta uma alternativa para o cálculo do fator de correção para LR , baseado em re-amostragem, via bootstrap. A ideia principal é estimar o valor esperado da estatística. Sejam B o número de re-amostras bootstrap e $(\mathbf{y}^{\mathbf{A}^1}, \mathbf{y}^{\mathbf{A}^2}, \dots, \mathbf{y}^{\mathbf{A}^B})$ as amostras artificiais geradas sob H_0 . Para cada pseudo-amostra $\mathbf{y}^{\mathbf{A}^b}$, $b = 1, \dots, B$ a estatística LR é calculada como na equação (2.14). A estatística LR versão bootstrap Bartlett é calculada como sendo.

$$LR_{boot}^* = q_{\overline{LR^{\mathbf{A}}}},$$

onde $\overline{LR^{\mathbf{A}}} = B^{-1} \sum_{b=1}^B LR^{\mathbf{A}^b}$.

A última abordagem é esperada ser mais eficiente computacionalmente do que a abordagem tradicional, i.e., esperamos que o correção Bartlett-bootstrap (computação numérica da correção Bartlett) produza inferências mais precisas com apenas 200 repetições, uma vez que estima a média da distribuição. Já para a primeira abordagem (estimativa do respectivo quantil crítico) espera-se um bom desempenho com pelo menos 1.000 replicações de bootstrap. Bayer & Cribari-Neto (2013) utilizaram um fator de correção de Bartlett-bootstrap no modelo de regressão Beta conduzindo a resultados competitivos com a correção de Bartlett nesses modelos.

Avaliação Numérica - Aperfeiçoamento de testes de hipóteses

Neste capítulo apresentamos simulações de Monte Carlo com o objetivo de avaliar o desempenho dos testes baseados nas estatísticas da razão de verossimilhanças (LR), razão de verossimilhanças bootstrap (LR_{boot}) e as suas versões modificadas via Bartlett (LR^* e $\widetilde{LR}$ e LR_{boot}^*) nos modelos em séries de potências não lineares generalizados (MSPNLGs).

Os desempenhos são avaliados em função da proximidade da probabilidade de rejeição sob a hipótese nula sendo esta verdadeira (Probabilidade de erro tipo I) aos seus respectivos níveis nominais. Também, avaliamos o poder de cada um dos testes propostos e a influência do número de parâmetros de perturbação no desempenho dos testes. Vale salientar que foram considerados diversos cenários, variando número de parâmetros de perturbação e tamanhos de amostra, para as simulações feitas. Vale salientar também que os cenários propostos nas simulações apresentadas são os mesmos apresentados em [Silva \(2010\)](#), em que foram introduzidas inicialmente as estatísticas de teste corrigida LR^* e $\widetilde{LR}$.

O estudo de simulação foi baseado nas seguintes distribuições dos MSPNLGs para a variável resposta: Poisson generalizada (GPO), binomial negativa generalizada (BNG) e Consul. Características de tais distribuições podem ser encontrados na Tabela 2.1. Para as simulações fixamos $\phi = 1$ e $\nu = 3$ para a BNG, em $\phi = 0,2$ para a GPO e em $\phi = 1$ para a Consul.

O estudo de simulação foi desenvolvido no `0x` ([Doornik, 2006](#)), usando como rotina de

maximização o algoritmo BFGS, descrita em (Nocedal & Wright, 1999, Cap. 8) e (Luenberger, 1973, Chap 12). Fixamos em 10000 o número de réplicas de Monte Carlo, em 600 o número de réplicas bootstrap que envolve a estatística bootstrap e em 300 o número de réplicas bootstrap para a estatística que envolve correção numérica de Bartlett. Consideramos os seguintes tamanhos de amostra $n = 20, 30, 40, 50$ e os seguintes níveis nominais $\alpha = 1\%, 5\%, 10\%$.

Para cada tamanho de amostra e cada nível nominal, são calculadas as taxas de rejeição dos testes baseados nas estatísticas $LR, LR^*, \widetilde{LR}, LR_{boot}^*$, isto é, estimamos via Monte Carlo, as seguintes probabilidades, a saber: $P(LR \geq \chi_{(\alpha,q)}^2), P(LR^* \geq \chi_{(\alpha,q)}^2), P(\widetilde{LR} \geq \chi_{(\alpha,q)}^2)$ e $P(LR_{boot}^* \geq \chi_{(\alpha,q)}^2)$, respectivamente, em que $\chi_{(\alpha,q)}^2$ é o percentil $(1 - \alpha)$ da distribuição χ_q^2 . Para o teste baseado na estatística LR_{boot} , a taxa de rejeição é obtida da estimação da probabilidade $P(LR_{boot} \geq \hat{q}_{1-\alpha})$ sendo $\hat{q}_{1-\alpha}$ dado pelo elemento ordenado $(B + 1 - k)$ de LR_{boot}^* , apresentado na Seção 2.5

Foram gerados dois cenários gerais para as simulações com o objetivo de avaliar o impacto do número de parâmetros de perturbação no desempenho dos testes. No caso linear, foram considerados os preditores apresentados na Tabela 3.1. Nosso interesse aqui é testar a seguinte hipótese: $H_0 : (\beta_5, \beta_6) = (0, 0)$ vs $H_1 : (\beta_5, \beta_6) \neq (0, 0)$. A variável resposta foi gerada assumindo que $\beta_5 = \beta_6 = 0$ e $\beta_i = 0, 05$ para $i \in \{1, 2, 3, 4, 7, 8\}$. As respectivas covariadas foram tomadas como amostras aleatórias das seguintes distribuições: $U(0, 1), F(2, 5)$, Cauchy, $N(0, 1)$, t_3 , $LN(0, 1)$ (Log-Normal), χ_3^2 e $F(3, 3)$, respectivamente.

Tabela 3.1: Preditores lineares para avaliar o desempenho dos testes

1. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \beta_7 x_{i7}$	$p = 8, p - q = 6$
2. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i5} + \beta_5 x_{i5} + \beta_6 x_{i6}$	$p = 7, p - q = 5$
3. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_5 x_{i5} + \beta_6 x_{i6}$	$p = 6, p - q = 4$
4. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_5 x_{i5} + \beta_6 x_{i6}$	$p = 5, p - q = 3$
5. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_5 x_{i5} + \beta_6 x_{i6}$	$p = 4, p - q = 2$
6. $\eta_i = \beta_0 + \beta_5 x_{i5} + \beta_6 x_{i6}$	$p = 3, p - q = 1$

Para o caso não linear, foram considerados os preditores apresentados na Tabela 3.2 cuja hipótese considerada é a mesma, ou seja: $H_0 : (\beta_5, \beta_6) = (0, 0)$ vs $H_1 : (\beta_5, \beta_6) \neq (0, 0)$ em que a variável resposta foi gerada assumindo $\beta_5 = \beta_6 = 0$ e $\beta_i = 0, 05$ para $i \in \{1, 2, 3, 4, 7, 8\}$. Neste caso, as covariadas foram tomadas como amostras aleatórias das seguintes distribuições: $LN(0, 1)$ (Log-normal), $F(2, 5)$, Cauchy, χ_3^2 , Beta(2, 3), $N(0, 2)$, $\exp(1)$, $N(0, 1)$, respectivamente.

Tabela 3.2: Preditores não lineares propostos para avaliar o desempenho dos testes

1. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \exp(\beta_8 x_{i8})$	$p = 8, p - q = 6$
2. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_5 x_{i5} + \beta_6 x_{i6} + \exp(\beta_8 x_{i8})$	$p = 7, p - q = 5$
3. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_5 x_{i5} + \beta_6 x_{i6} + \exp(\beta_8 x_{i8})$	$p = 6, p - q = 4$
4. $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_5 x_{i5} + \beta_6 x_{i6} + \exp(\beta_8 x_{i8})$	$p = 5, p - q = 3$
5. $\eta_i = \beta_0 + \beta_5 x_{i5} + \beta_6 x_{i6} + \exp(\beta_8 x_{i8})$	$p = 4, p - q = 2$
6. $\eta_i = \beta_5 x_{i5} + \beta_6 x_{i6} + \exp(\beta_8 x_{i8})$	$p = 3, p - q = 1$

3.1 Cenário Linear

Na Tabela 3.3 encontram-se as taxas de rejeição dos testes baseados nas estatísticas LR , LR^* , $\widetilde{LR}$, LR_{boot} e LR_{boot}^* para os modelos lineares, considerando vários tamanhos de amostra, 2 parâmetros de interesse ($q = 2$) e 6 parâmetros de perturbação ($p - q = 6$) (modelo 1 da Tabela 3.1). Para todos os modelos adotados (BNG, Consul e GPO), o teste baseado na estatística LR é bastante liberal. Por exemplo, para $n = 20$ e $\alpha = 10\%$, temos que as taxas de rejeição do teste LR são respectivamente dadas por 15,4%, 17,8% e 15,8%. Como é de se esperar, à medida que o tamanho de amostra aumenta, as taxas de rejeição dos testes ficam mais próximas dos níveis nominais considerados. Vale ressaltar ainda que mesmo para tamanho de amostra $n = 50$, as taxas de rejeição do teste LR são 11,1%, 11,5% e 10,9%, respectivamente. Portanto, o teste LR continua sendo ligeiramente liberal quando o tamanho de amostra é grande. enquanto que os testes baseados nas estatísticas corrigidas LR^* , $\widetilde{LR}$ e LR_{boot}^* apresentam taxas de rejeição mais próximas dos níveis nominais. Por exemplo, para $n = 20$, $\alpha = 10\%$, temos que as taxas de rejeição dos testes são, respectivamente 10,8%, 9,8% e 10,4% para o modelo BNG; 9,9%, 9,9% e 9,5% para o modelo Consul e 10,8%, 10,0% e 9,8% para o modelo GPO.

Com o objetivo de analisar o impacto do número de parâmetros de perturbação no desempenho dos testes fixamos o tamanho de amostra em $n = 30$ e diferentes valores de p foram considerados. Os resultados deste estudo encontra-se na Tabela 3.4. Observamos que os testes baseados na estatística LR são liberais à medida que aumentamos o número de parâmetros de perturbação, isto é, suas taxas de rejeição tornam-se consideravelmente superiores aos níveis nominais correspondentes. Já para os demais testes, as taxas de rejeição permanecem mais estáveis. Para o modelo BNG com $p = 8$ e $\alpha = 10\%$, por exemplo, as taxas de rejeição dos testes baseados nas estatísticas LR , LR^* , $\widetilde{LR}$, LR_{boot} e LR_{boot}^* são, respectivamente de 13,3%, 10,9%, 10,5%, 11,4% e 9,3%. O impacto do número de parâmetros para o teste LR é muito marcante pois apresentou taxas de rejeição bem acima dos níveis nominais tornando-o severamente liberal. Entre os testes corrigidos, esse impacto é notavelmente menos marcante

nos testes baseados nas estatísticas LR^* , $\widetilde{LR}$ e LR_{boot}^* pois as taxas de rejeição destes testes permanecem mais estáveis.

Os resultados apresentados na Tabela 3.5 foram obtidos levando em consideração a hipótese alternativa $H_1 : \beta_5 = \beta_6 \neq 0$ para $n = 30$, $p = 4$, $\alpha = 10\%$ e diferentes valores de $\beta_5 = \beta_6 = \beta^{(0)}$ com $\beta^{(0)}$ variando de 0,05 a 0,5. Como observamos que o teste LR é liberal, as simulações foram obtidas com os valores críticos estimados, ou seja com os quantis da distribuição empírica de LR , em vez dos valores tabulados para que todos os testes tivessem o mesmo tamanho. Os resultados indicam que não há nenhuma perda de poder derivada do fato de usar os fatores de correção de Bartlett calculados teoricamente e numericamente.

Nas Figuras 3.1, 3.2 e 3.3 são apresentados gráficos da distorção (diferença entre a nível nominal e a taxa de rejeição obtida) dos testes considerados para os modelos lineares BNG, Consul e GPO, respectivamente, com $n = 30$ e vários níveis de significância, quando aumentarmos o número de parâmetros de regressão desde $p = 3$ até $p = 8$. Os resultados sugerem que, como já foi dito, o teste LR tende a ser liberal quando p aumenta. Observa-se também que os testes LR^* , $\widetilde{LR}$ e LR_{boot}^* apresentam taxas de rejeição mais próximas dos respectivos níveis nominais do que o teste LR original, tornando-os em testes alternativos mais acurados do que o teste LR . Dentre os testes baseados nas estatísticas sem correção, em geral, o teste LR_{boot} apresentou as taxas de rejeição mais próximas dos níveis nominais do que o teste LR para os diversos casos do cenário dos preditores lineares.

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

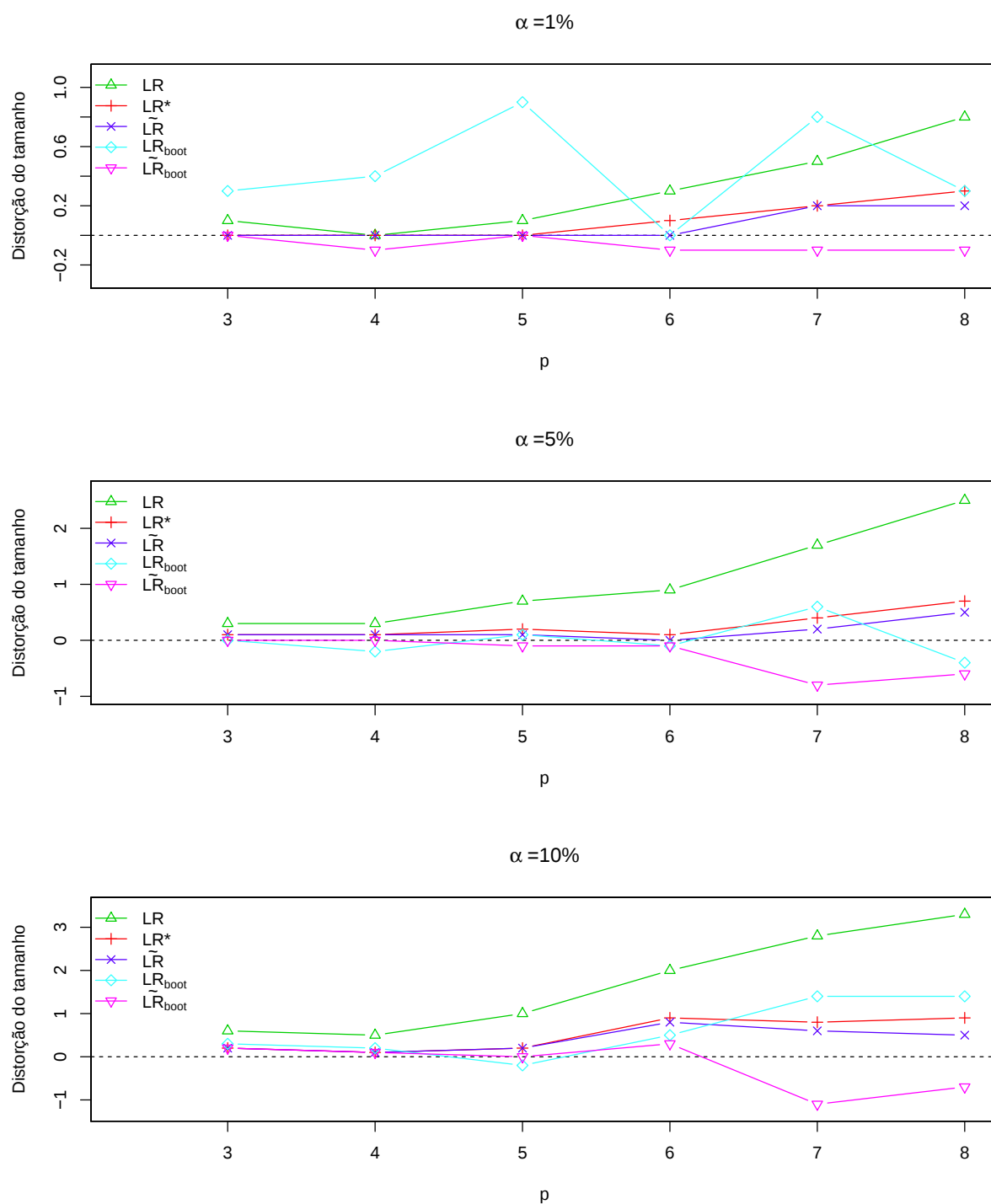
Tabela 3.3: Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ para as estatísticas LR , LR^* , $\widehat{LR}$, LR_{boot} e LR_{boot}^* para modelos lineares com $p = 8$ e vários valores de n .

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

Tabela 3.4: Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ para os testes $LR, LR^*, \widehat{LR}, LR_{boot}$ e LR_{boot}^* em modelos lineares com $n = 30$ e vários valores de p .

$\beta^{(0)}$	Modelo BNG						Modelo Consul					
	LR	LR*	LR	LR _{boot}	LR*	LR _{boot}	LR	LR*	LR	LR _{boot}	LR*	LR _{boot}
0,05	14,6	14,1	14,1	14,1	14,2	14,2	12,6	11,9	11,9	12,0	11,9	11,9
0,1	31,2	30,6	30,6	30,6	30,8	30,8	20,6	19,8	19,8	19,9	19,8	19,8
0,15	59,3	58,8	58,8	58,8	58,9	58,9	36,8	35,9	35,8	35,8	35,8	35,8
0,2	85,2	85,0	85,0	84,9	85,0	85,0	57,7	56,9	56,9	56,9	56,9	56,9
0,25	97,0	96,9	96,9	97,0	96,9	96,9	76,4	75,8	75,8	75,9	75,9	75,9
0,3	99,7	99,7	99,7	99,5	99,6	99,6	89,1	88,7	88,7	88,7	88,7	88,7
0,35	100,0	100,0	100,0	100,0	100,0	100,0	95,8	98,6	98,6	98,6	96,6	96,6
0,4	100,0	100,0	100,0	100,0	100,0	100,0	98,3	98,3	98,3	98,3	98,3	98,3
0,45	100,0	100,0	100,0	100,0	100,0	100,0	99,5	99,5	99,5	99,5	99,5	99,5
0,5	100,0	100,0	100,0	100,0	100,0	100,0	99,8	99,8	99,8	99,8	99,8	99,8

Tabela 3.5: Poder dos testes $LR, LR^*, \widetilde{LR}, LR_{boot}$ e LR_{boot}^* com $n = 30, p = 4$ e $\alpha = 5\%$ no caso linear


 Figura 3.1: Distorção de tamanho dos testes no modelo linear BNG, com $n = 30$.

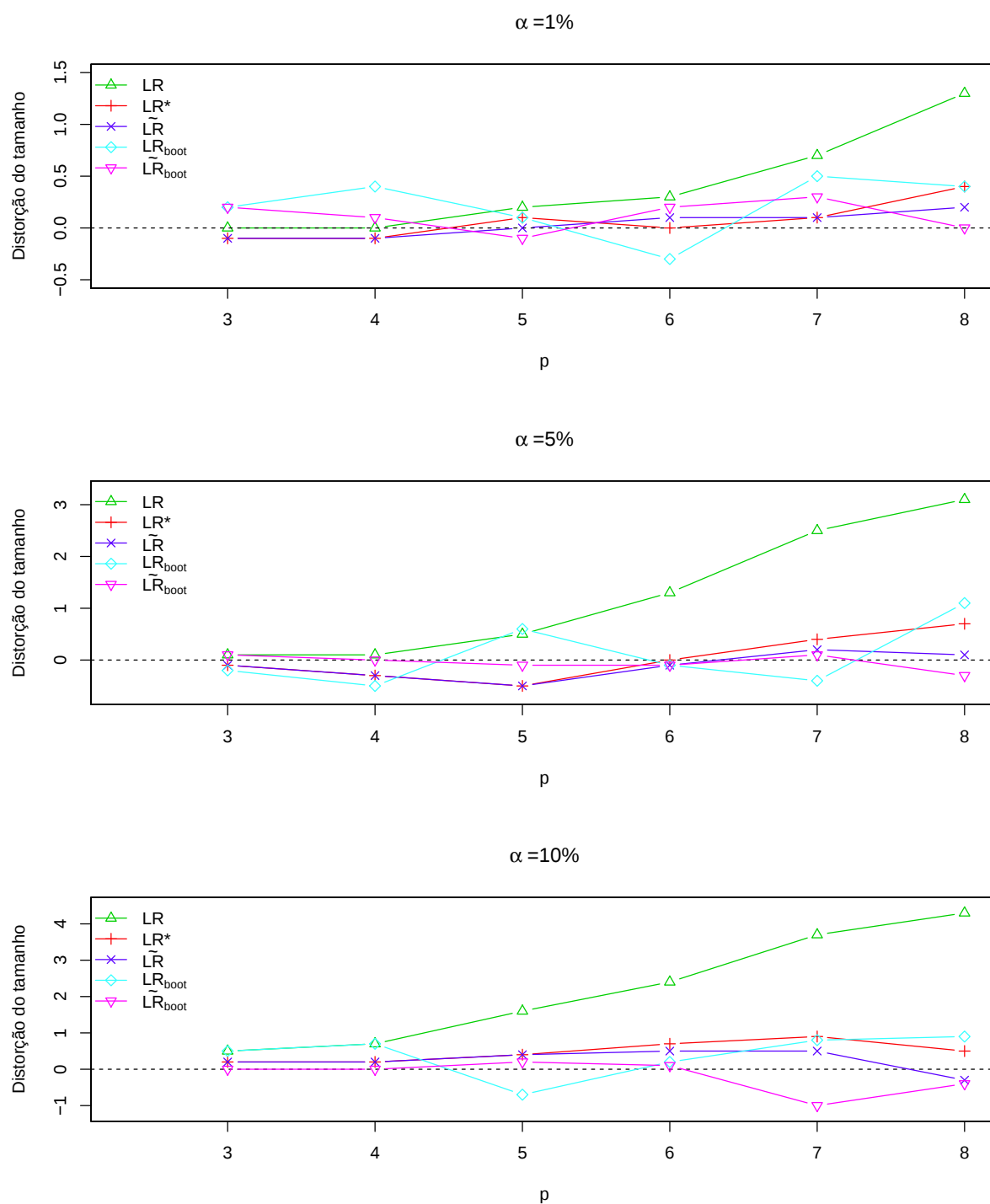


Figura 3.2: Distorção de tamanho dos testes no modelo linear Consul, com $n = 30$.

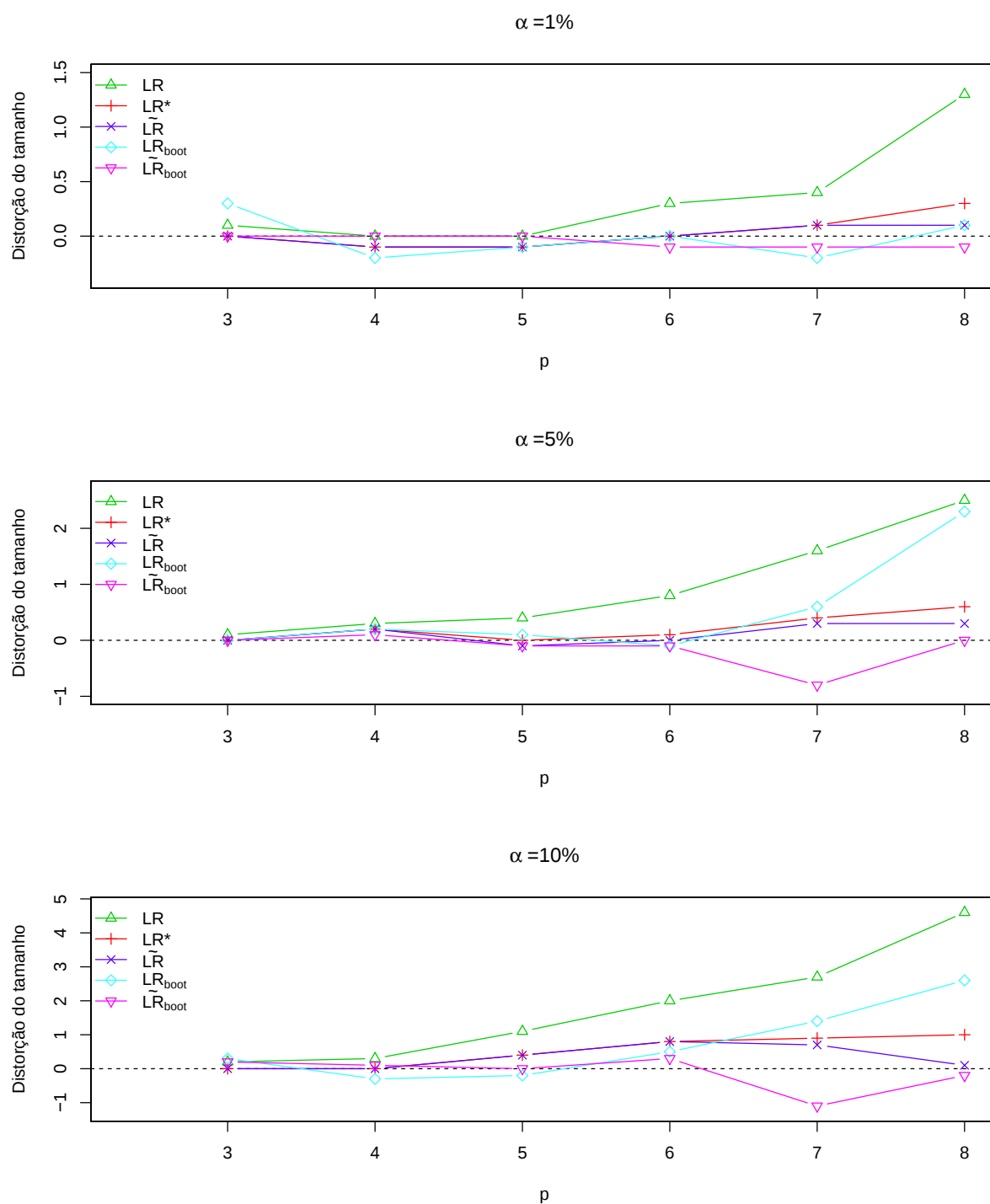


Figura 3.3: Distorção de tamanho dos testes no modelo linear GPO, com $n = 30$.

3.2 Cenário Não linear

Na Tabela 3.6 encontram-se as taxas de rejeição dos testes baseados nas estatísticas LR , LR^* , $\widetilde{LR}$, LR_{boot} e LR_{boot}^* para os modelos não lineares, considerando vários tamanhos de amostra, 2 parâmetros de interesse ($q = 2$) e 6 parâmetros de perturbação ($p - q = 6$) (modelo 1 da Tabela 3.2). Para todos os modelos adotados (BNG, Consul e GPO), o teste baseado na estatística LR é bastante liberal. Por exemplo, para $n = 30$ e $\alpha = 5\%$, temos que as taxas de rejeição são respectivamente dadas por 6,4%, 7,1% e 6,3%. Como é de se esperar, à medida que o tamanho de amostra aumenta, as taxas de rejeição dos testes ficam mais próximas dos níveis nominais considerados. Vale ressaltar ainda que mesmo para tamanho de amostra $n = 50$, as taxas de rejeição do teste LR são 5,5%, 5,9% e 5,5%, respectivamente. Portanto, o teste LR continua sendo ligeiramente liberal quando o tamanho de amostra é grande. Enquanto que os testes baseados nas estatísticas corrigidas LR^* , $\widetilde{LR}$ e LR_{boot}^* apresentam taxas de rejeição mais próximas dos níveis nominais. Por exemplo, para $n = 30$, $\alpha = 5\%$, temos que as taxas de rejeição dos testes são, respectivamente 5,1%, 5,0% e 5,0% para o modelo BNG; 4,8%, 4,8% e 4,5% para o modelo Consul e 5,0%, 5,0% e 5,0% para o modelo GPO.

Com o objetivo de analisar o impacto do número de parâmetros de perturbação no desempenho dos testes fixamos o tamanho de amostra em $n = 30$ e diferentes valores de p foram considerados. Os resultados deste estudo encontram-se na Tabela 3.7. Observamos que os testes baseados na estatística LR são liberais à medida que aumentamos o número de parâmetros de perturbação, isto é, suas taxas de rejeição tornam-se consideravelmente superiores aos níveis nominais correspondentes. Já para os demais testes, as taxas de rejeição são mais estáveis. Para o modelo Consul com $p = 8$ e $\alpha = 5\%$ por exemplo, as taxas de rejeição dos testes baseados nas estatísticas LR , LR^* , $\widetilde{LR}$, LR_{boot} e LR_{boot}^* são, respectivamente de 7,9%, 5,0%, 4,6%, 5,8% e 4,5%. O impacto do número de parâmetros para o teste LR é muito marcante pois apresentou taxas de rejeição bem acima dos níveis nominais tornando-o severamente liberal. Entre os testes corrigidos, esse impacto é notavelmente menos marcante nos testes baseados nas estatísticas LR^* , $\widetilde{LR}$ e LR_{boot}^* pois as taxas de rejeição destes testes permanecem mais estáveis.

Os resultados apresentados na Tabela 3.8 foram obtidos levando em consideração a hipótese alternativa $H_1 : \beta_5 = \beta_6 \neq 0$ para $n = 30$, $p = 4$, $\alpha = 10\%$ e diferentes valores de $\beta_5 = \beta_6 = \beta^{(0)}$ com $\beta^{(0)}$ variando de 0,05 a 0,5. Como observamos que o teste LR é liberal, as simulações foram obtidas com os valores críticos estimados, ou seja com os quantis da distribuição empírica de LR , em vez dos valores tabulados para que todos os testes tivessem o mesmo tamanho. Os resultados indicam que não há nenhuma perda de poder derivada do

fato de usar os fatores de correção de Bartlett calculados teoricamente e numericamente.

Nas Figuras 3.4, 3.5 e 3.6 são apresentados gráficos da distorção (diferença entre a nível nominal e a taxa de rejeição obtida) dos testes considerados para os modelos não lineares BNG, Consul e GPO, respectivamente, com $n = 30$ e vários níveis de significância, quando aumentarmos o número de parâmetros de regressão desde $p = 3$ até $p = 8$. Os resultados sugerem que, como já foi dito, o teste LR tende a ser liberal quando p aumenta. Observa-se também que os testes LR^* , $\widetilde{LR}$ e LR_{boot}^* ficam sempre mais perto do respectivo nível nominal do que o teste original LR , tornando-os em testes alternativos mais acurados do que o teste LR . Já para o teste LR_{boot} , existiram diversos cenários ($\alpha = 5\%$ e $\alpha = 10\%$ para o modelo BNG e $\alpha = 1\%$ e $\alpha = 5\%$ para o modelo GPO) em que foi um teste com um desempenho parecido ao apresentado pelo teste original LR mas consideravelmente inferior ao desempenho apresentados nos testes baseados nas estatísticas corrigidas LR^* , $\widetilde{LR}$ e LR_{boot}^* . Para os demais casos considerados, o teste LR_{boot} teve um desempenho melhor do que o teste LR .

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

Tabela 3.6: Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ para as estatísticas LR , LR^* , $\widetilde{LR}$, LR_{boot} e LR^*_{boot} para modelos não lineares com $p = 8$ e vários valores de n .

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

Tabela 3.7: Taxas de rejeição de $H_0 : \beta_5 = \beta_6 = 0$ para os testes $LR, LR^*, \widetilde{LR}, LR_{boot}$ e LR_{boot}^* em modelos não lineares com $n = 30$ e vários valores de p .

$\beta^{(0)}$	Modelo BNG						Modelo Consul						Modelo GPO					
	LR			LR*			LR			LR*			LR			LR*		
	LR	LR*	LR	LR	LR*	LR _{boot}	LR	LR*	LR _{boot}	LR	LR*	LR _{boot}	LR	LR*	LR	LR	LR _{boot}	LR _{boot} *
0,05	12,0	11,8	11,8	12,1	12,0	12,0	10,9	10,3	10,2	11,2	10,1	11,8	11,6	11,6	11,6	11,5	11,5	10,7
0,1	20,2	19,9	19,9	19,3	19,0	19,0	14,8	14,1	14,1	14,7	13,9	18,0	17,7	17,7	17,7	16,5	16,5	17,8
0,15	34,9	34,5	34,5	35,0	34,1	34,1	21,7	20,8	20,8	22,3	20,3	28,9	28,7	28,7	28,7	27,3	27,3	28,7
0,2	53,7	53,3	53,3	53,4	53,3	53,3	30,7	29,7	29,6	28,9	29,4	43,7	43,7	43,7	43,7	43,1	43,1	43,5
0,25	71,9	71,7	71,6	71,9	71,1	71,1	42,6	41,5	41,5	41,5	41,1	60,5	60,4	60,4	60,4	61,0	61,0	60,2
0,3	86,1	86,0	86,0	85,6	85,7	85,7	55,8	54,7	54,6	53,7	54,0	74,7	74,8	74,8	74,8	74,4	74,4	74,3
0,35	94,2	94,1	94,1	94,0	94,1	94,1	68,1	67,3	67,3	67,4	66,6	86,2	86,2	86,2	86,2	85,6	85,6	85,9
0,4	98,3	98,3	98,3	98,0	98,2	98,2	79,5	78,9	78,9	78,7	78,0	93,3	93,4	93,4	93,4	93,1	93,1	93,2
0,45	99,5	99,5	99,5	99,3	99,4	99,4	85,4	84,6	84,6	84,9	84,6	96,3	96,2	96,2	96,2	96,1	96,1	96,2
0,5	99,9	99,9	99,9	99,9	99,8	99,8	91,8	91,3	91,3	91,4	91,2	98,5	98,4	98,4	98,4	98,4	98,4	98,4

Tabela 3.8: Poder dos testes $LR, LR^*, \widetilde{LR}, LR_{boot}$ e LR_{boot}^* com $n = 30, p = 4$ e $\alpha = 10\%$ no caso não linear

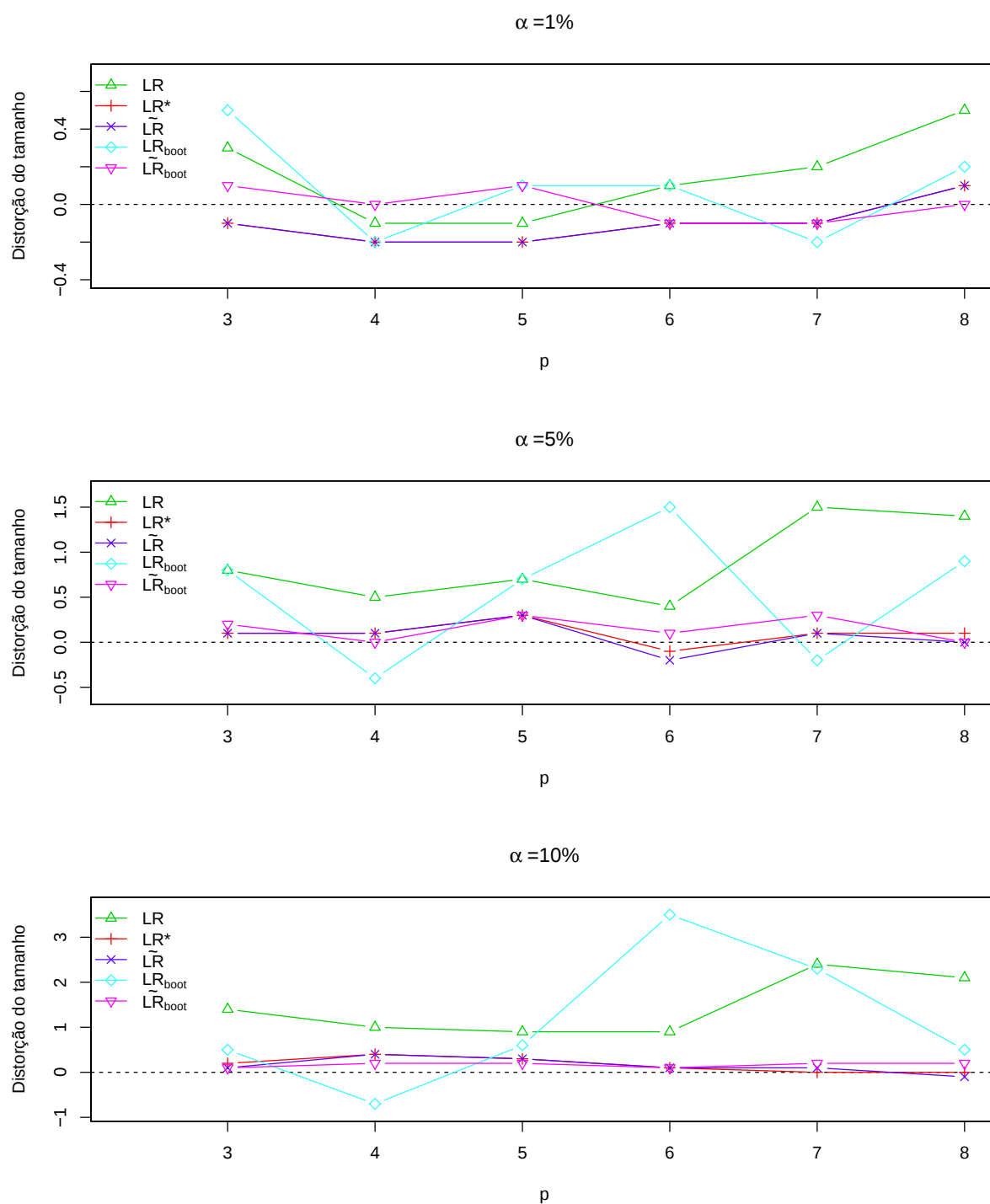


Figura 3.4: Distorção de tamanho dos testes no modelo não linear BNG, com $n = 30$.

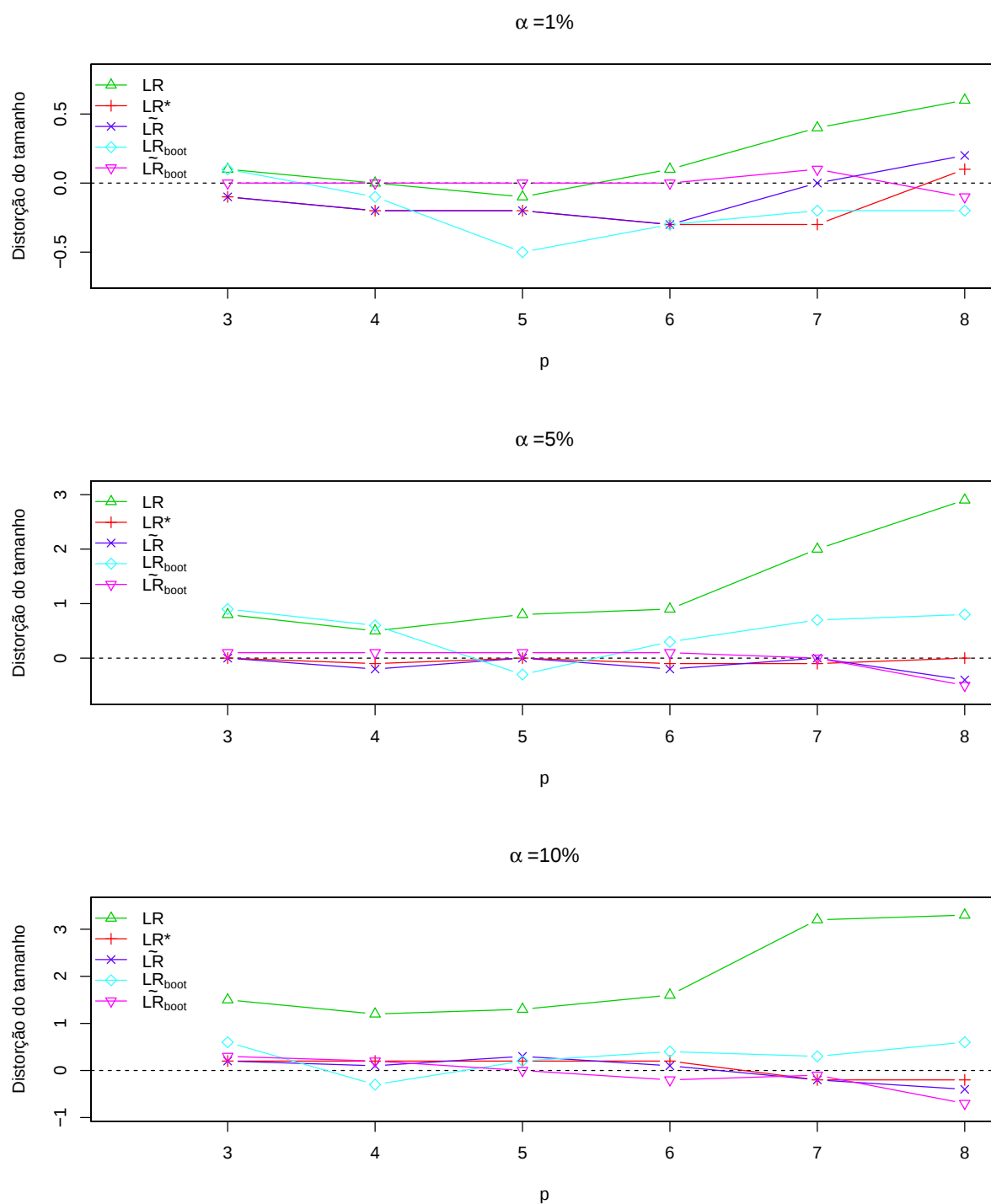


Figura 3.5: Distorção de tamanho dos testes no modelo não linear Consul, com $n = 30$.

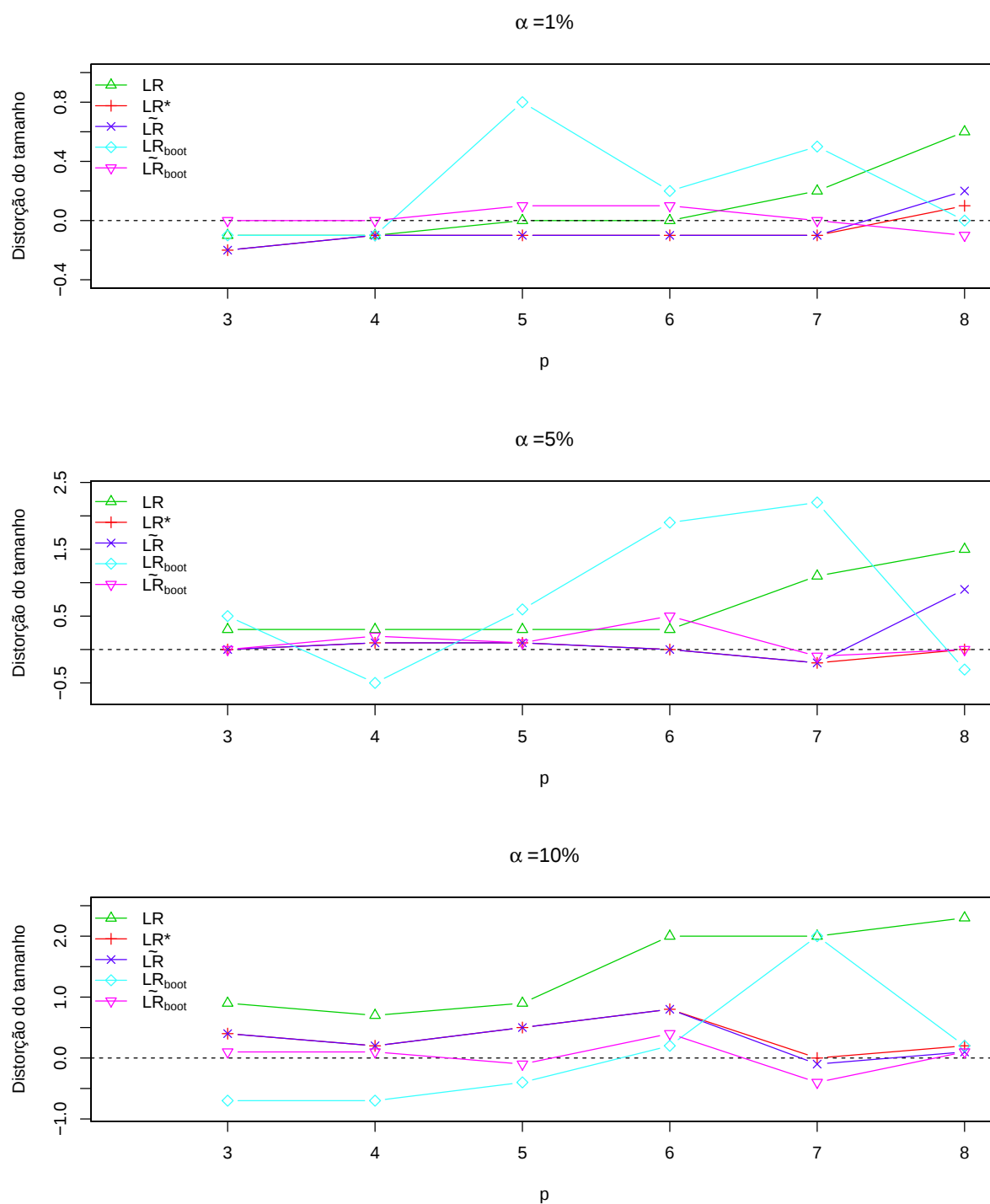


Figura 3.6: Distorção de tamanho dos testes no modelo não linear GPO, com $n = 30$.

Técnicas de diagnóstico nos MSPNLGs

Uma das etapas mais importantes na hora de ajustar um modelo estatístico, baseado em dados reais, é a verificação de certas suposições feitas acerca do modelo para garantir credibilidade nas suas conclusões e/ou previsões. Destaca-se nesta etapa, que objetiva verificar possíveis afastamentos de tais supostos, a existência de observações atípicas ou influentes que possam causar interferência desproporcional nos resultados. Tal etapa, conhecida como *análise de diagnóstico*, é fundamental para validar as conclusões que possam ser obtidas usando o modelo estatístico empregado.

Os primeiros estudos feitos nesta área correspondem à análise de resíduos como forma de avaliar o ajuste do modelo. Estudos baseados nos resíduos são importantes dado que é preciso avaliar as suposições dos erros (Cox & Snell, 1968). Os primeiros estudos na área dos resíduos estavam focados ao estudo dos modelos lineares normais e aos MLGs (Modelos lineares generalizados). Posteriormente e visando produzir resíduos mais confiáveis, foram introduzido outros resíduos como o resíduo padronizado, vide Cook & Weisberg (1982) e Belsey et al. (1980) onde é proposto um resíduo padronizado para os modelos lineares e Pregibon (1981), onde é introduzido o resíduo de componente de desvio. McCullagh (1987) apresenta o resíduo de componente de desvio padronizado. Para avaliar também os resíduos, em Atkinson (1981) foi proposta a construção de gráficos de bandas de confiança visando comparar os percentis dos resíduos com os percentis respectivo da distribuição normal padrão. Tal téc-

nica é conhecida atualmente como gráficos de envelope (QQ-plots). No presente trabalho, serão chamados de *pontos aberrantes* os pontos que foram sinalizados como afastados (fora das bandas de confiança) pelos resíduos estudados.

É importante também a obtenção de pontos que sejam influentes, ou seja, pontos que exercem um peso desproporcional nas estimativas dos parâmetros do modelo. Os primeiros passos nesta direção foram dados na década dos 70 no caso de modelo normal linear. Um dos principais resultados desta época é o estudo da diagonal principal da matriz $H = X(X^\top X)^{-1}X^\top$ onde X representa a matriz modelo, introduzido em [Hoaglin & Welsch \(1978\)](#). Posteriores estudos para identificar os pontos de alavanca (nome dado aos pontos que tivessem um peso desproporcional no próprio valor ajustado) foram dados em [St Laurent & Cook \(1992\)](#), onde estendem a definição de pontos de alavanca para modelos normais não lineares, e [Wei et al. \(1998\)](#), onde é apresentado o conceito de alavancagem generalizada que é uma metodologia bem mais geral para discriminar tais pontos. Vale salientar que a última metodologia foi deduzida para modelos cuja resposta seja contínua. Embora tenham sido obtidos resultados satisfatórios neste trabalho, a aplicação da metodologia de alavancagem generalizada só é possível nos MSPNLGs no caso em que ϕ for suposto como fixo e conhecido.

Quando se está ajustando um modelo a um conjunto de dados, é desejável que as estimativas obtidas a partir do modelo proposto sejam resistentes a pequenas perturbações. Por isso é importante que se faça um estudo sobre a robustez dos resultados obtidos em termos dos vários aspectos que envolvem a formulação do modelo e as estimativas dos seus parâmetros, ou seja, fazer uma análise de diagnóstico, que consiste, por exemplo, de métodos para avaliar o grau de sensibilidade das inferências a pequenas perturbações nos dados ou mesmo no modelo proposto.

Assim, com o intuito de detectar observações atípicas ou influentes e realizar a eliminação individual dos pontos para avaliar a mudança nas estimativas dos parâmetros com os dados retirados, é introduzida a metodologia de *influência global*, proposta em [Cook \(1977\)](#) originalmente para modelos normais e depois generalizada para outra classe de modelos ([Paula, 2003](#)). A detecção de pontos que sejam influentes, i.e., que possam ter uma influência muito forte na estimação dos parâmetros da regressão, é muito importante na hora de ajustar modelos de regressão. Mesmo sendo uma ideia natural (deleção de pontos) podem não ser percebidas certas interações entre as observações que possam ser influentes, i.e., podem não estar sendo detectadas certas interações entre as observações que influenciam na estimação. Tal efeito é conhecido, em inglês, com o nome de *masking effect*. Além disso, o custo computacional da metodologia de influência global pode ser muito alta.

Considerando tal efeito (masking effect), em [Cook \(1986\)](#) foi proposta a metodologia de *Influência Local* que propõe avaliar a influência conjunta dos pontos sob pequenas mudanças (perturbações) nos dados ou no modelo, nas estimativas de máxima verossimilhança dos parâmetros. Nesta análise de diagnóstico, assumimos que o modelo postulado é correto, e desejamos, por exemplo, comparar as estimativas obtidas através desse modelo com as respectivas estimativas decorrentes de uma pequena perturbação no mesmo. Assim, pode-se verificar a existência de pontos que sob modificações modestas causam mudanças desproporcionais nos resultados, ao invés da avaliação pela retirada individual ou conjunta de pontos.

Estudos de diagnóstico, em vários modelos de regressão, vêm sendo amplamente aplicados na literatura. Por exemplo, em [Beckman et al. \(1987\)](#) foi feito um estudo de influência local para estudar a variância em modelos mistos. Em [Galea et al. \(1997\)](#) foi feita a influência local nos modelos lineares elípticos. Em [Galea et al. \(2005\)](#) Foram desenvolvidas as metodologias de alavancagem generalizada, influência local e padronização de resíduos nos modelos não lineares simétricos. Em [Espinheira et al. \(2008\)](#) são feitos estudos de diagnóstico no modelo de regressão Beta. Em [Patriota et al. \(2010\)](#) são deduzidas técnicas de diagnóstico nos modelos normais multivariados reparametrizados. Em [Lemonte & Patriota \(2011\)](#) são aplicadas técnicas de diagnóstico na regressão não linear Birmbaum-Saunders, generalizando os resultados obtidos em [Galea et al. \(2004\)](#). Em [Borsoi et al. \(2011\)](#) foi desenvolvida a metodologia de influência local nos modelos espaciais Gaussianos lineares. Em [De Bastiani et al. \(2014\)](#) foi desenvolvido um estudo de influência local em modelos lineares espaciais elípticos.

Neste capítulo serão apresentados algumas técnicas de diagnóstico nos modelos em séries de potencia não lineares generalizados (MSPNLG). Será tratada a metodologia de alavancagem generalizada obtendo assim uma forma para discriminar a importância individual das observações. Também será feito a análise dos resíduos nesta classe de modelos. Finalmente, as metodologias de influência local e influência global, que estudam o impacto quando alguns pontos são eliminados e diversos esquemas de perturbação nos dados, respectivamente, serão aplicadas. Vale salientar que para os resultados apresentados nesta dissertação, referentes aos estudos de diagnóstico e influência, foi assumido ϕ como sendo fixo e conhecido.

4.1 Análise de Resíduos

Quando se procura ajustar um modelo a um conjunto de dados, a validação desse ajuste passa pela análise de uma estatística especial, que tem finalidade de medir a qualidade do modelo ajustado. Uma vez escolhido o modelo e realizado o ajuste, a análise de resíduos é

uma maneira eficiente de verificar a qualidade do ajuste do modelo. Esta, por sua vez, é feita para pesquisar problemas com o modelo ajustado, como por exemplo, avaliar a adequação da distribuição proposta para a variável resposta, detectar a presença de observações aberrantes (que interferem no intercepto do modelo e conseqüentemente alteram os valores ajustados), identificar a relevância de um fator adicional omitido no modelo, assim como verificar se há indícios de afastamentos sérios das suposições para a distribuição para o erro.

Por outro lado, sabemos que os resíduos são usados para identificar discrepâncias entre o modelo ajustado e o conjunto de dados logo, é conveniente procurar uma definição para o resíduo que leve em conta a contribuição de cada observação sobre a medida usada de qualidade do ajuste.

4.1.1 Resíduo Componente de Desvio

Uma medida de qualidade de ajuste bastante utilizada nos modelos lineares generalizados é a função desvio (Paula, 2003). Os resíduos construídos a partir dos componentes da função desvio (McCullagh et al., 1989) baseiam-se na distância para cada observação entre o máximo do logaritmo da função de verossimilhança do modelo saturado (com n parâmetros) e o máximo do logaritmo da função de verossimilhança do modelo investigado (com p parâmetros). A função de desvio para o caso dos MSPNLGs é definida como sendo

$$D(\mathbf{y}, \hat{\boldsymbol{\mu}}) = 2 \sum_{i=1}^n \left[y_i \log \left\{ \frac{g(y_i, \phi)}{g(\hat{\mu}_i, \phi)} \right\} + \log \left\{ \frac{f(\hat{\mu}_i, \phi)}{f(y_i, \phi)} \right\} \right].$$

O resíduo componente do desvio é definido como sendo a raiz quadrada da diferença entre os logaritmos das funções de verossimilhança sob o modelo saturado e o modelo investigado para cada observação, com o mesmo sinal $(y_i - \hat{\mu}_i)$. Assim, o resíduo no caso dos MSPNLGs fica dado por

$$t_{D_i} = \sqrt{2} \text{Sinal}(y_i - \hat{\mu}_i) \left\{ y_i \log \left\{ \frac{g(y_i, \phi)}{g(\hat{\mu}_i, \phi)} \right\} + \log \left\{ \frac{f(\hat{\mu}_i, \phi)}{f(y_i, \phi)} \right\} \right\}^{\frac{1}{2}},$$

em que $\hat{\mu}_i = h^{-1}(\hat{\eta}_i)$ é a estimativa de μ para o modelo postulado.

Com isso, pode-se dizer que o resíduo t_{D_i} representa uma distância entre a observação y_i e seu valor ajustado $\hat{\mu}_i$, medida na escala do logaritmo da função de verossimilhança.

4.1.2 Resíduo Componente de Desvio Padronizado

Com base na proposta de [De Souza & Paula \(2002\)](#), podemos usar a alavancagem generalizada para considerar uma correção nos componentes da função desvio. Assim, é obtido o resíduo componente de desvio padronizado, dado por

$$t_{D_i}^* = \frac{t_{D_i}}{\sqrt{1 - \hat{h}_{ii}}}, \quad \text{para } i = 1, 2, \dots, n,$$

em que $\hat{h}_{ii}$ corresponde com o i -ésimo elemento da diagonal da matriz $\hat{H} = \hat{W}^{\frac{1}{2}} \tilde{X} (\tilde{X}^\top \hat{W} \tilde{X})^{-1} \tilde{X}^\top \hat{W}^{\frac{1}{2}}$.

4.2 Alavancagem Generalizada

Os pontos de alavanca são pontos que elevam a variância dos valores ajustados, o que prejudica na previsão dos dados. Para a detecção desses pontos vamos estudar a influência de cada observação em seu próprio valor ajustado. Segundo [Wei et al. \(1998\)](#) a matriz $\partial \hat{y} / \partial y$ pode ser obtida em forma geral como

$$GL(\beta) = \frac{\partial \hat{y}}{\partial y} = \left\{ D_\beta (-\ddot{L}_{\beta\beta})^{-1} \ddot{L}_{y\beta} \right\} \bigg|_{\beta=\hat{\beta}}, \quad (4.1)$$

em que

$$D_\beta = \frac{\partial \mu}{\partial \beta^\top}, \quad \ddot{L}_{\beta\beta} = \frac{\partial^2 \ell(\beta)}{\partial \beta \partial \beta^\top} \quad \text{e} \quad \ddot{L}_{y\beta} = \frac{\partial^2 \ell(\beta)}{\partial \beta \partial y^\top}.$$

O (r, s) -ésimo elemento da matriz de informação observada é expresso como

$$\frac{\partial^2 \ell(\beta; y)}{\partial \beta_r \partial \beta_s} = \sum_i \{d_{1i} \tilde{x}_{is} \tilde{x}_{ir} + d_{0i} \tilde{x}_{irs}\}$$

em que $d_{0i} = y_i t_i - q_i$, $d_{1i} = \frac{y_i t'_i - q'_i}{h'_i}$. Fazendo $T^{(1)} = \text{diag} \{t'_1, \dots, t'_n\}$ e $Q^{(1)} = (q'_1, \dots, q'_n)^\top$. Temos em notação matricial

$$\ddot{L}_{\beta\beta} = \tilde{X}^\top \left(T^{(1)} y - Q^{(1)} \right) H_1 \tilde{X} + [(Ty - Q)^\top] [\tilde{X}].$$

Em que $H_1 = (1/h'_1, \dots, 1/h'_n)$ é um vetor de $1 \times n$,

Vale salientar que $\tilde{X}$ é um *array* de dimensão $n \times p \times p$ e portanto $[\cdot][\cdot]$ representa a multiplicação de uma matriz por um array, chamado de *produto colchete* e definido em [Wei \(1998\)](#),

Bates & Watts (1988) e Seber & Wild (1989).

Além disso,

$$D_{\beta} = \frac{\partial \mu}{\partial \beta^{\top}} = \frac{\partial \mu}{\partial \eta} \frac{\partial \eta}{\partial \beta} = H\tilde{X},$$

em que $H = \text{diag}\{1/h'_1, \dots, 1/h'_n\}$ é uma matriz de $n \times n$, e o termo (r, s) da matriz $\ddot{L}_{\beta y}$ é dado por

$$\begin{aligned} \frac{\partial^2 \ell(\beta; y)}{\partial \beta_r \partial y_s} &= \frac{\partial}{\partial y_s} \left(\frac{\partial}{\partial \beta_r} \sum_{i=1}^n [\log\{a(y_i, \phi)\} + y_i \log\{g(\mu_i, \phi)\} - \log\{f(\mu_i, \phi)\}] \right) \\ &= \frac{\partial}{\partial y_s} \left(\sum_{i=1}^n y_i \frac{1}{g(\mu_i, \phi)} \frac{\partial g(\mu_i, \phi)}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_r} - \sum_{i=1}^n \frac{1}{f(\mu_i, \phi)} \frac{\partial f(\mu_i, \phi)}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_r} \right) \\ &= \frac{\partial}{\partial y_s} \left(\sum_{i=1}^n y_i \frac{1}{g(\mu_i, \phi)} g'(\mu_i, \phi) \frac{1}{h'(\mu_i)} \tilde{x}_{ir} - \sum_{i=1}^n \frac{1}{f(\mu_i, \phi)} f'(\mu_i, \phi) \frac{1}{h'(\mu_i)} \tilde{x}_{ir} \right) \\ &= \frac{1}{g(\mu_s, \phi)} g'(\mu_s, \phi) \frac{1}{h'(\mu_s)} \tilde{x}_{sr} \\ &= t(\mu_s, \phi) \tilde{x}_{sr}. \end{aligned}$$

Portanto, em notação matricial, tem-se que $\ddot{L}_{\beta y} = \tilde{X}^{\top} T$.

Substituindo as quantidades obtidas em (4.1), obtém-se

$$GL(\hat{\beta}) = \left\{ H\tilde{X} \left(-\tilde{X}^{\top} \left(T^{(1)} y - Q^{(1)} \right) H_1 \tilde{X} - [(Ty - Q)^{\top}] [\tilde{X}] \right)^{-1} \tilde{X}^{\top} T \right\}.$$

4.3 Influência Global

É importante detectar pontos que possam causar uma influência desproporcional nas estimativas de máxima verossimilhança dos parâmetros obtidos em um certo modelo proposto. Uma ideia natural na hora de tentar avaliar tal influência dos pontos, é tirar da amostra os pontos a serem examinados e verificar a diferença entre as estimativas de máxima verossimilhança baseada em todos os pontos e baseada nos pontos sem a observação que está sendo avaliada. Tal ideia foi apresentada em Cook (1977) e foi chamada de *Influência Global*.

Seja $\ell(\beta; y)$ o logaritmo da função de verossimilhança do vetor de parâmetros β . Para avaliar a sensibilidade do vetor paramétrico β quando a i -ésima observação é retirada, usa-se a medida de influência LD_i , que mede o afastamento da verossimilhança, permitindo detectar

possíveis pontos influentes. Define-se LD_i como sendo

$$LD_i = 2 \left\{ \ell(\hat{\beta}; \mathbf{y}) - \ell(\hat{\beta}_{(i)}; \mathbf{y}) \right\}, \quad (4.2)$$

em que $\hat{\beta}_{(i)}$ corresponde à estimativa $\hat{\beta}$ quando a i -ésima observação é excluída. Assim, LD_i objetiva avaliar o impacto da retirada da observação, em $\ell(\cdot)$.

O problema aparece quando não é possível obter uma forma analítica para LD_i . É comum utilizar uma aproximação baseada expansão de $\ell(\beta, \mathbf{y})$ em série Taylor ao redor de $\hat{\beta}$. Assim, é obtido que

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

Agora, o termo $\ddot{L}_{\beta\beta}$ pode ser substituído pelo valor respectivo da informação de Fisher e o termo β pode ser substituído por $\hat{\beta}_{(i)}$, assim chega-se na expressão

$$LD_i \approx (\hat{\beta} - \hat{\beta}_{(i)})^\top \tilde{X}^\top \hat{W} \tilde{X} (\hat{\beta} - \hat{\beta}_{(i)}), \quad (4.3)$$

sendo $\hat{W} = \hat{L}\hat{V}\hat{L}$. 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 $\hat{\beta}$. Assim, temos que

$$\beta_{(i)} = \hat{\beta} + \left\{ -\ddot{L}_{\beta\beta}(\hat{\beta})^{-1} \right\} \ell_{(i)}(\hat{\beta}),$$

logo,

$$\hat{\beta}_{(i)} = \hat{\beta} - \frac{(y_i - \hat{\mu}_i)}{\hat{V}_i^{1/2}(1 - \hat{h}_{ii})} \left(\tilde{X}^\top \hat{W} \tilde{X} \right)^{-1} \tilde{x}_i \tilde{w}_i^{1/2}, \quad (4.4)$$

sendo $\hat{h}_{ii}$ representa o i -ésimo elemento da diagonal da matriz $\hat{H} = \hat{W}^{1/2} \tilde{X} (\tilde{X}^\top \hat{W} \tilde{X})^{-1} \tilde{X}^\top \hat{W}^{1/2}$, ver Apêndice B.

Portanto, inserindo (4.4) em (4.3), obtemos

$$LD_i = t_{S_i}^2 \frac{\hat{h}_{ii}}{1 - \hat{h}_{ii}},$$

sendo t_{S_i} dado por

$$t_{S_i} = \frac{y_i - \hat{\mu}_i}{\sqrt{\widehat{\text{Var}}(Y_i)(1 - \hat{h}_{ii})}}, \quad i = 1, 2, \dots, n.$$

4.4 Influência Local

Quando se está ajustando um modelo a um conjunto de dados, é desejável que as estimativas obtidas a partir do modelo proposto sejam resistentes a pequenas perturbações no modelo ou nas observações. Por isso é importante que se faça um estudo sobre a robustez dos resultados obtidos, em termos de vários aspectos que envolvem a formulação do modelo e as estimativas dos parâmetros, ou seja, fazer uma análise de diagnóstico que consiste por exemplo, de métodos para avaliar o grau de sensibilidade das inferências a pequenas perturbações nos dados ou mesmo no modelo proposto.

Uma análise de influência é feita assumindo o modelo como correto e estuda-se a robustez das conclusões a perturbações nos dados ou no modelo. Uma observação é dita influente se ela produz alterações desproporcionais nos resultados da análise, quando a mesma for excluída ou submetida a uma pequena perturbação. Perturbação é qualquer modificação nas suposições do modelo ou nos dados que podem causar diferenças substanciais nos resultados da análise.

Seja $\ell(\beta; \mathbf{y})$ o logaritmo da função de verossimilhança para uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ e para um vetor de parâmetros de dimensão $(p+1) \times 1$ dada em (2.4). A metodologia de influência local baseia-se na introdução de um vetor de perturbação $\omega = (\omega_1, \dots, \omega_n)^\top$, com $\omega \in \Omega$ e $\Omega \subset \mathbb{R}^n$ para assim, poder comparar os estimadores de máxima verossimilhança a fim de determinar os efeitos das perturbações nas estimativas fornecidas pelo modelo. Assim, o logaritmo da função de verossimilhança do modelo perturbado possui então a forma $\ell(\beta_\omega)$. Portanto, seja $\hat{\beta}_\omega$ a estimativa de máxima verossimilhança do vetor de parâmetros β sob o modelo perturbado e $\hat{\beta}$ o estimado usual de máxima verossimilhança sob o modelo sem perturbação. Há um grande interesse em conhecer o tamanho da mudança entre $\hat{\beta}_\omega$ e $\hat{\beta}$.

Especificamente, o conceito de influência local está baseado na curvatura da superfície do logaritmo da função de verossimilhança da verossimilhança (Cook, 1986). Assim, considere o afastamento da verossimilhança da seguinte forma

$$LD(\omega) = 2 \left\{ \ell(\hat{\beta}; \mathbf{y}) - \ell(\hat{\beta}_\omega; \mathbf{y}) \right\}, \quad (4.5)$$

em que $LD(\omega) > 0$, $\hat{\beta}$ é a estimativa de máxima verossimilhança do modelo não perturbado e $\hat{\beta}_\omega$ é a estimativa de máxima verossimilhança do modelo perturbado. $LD(\omega)$ permite calcular a diferença das verossimilhanças avaliadas em seus estimadores ($\hat{\beta}$ e $\hat{\beta}_\omega$). Assim, valores altos de $LD(\omega)$ indicam que as estimações são altamente sensíveis à perturbação. A ideia principal

é estudar o comportamento da função $LD(\omega)$ em torno do vetor de não perturbação ω_0 , isto é, o vetor ω_0 é tal que $\ell(\hat{\beta}_{\omega_0}; \mathbf{y}) = \ell(\hat{\beta}; \mathbf{y})$.

Em particular, tem-se $LD(\omega_0) = 0$ desde que $LD(\omega) \geq 0$, e o ponto ω_0 é um mínimo absoluto da função $LD(\omega)$. Em [Cook \(1986\)](#) é estabelecido que (4.5) pode ser investigado tomando seções planas, isto é, considerando valores ao longo de uma reta em Ω na direção de um vetor unitário $\mathbf{d}$ que passa por ω_0 . Ou seja, ao longo da reta $\omega(a) = \omega_0 + a\mathbf{d}$ com $a \in \mathbb{R}$. Como resultado, é obtida uma curva plana $(a, LD(\omega(a)))$ cujo comportamento, como função de a , ao redor de $a = 0$, é estudado. Assim, a curvatura normal na direção $\mathbf{d}$ no ponto ω_0 é estudada por [Cook \(1986\)](#) como uma medida de influência e é dada por

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

em que $-\ddot{L}_{\beta\beta}$ é a matriz de informação observada de Fisher e Δ é uma matriz $p \times n$ que depende explicitamente do esquema de perturbação, com elementos dados por $\Delta_{ij} = \partial^2 \ell(\beta_\omega; \mathbf{y}) / \partial \beta_i \omega_j$ com $i = 1, \dots, p+1$ e $j = 1, \dots, n$, avaliados em $\beta = \hat{\beta}$ e $\omega = \omega_0$.

[Cook \(1986\)](#) sugere utilizar $\mathbf{d}_{\max}$ que é maior autovalor da matriz $\ddot{F} = \Delta^\top (-\ddot{L}_{\beta\beta})^{-1} \Delta$. O gráfico de $|\mathbf{d}_{\max}|$ contra a ordem das observações pode revelar quais pontos que, sob pequenas perturbações, exercem maior influência em $LD(\omega)$, na vizinhança de ω_0 . Tais observações podem ser responsáveis por alterações consideráveis nas estimativas dos parâmetros sob pequenas perturbações nos dados ou no modelo.

[Escobar & Meeker Jr \(1992\)](#) sugerem tomar como medida de influência os elementos da diagonal principal da matriz $\ddot{F}$, enquanto que, [Lesaffre & Verbeke \(1998\)](#) propõem que a curvatura normal seja avaliada na direção da i -ésima observação, ou seja, avaliar $C_d(\beta)$ no vetor $\mathbf{d}_i$ de dimensão $n \times 1$, composto por um na i -ésima linha e zero nas demais. Tal curvatura é denotada por C_i em que $C_i = 2|f_{ii}|$ sendo f_{ii} os elementos da diagonal principal da matriz $\ddot{F}$. Um gráfico de C_i versus índice é sugerido, em que observações com $C_i > 2\bar{C}$ em que $\bar{C} = \sum_{j=1}^n C_j / n$, merecem atenção especial.

Na presente dissertação apresentamos três tipos de esquemas de perturbação, considerando os MSPNLGs definidos em (2.1) e (2.2) e cujo logaritmo da função de verossimilhança é dada por (2.4). Foram considerados os cenários de perturbação aditiva na resposta, perturbação de casos ponderados e perturbação aditiva no preditor. Para cada esquema de perturbação, obtemos a matriz Δ avaliada em $\beta = \hat{\beta}$ e $\omega = \omega_0$ apresentada sob a forma $\Delta = \partial^2 \ell(\beta_\omega; \mathbf{y}) / \partial \beta \partial \omega^\top$.

4.4.1 Perturbação Aditiva na Resposta

Considere um esquema de perturbação aditiva na variável resposta, ou seja,

$$y_{i\omega} = y_i + \omega_i s(y_i), \quad i = 1, \dots, n,$$

em que $s(y_i)$, é um fator de escala, comumente utilizado com o intuito de padronizar os componentes de ω e pode ser estimado pelo desvio padrão de y_i . Assim, o logaritmo da função de verossimilhança do modelo perturbado assume a forma

$$\ell(\boldsymbol{\beta}_\omega; \mathbf{y}) = \ell(\boldsymbol{\beta}_\omega) = \sum_{i=1}^n [\log\{a(y_{i\omega}, \phi)\} + y_{i\omega} \log\{g(\mu_i, \phi)\} - \log\{f(\mu_i, \phi)\}],$$

em que as funções $a(y_i, \phi)$, $g(\mu_i, \phi)$ e $f(\mu_i, \phi)$ são dadas em (2.4). Ao derivar $\ell(\boldsymbol{\beta}_\omega; \mathbf{y})$ em relação a β_s , tem-se que

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\beta}_\omega; \mathbf{y})}{\partial \beta_s} &= \sum_{i=1}^n \left[y_{i\omega} \frac{g'(\mu_i, \phi)}{g(\mu_i, \phi)} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_s} - \frac{f'(\mu_i, \phi)}{f(\mu_i, \phi)} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_s} \right] \\ &= \sum_{i=1}^n (y_{i\omega} t(\mu_i, \phi) - q(\mu_i, \phi)) \tilde{x}_{is} \quad s = 1, \dots, p. \end{aligned}$$

Agora, derivando com relação ao elemento ω_l , $l = 1, \dots, n$ obtemos

$$\frac{\partial^2 \ell(\boldsymbol{\beta}_\omega; \mathbf{y})}{\partial \beta_s \partial \omega_l} = s(y_l) t(\mu_l, \phi) \tilde{x}_{ls}.$$

Para a perturbação aditiva na resposta, $\omega_0 = (0, \dots, 0)^\top$, a matriz Δ avaliada em $\boldsymbol{\beta}$ e $\omega = \omega_0$ fica dado por $\hat{\Delta} = [TS\tilde{X}]^\top$ em que T e $\tilde{X}$ foram definidos no Capítulo 2. Finalmente, a matriz de tamanho $n \times n$, $S = \text{diag}\{s(y_1), \dots, s(y_n)\}$.

4.4.2 Perturbação de Casos Ponderados

Considere agora um esquema de perturbação da i -ésima observação no qual o logaritmo da função de verossimilhança do modelo perturbado seja expresso na forma

$$\ell(\boldsymbol{\beta}_\omega; \mathbf{y}) = \ell(\boldsymbol{\beta}_\omega) = \sum_{i=1}^n \omega_i \ell(\boldsymbol{\beta}; y_i) = \sum_{i=1}^n [\omega_i \log\{a(y_i, \phi)\} + y_i \omega_i \log\{g(\mu_i, \phi)\} - \omega_i \log\{f(\mu_i, \phi)\}]$$

com $0 \leq \omega_i \leq 1$. Derivando $\ell(\boldsymbol{\beta}_\omega; \mathbf{y})$ em relação a β_r , $r = 1, \dots, p$

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\beta}_\omega; \mathbf{y})}{\partial \beta_r} &= \sum_{i=1}^n \left[\omega_i y_i \frac{g'(\mu_i, \phi)}{g(\mu_i, \phi)} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_s} - \omega_i \frac{f'(\mu_i, \phi)}{f(\mu_i, \phi)} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_s} \right] \\ &= \sum_{i=1}^n [\omega_i y_i t(\mu_i, \phi) \tilde{x}_{ir} - \omega_i q(\mu_i, \phi) \tilde{x}_{ir}]. \end{aligned}$$

Assim,

$$\frac{\partial^2 \ell(\boldsymbol{\beta}_\omega; \mathbf{y})}{\partial \beta_r \partial \omega_l} = [y_l t(\mu_l, \phi) \tilde{x}_{lr} - q(\mu_l, \phi) \tilde{x}_{lr}], \quad l = 1, \dots, n.$$

Logo, para a perturbação de casos ponderados, $\omega_0 = (1, \dots, 1)^\top$, a matriz Δ avaliada em $\boldsymbol{\beta}$ e em $\omega = \omega_0$ fica dado por $\hat{\Delta} = \tilde{X}^\top (T\mathbf{y} - Q)$ em que T, Q e $\tilde{X}$ foram definidas no Capítulo 2

4.4.3 Perturbação Aditiva no Preditor

Um esquema de perturbação em uma determinada variável explicativa x_n pode ser descrita adicionando um vetor de perturbação ω ponderado por um vetor de escala s_{x_n} , que pode ser o desvio padrão da variável em questão, ou seja $x_{ip\omega} = x_{ip} + w_i s(x_n)$, $i = 1, \dots, n$. Seja $\eta_{i\omega} = \eta(x_{i\omega}, \boldsymbol{\beta})$ com $x_{i\omega} = (x_{i1}, \dots, x_{ip\omega}, \dots, x_{ip})^\top$. O logaritmo da função de verossimilhança para modelo com perturbação aditiva na p -ésima variável explicativa é expressa por

$$\ell(\boldsymbol{\beta}_\omega; \mathbf{y}) = \ell(\boldsymbol{\beta}_\omega) = \sum_{i=1}^n [\log\{a(y_i, \phi)\} + y_i \log\{g(\mu_{i\omega}, \phi)\} - \log\{f(\mu_{i\omega}, \phi)\}].$$

Derivando $\ell(\boldsymbol{\beta}_\omega; \mathbf{y})$ com relação ao β_r , obtemos

$$\frac{\partial \ell(\boldsymbol{\beta}_\omega; \mathbf{y})}{\partial \beta_r} = \sum_{i=1}^n \left[y_i \frac{g'(\mu_{i\omega}, \phi)}{g(\mu_{i\omega}, \phi)} - \frac{f'(\mu_{i\omega}, \phi)}{f(\mu_{i\omega}, \phi)} \right] \frac{1}{h'(\mu_{i\omega})} \frac{\partial \eta_{i\omega}}{\partial \beta_r}, \quad r = 1, \dots, p. \quad (4.6)$$

Derivando (4.6) em relação a ω_l , obtemos.

$$\begin{aligned} \frac{\partial^2 \ell(\boldsymbol{\beta}_\omega; \mathbf{y})}{\partial \beta_r \partial \omega_l} &= \sum_{i=1}^n \left[y_i \frac{g''(\mu_{i\omega}, \phi)g(\mu_{i\omega}, \phi) - g'(\mu_{i\omega}, \phi)g'(\mu_{i\omega}, \phi)}{(g(\mu_{i\omega}, \phi))^2} \right. \\ &\quad \left. - \frac{f''(\mu_{i\omega}, \phi)f(\mu_{i\omega}, \phi) - f'(\mu_{i\omega}, \phi)f'(\mu_{i\omega}, \phi)}{(f(\mu_{i\omega}, \phi))^2} \right] f_{1i} \left(\frac{1}{h'(\mu_{i\omega})} \right)^2 \frac{\partial \eta_{i\omega}}{\partial \beta_r} \\ &\quad + \left[y_i \frac{g'(\mu_{i\omega}, \phi)}{g(\mu_{i\omega}, \phi)} - \frac{f'(\mu_{i\omega}, \phi)}{f(\mu_{i\omega}, \phi)} \right] \left[\frac{\partial}{\partial \omega_l} \left\{ \frac{1}{h'(\mu_{i\omega})} \right\} \frac{\partial \eta_{i\omega}}{\partial \beta_r} + \frac{1}{h'(\mu_{i\omega})} f_{2ir} \right], \end{aligned}$$

sendo $f_{1i} = \frac{\partial \eta_{i\omega}}{\partial \omega_l}$ e $f_{2ir} = \frac{\partial^2 \eta_{i\omega}}{\partial \beta_r \partial \omega_l}$, com $l = 1, \dots, n$ e $r = 1, \dots, p$.

Para a perturbação aditiva no preditor, $\omega_0 = (0, \dots, 0)^\top$, a matriz Δ avaliada em $\beta = \hat{\beta}$ e $\omega = \omega_0$ ficará dependendo da forma de η . Em notação matricial fica dada por

$$\begin{aligned}\hat{\Delta} &= \tilde{X}^\top \{ \Phi(g_1(\mu_\omega, \phi)) \mathbf{y} + \Psi(f_1(\mu_\omega, \phi)) \} H^2 F_1 \\ &\quad + \tilde{X}^\top \{ \Phi(g_2(\mu_\omega, \phi)) \mathbf{y} + \Psi(f_2(\mu_\omega, \phi)) \} H^{(1)} \\ &\quad + [F_2] [\Phi(g_2(\mu_\omega, \phi)) \mathbf{y} + \Psi(f_2(\mu_\omega, \phi))] H,\end{aligned}$$

em que

$$\begin{aligned}\Phi(g_1(\mu, \phi)) &= \text{diag} \left\{ g_1(\mu_1, \phi), \dots, g_1(\mu_n, \phi) \right\}, \text{ e} \\ \Psi(f_1(\mu, \phi)) &= \left(g_1(\mu_1, \phi), \dots, g_1(\mu_n, \phi) \right)^\top,\end{aligned}$$

com $F_1 = \text{diag}\{f_{11}, \dots, f_{1n}\}$ e F_2 é um array de dimensão $n \times n \times p$ cuja i -ésima matriz é dada pela matriz F_{2i} cujo (i, r) -ésimo elemento é f_{2ir} . Além disso, temos.

$$\begin{aligned}g_1(\mu_i, \phi) &= \frac{g''(\mu_i, \phi)g(\mu_i, \phi) - g'(\mu_i, \phi)g'(\mu_i, \phi)}{(g(\mu_i, \phi))^2} \text{ e} \\ g_2(\mu_i, \phi) &= \frac{g'(\mu_i, \phi)}{g(\mu_i, \phi)}.\end{aligned}$$

Mesma notação será empregada para a função $f(\mu_i, \phi)$.

Avaliação Numérica - Técnicas de Diagnóstico

A análise de influência e diagnóstico é uma metodologia freqüentemente utilizada em modelos de regressão. A detecção de pontos que sejam aberrantes ou influentes é muito importante para evitar chegar em conclusões ou previsões erradas, pois os parâmetros que indexam esses modelos de regressão podem estar sendo afetados, de forma desmedida, por tais pontos. No capítulo 4 foram desenvolvidas técnicas de diagnóstico para os modelos em séries de potências não lineares generalizados (MSPNLGs) com o intuito de avaliar o impacto dos dados nas estimativas dos parâmetros nos cenários em que a variável resposta tenha uma distribuição pertencente aos MSPNLGs.

Assim, com a finalidade de apresentar as técnicas de diagnóstico abordadas nesta dissertação, realizamos simulações computacionais com tamanho de amostra fixo em $n = 50$. A variável resposta foi gerada seguindo as distribuições Poisson generalizada (GPO) e Consul, pertencentes aos MSPNLGs de acordo com o preditor não linear, dado por

$$h(\mu_i) = \beta_0 + \beta_1 x_{i1} + \exp\{\beta_2 x_{i2}\},$$

em que $h(\mu_i) = \log(\mu_i)$. Ou seja, foram geradas as variáveis $y_i \sim \text{GPO}(\mu_i)$ e $y_i \sim \text{Consul}(\mu_i)$ onde

$$\mu_i = \exp\{\beta_0 + \beta_1 x_{i1} + \exp\{\beta_2 x_{i2}\}\}$$

para $i = 1, \dots, 50$. Os valores iniciais dos parâmetros foram $\beta_i = 0.2$ para $i \in \{0, 1, 2\}$, onde as covariadas x_1 e x_2 foram tomadas como amostras aleatórias das distribuições $U(0, 1)$ e $N(0, 1)$, respectivamente. Foram também considerados $\phi = 0, 2$ para a GPO e $\phi = 1$ para a Consul.

A metodologia geral aqui adotada para abordar o problema da avaliação das técnicas propostas é descrito a seguir: Foram gerados 49 dados das respectivas distribuições utilizando o parâmetro μ e foi gerada uma última observação utilizando o parâmetro 2μ . Assim, os métodos desenvolvidos nesta dissertação serão efetivos na medida que o ponto #50 (ponto que foi gerado com o parâmetro de media diferente) seja sinalizado. Finalmente, a avaliação será feita mediante as gráficas correspondentes nas metodologias desenvolvidas. Tais gráficos foram feitos no R.

Nas Figuras 5.1 e 5.2 são apresentados os gráficos dos resíduos versus os índices (e os valores ajustados) para o modelo GPO, respectivamente. Vemos que, em geral, os gráficos baseados nos resíduos de componente de desvio são pouco voláteis, no sentido de não se afastar demais das bandas consideradas. Podemos verificar que o ponto #50 é candidato a pontos aberrante.

Na Figura 5.3 é apresentado o gráfico de envelope para os resíduos de componente de desvio padronizados. Pode-se observar que há 2 valores extremos que ficam fora das bandas de 95% de confiança e portanto não fornece indícios de afastamentos sérios da suposição da distribuição GPO para a variável resposta. O gráfico de alavancagem generalizada versus os índices é dado na Figura 5.4a. Pode-se observar que os pontos #29, #41, #50 são pontos notavelmente afastados, i.e., esse pontos são candidatos a serem pontos de alavanca. Na Figura 5.4b é apresentado o gráfico de influência global onde o ponto #50 foi sinalizado de forma muito marcante dada a escala do gráfico obtido.

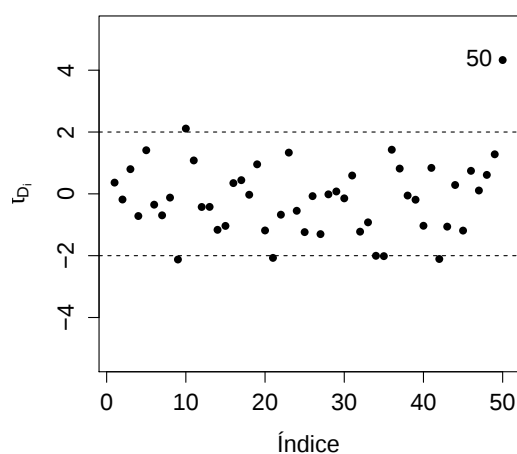
Finalmente, para o modelo GPO, os resultados de influência local sob vários esquemas de perturbação a saber: perturbação de casos ponderados, perturbação aditiva na resposta e predição aditiva no preditor, são apresentados na Figura 5.5. Os pontos #29, #41 e #50 foram considerados como influentes pelo esquema aditivo na resposta. Os pontos #3 e #50 foi considerado como influente pelo esquema de casos ponderados. Finalmente, os pontos #26 e #29 foram sinalizados pelo esquema de perturbação aditiva no preditor.

Já para o caso da distribuição Consul, os gráficos dos resíduos considerados contra o índice, reportado na Figura 5.6, sinalizam várias vezes o ponto #50. Na Figura 5.7 são apresentados os gráficos dos resíduos de componente de desvio e componente de desvio padronizado versus os seus respectivos valores ajustados onde o ponto #50 foi sinalizado novamente.

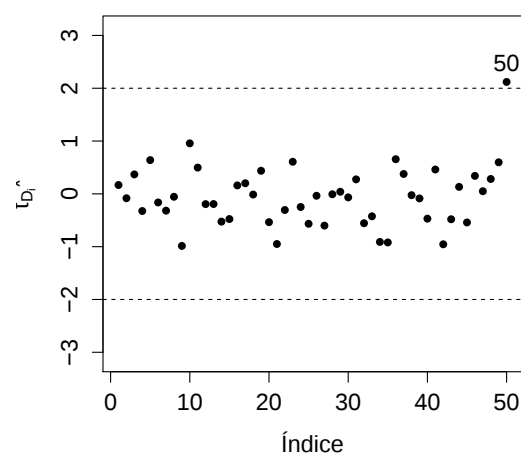
Na Figura 5.8 é apresentado o gráfico de envelope para os resíduos de componente de desvio padronizados que não fornece indícios de afastamentos sérios da suposição da distribuição GPO para a variável resposta. Na figura 5.9a são mostrados os gráficos da metodologia de alavancagem generalizada versus os índices. Os pontos #26, #29, #41 e #50 foram sinalizados como sendo pontos de alavancagem, i.e., tem uma influência considerável no intercepto do modelo. Na Figura 5.9b são apresentados os gráficos de influência global versus os índices onde foi sinalizado o ponto #50 de forma destacada.

Finalmente, na Figura 5.10 são apresentados os gráficos de influência local para os três cenários de perturbação que foram escolhidos. Os 8 pontos #28, #41, #47, #49 e #50 foram sinalizados no caso de perturbação aditiva na resposta, enquanto o ponto #50 foi o único sinalizado no esquema de casos ponderados. Finalmente, os pontos #26 e #29 foram sinalizado pelo esquema de perturbação aditiva no preditor.

Vemos que, em geral, os métodos de influência e diagnóstico tiveram êxito na hora de identificar o ponto #50 como sendo influente nos casos considerados, ou seja, os casos em que a variável foi distribuída segundo as distribuições GPO e Consul. Vale salientar que os pontos #29 e #41 foram também sinalizados por algumas das metodologias empregadas. Fato que pode ser devido a maneira que foram que as covariadas foram geradas já que os mesmos pontos foram sinalizados nos dois cenários considerados. Mesmo assim, vemos que as metodologias apresentadas corretamente apontaram ao ponto #50. Concluimos portanto, baseados nos resultados nos dados simulados, a utilidade dos métodos desenvolvidos nesta dissertação referente aos estudos de diagnóstico e influência nos MSPNLGs.

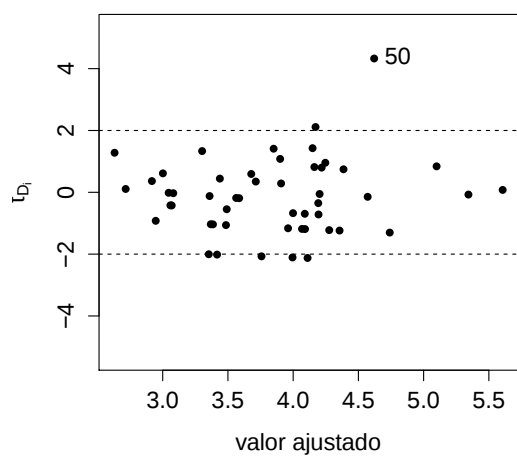


(a) Resíduos de componente de desvio

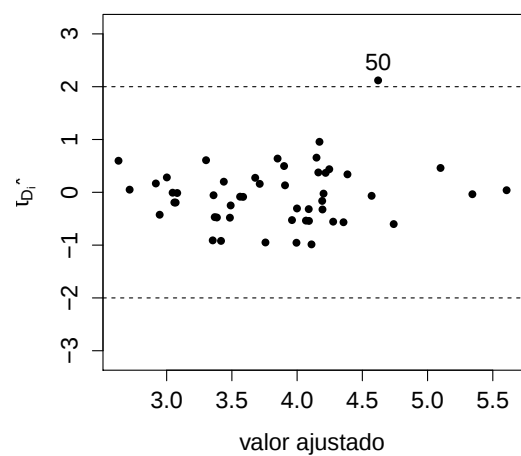


(b) Resíduos de Comp. de desvio padronizados

Figura 5.1: Gráficos dos resíduos contra os índices para o modelo GPO ajustado



(a) Resíduos de componente de desvio



(b) Resíduos de Comp. de desvio padronizados

Figura 5.2: Gráficos dos resíduos contra os valores ajustados para o modelo GPO ajustado

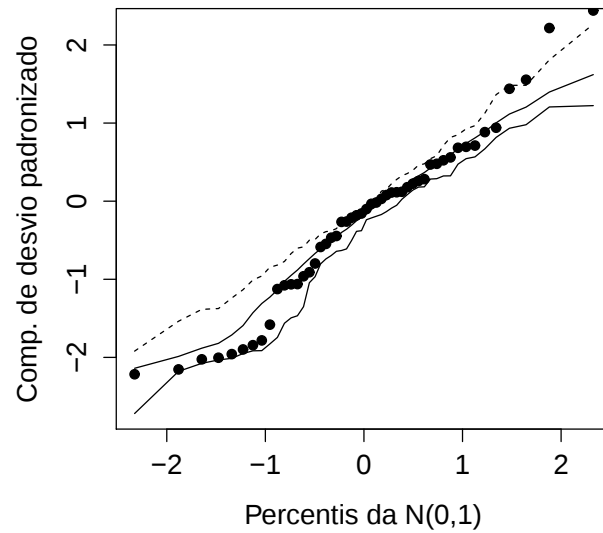


Figura 5.3: Gráfico normal de probabilidade para o caso da distribuição GPO

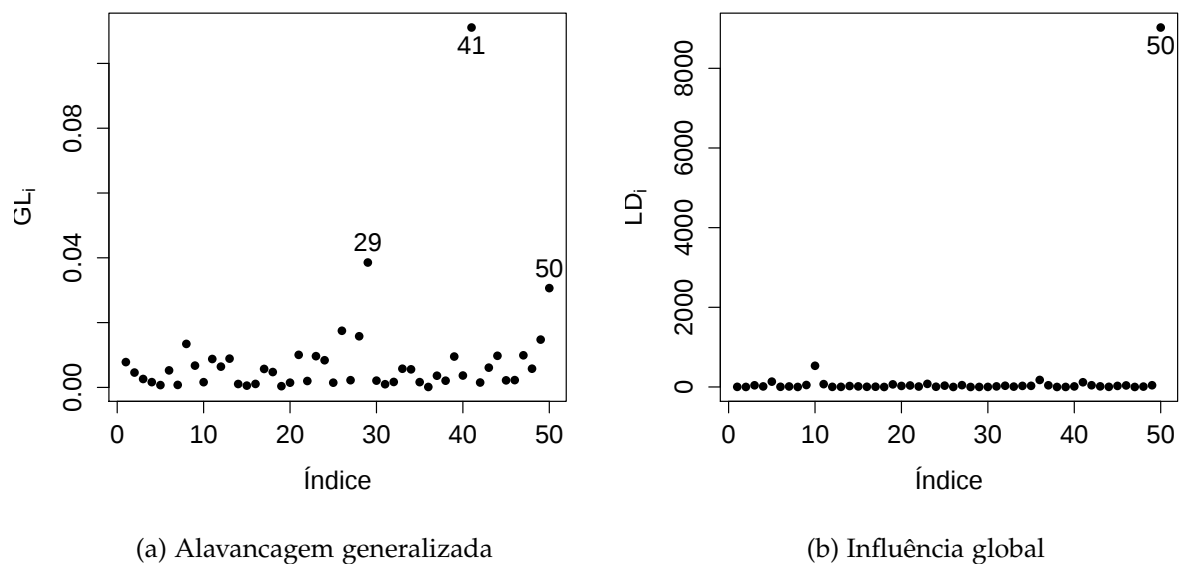
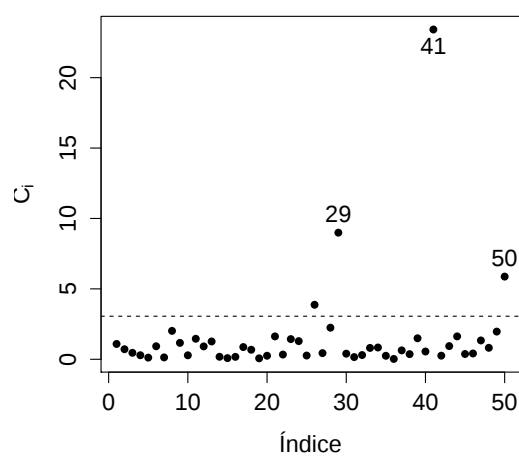
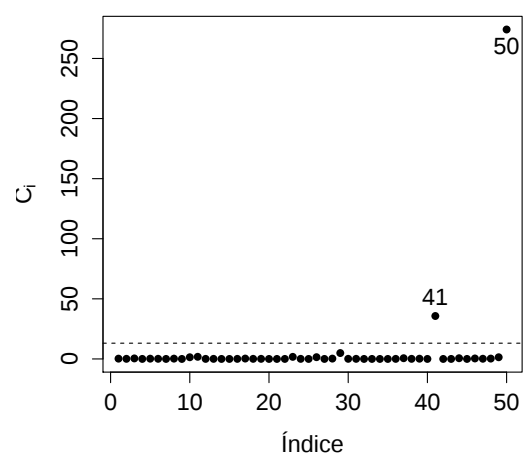


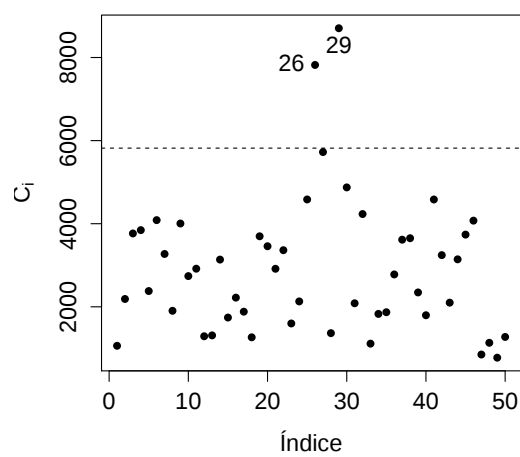
Figura 5.4: Gráficos de alavancagem generalizada e da Influência local para o caso da distribuição GPO



(a) Perturbação aditiva na resposta

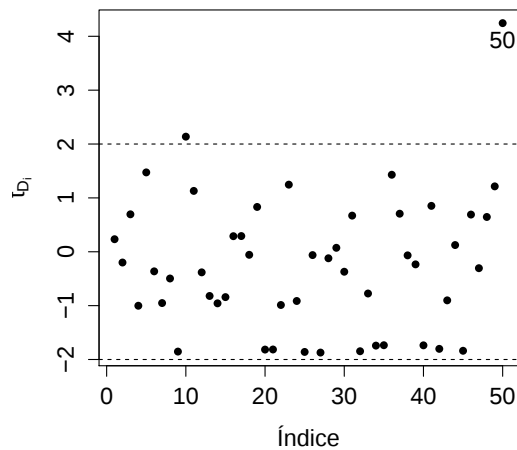


(b) Perturbação de casos ponderados

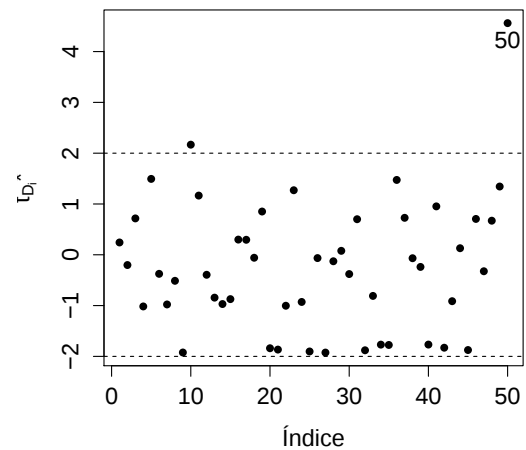


(c) Perturbação aditiva no preditor

Figura 5.5: Gráficos da Influência local para o caso da distribuição GPO

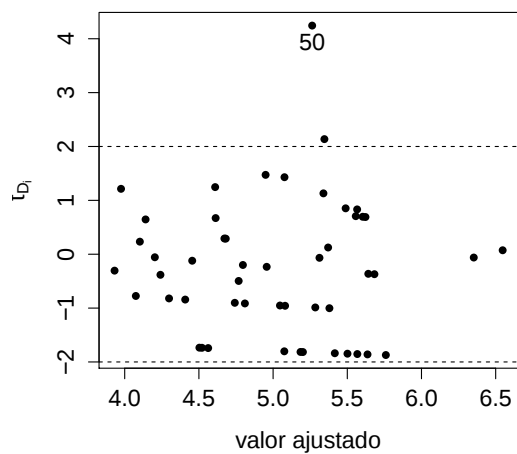


(a) Resíduos de componente de desvio

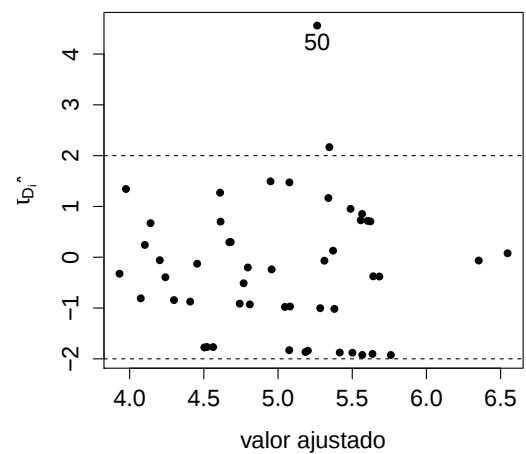


(b) Resíduos de Comp. de desvio padronizados

Figura 5.6: Gráficos dos resíduos contra os índices para o modelo Consul ajustado



(a) Resíduos de componente de desvio



(b) Resíduos de Comp. de desvio padronizado

Figura 5.7: Gráficos dos resíduos contra os valores ajustados para o modelo Consul ajustado

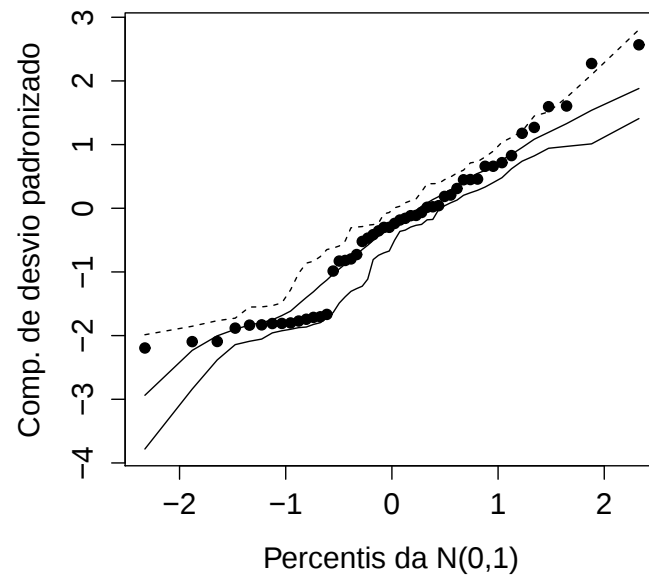


Figura 5.8: Gráfico normal de probabilidade para o caso da distribuição Consul

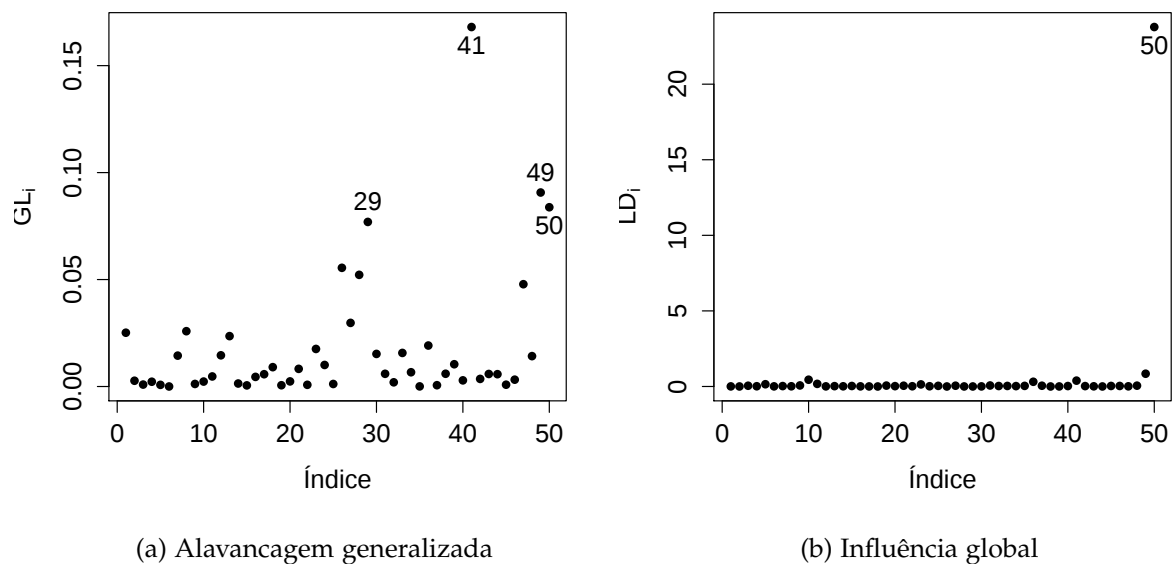
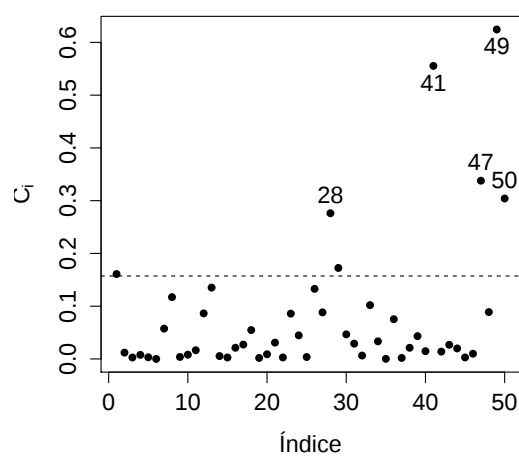
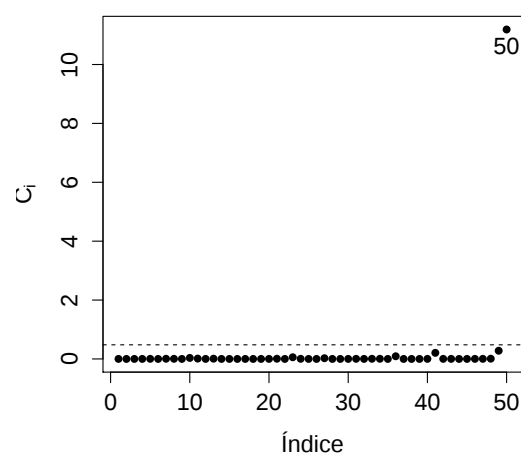


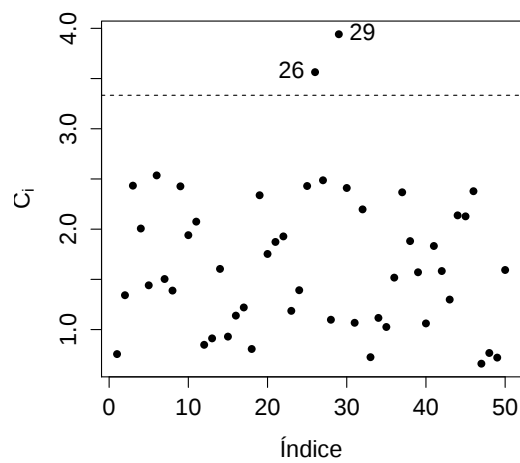
Figura 5.9: Gráficos de alavancagem generalizada e da Influência local para o caso da distribuição Consul



(a) Perturbação aditiva na resposta



(b) Perturbação de casos ponderados



(c) Perturbação aditiva no preditor

Figura 5.10: Gráficos da Influência local para o caso da distribuição Consul

Aplicação dos Métodos de Diagnóstico

Neste capítulo ilustraremos os Modelos em séries de potência não lineares generalizados (MSPNLGs) com um exemplo de dados reais e aplicaremos as técnicas de diagnóstico desenvolvidas nesta dissertação aos dados apresentados no [Apêndice C](#), correspondentes ao número de espécies de peixes em um lago (y) e o logaritmo da área do lago em km^2 (x). Esses dados foram analisados inicialmente por [Barbour & Brown \(1974\)](#), e posteriormente por [Rigby et al. \(2008\)](#), por [Cordeiro et al. \(2009\)](#) e por [Silva \(2010\)](#). Em [Cordeiro et al. \(2009\)](#) foi discutida a flexibilidade dos MSPNLGs para ajustar os dados na Tabela [C.1](#) adotando o modelo delta binomial (DB) cujo o preditor é dado por

$$\eta_i = \beta_0 + \beta_1 \log(x_i) + \beta_2 \{\log(x_i)\}^2, \quad (6.1)$$

onde $\eta_i = h(\mu_i)$, e $h(\mu_i) = \log(\mu_i - m)$ é a função de ligação empregada, onde m é o suporte da distribuição DB. Neste caso $m = 5$.

No decorrer deste trabalho foi percebido que para tais dados e para o modelo ajustado [\(6.1\)](#), os p -valores correspondentes dos parâmetros β_0 , β_1 e β_2 foram de 1.644e-11, 8.230e-01 e 5.618e-02, respectivamente. Ou seja, o parâmetro β_1 é não significativo. Portanto, nossa abordagem para a aplicação, referente as técnicas de diagnóstico e influência, será feita utilizando

o preditor

$$\eta_i = \beta_0 + \beta_1 \{\log(x_i)\}^2, \quad (6.2)$$

já que no modelo (6.2) os respectivos p -valores para os parâmetros β_0 e β_1 , foram de 3.335e-24 e 7.782e-08 e portanto, são parâmetros significativos. Assim, foi assumido que $y_i \sim DB(\mu_i)$ onde $\mu_i = m + \exp\{\eta_i\}$, com η_i dado por (6.2). Na Tabela 6.1 apresentamos os valores do ajuste dos modelos (6.1) e (6.2) ao banco de dados referente ao número de espécies de peixe dados na Tabela C.1.

Tabela 6.1: Valores do ajuste dos modelos (6.1) e (6.2)

	Parâmetros	Estimativa	Erro padrão	t -valor	p -valor
Modelo (6.2)	β_0	2.42343	0.153924	15.744	3.335e-24
	β_1	0.01634	0.002709	6.033	7.782e-08
Modelo (6.1)	β_0	2.48233	0.305867	8.1157	1.644e-11
	β_1	-0.02547	0.113429	-0.2246	8.230e-01
	β_2	0.01838	0.009456	1.9439	5.618e-02

Na Figura 6.1 encontra-se o gráfico de dispersão da área do lago (km^2) contra o número de espécies de peixes, para o modelo DB assumido para a variável resposta e o preditor adotado. Na Tabela 6.2 encontram-se as estimativas dos parâmetros β_0, β_1 junto com os seus respectivos erros padrão e p -valores. Notemos que existem algumas observações se destacando em relação as demais, a saber: #26, #32, #33, #38 e #39. As observações #26, #32 e #33 são destacadas por terem área do lago muito grande e uma grande quantidade de espécies de peixe, enquanto as observações #38 e #39 têm áreas do lago de maior tamanho mas uma menor quantidade de espécies de peixe.

Objetivando avaliar a influência dos pontos, a seguir são apresentados os resultados dos estudos de diagnóstico para o banco de dados usando o modelo (6.2). Foram aplicadas as metodologias apresentados de análise dos resíduos com resíduos de componente de desvio e resíduo de componente de desvio de componente padronizado, alavancagem generalizada, influência global e influência local.

Na Figura 6.2 e 6.3 encontram-se os gráficos dos resíduos de componente de desvio (t_{D_i}) e gráficos de componente de desvio padronizados ($t_{D_i}^*$) contra os índices para o modelo DB ajustado aos dados dos peixes da Tabela C.1. Neste caso, os pontos #36 e #66 foram sinalizados como candidatos a serem pontos aberrantes. Na Figura 6.4 é apresentado o gráfico de normal de probabilidades com envelopes para o modelo (6.2) no caso dos resíduo de componente de desvio padronizado, que não fornece indícios de afastamento sérios da suposição de distri-

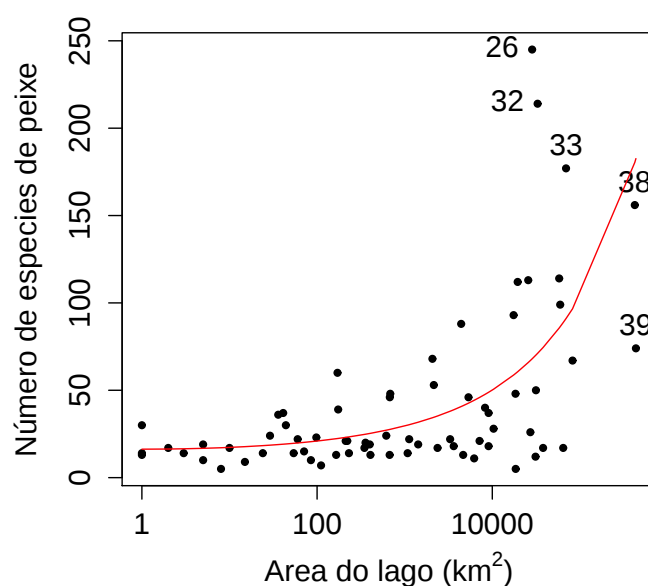


Figura 6.1: Gráfico de dispersão do número de espécies de peixe contra a área do lago

buição DB para a reposta. portanto o modelo (6.2) é um modelo adequado para descrever os dados.

Os resultados da alavancagem generalizada são reportados na Figura 6.5b. Neste caso, os pontos #38 e #39 foram considerados como pontos influentes quando gráfica-se a alavancagem contra os índices. Os resultados de metodologia de influência local, apresentados na figura 6.6 discriminam aos pontos #3 e #18 como sendo influentes, quando comparada com os índices.

Finalmente, os resultados da metodologia de influência local nos seus vários cenários (esquemas de perturbação) são apresentados na Figura 6.6. Neste esquema de avaliação da influência, os pontos #3, #26, #27, #32, #39 e #38 resultaram influentes de acordo com estas metodologias. Vale salientar que no esquema de perturbação aditiva na resposta não foram sinalizados pontos influentes.

Precisamos verificar o peso das observações destacadas nas estimativas dos parâmetros. Vemos assim que os pontos #3, #18, #26, #27, #32, #36, #38, #39 e #66 foram sinalizados por várias das metodologias empregadas. Visando avaliar a influência desses pontos, iremos calcular os estimadores de máxima verossimilhança dos parâmetros do modelo (6.2) tirando separadamente os ponto já mencionados com o intuito a de avaliar o impacto da retirada individual

das observações. Os resultados são apresentados na Tabela (6.2).

Tabela 6.2: Estimativas de máxima verossimilhança dos parâmetros do modelo DB ajustado aos dados dos peixes da Tabela C.1

Pontos retirados	β_0			β_1		
	EMV	Erro padrão	p -valor	EMV	Erro padrão	p -valor
Nenhum	2.42343	0.153924	3.335e-24	0.01634	0.002709	7.782e-08
3	2.3640	0.151600	8.124e-24	0.01680	0.00262	1.784e-08
18	2.36476	0.158208	7.237e-23	0.01713	0.002818	6.718e-08
26	2.44819	0.152231	1.612e-24	0.01512	0.002649	2.930e-07
27	2.42000	0.152559	3.320e-24	0.01602	0.002653	7.979e-08
32	2.44293	0.153796	3.092e-24	0.01542	0.002705	3.009e-07
36	2.41668	0.145911	3.376e-25	0.01673	0.002564	1.116e-08
38	2.42242	0.156419	1.162e-23	0.01639	0.002801	1.672e-07
39	2.41727	0.154787	7.519e-24	0.01654	0.002756	9.188e-08
66	2.46475	0.152915	1.430e-24	0.01584	0.002615	7.392e-08

Vemos que em geral, existem mudanças pequenas nos parâmetros quando são retirados os pontos sinalizados pelas metodologias empregadas. Vale salientar também que, mesmo retirando os pontos sinalizados, e levando em consideração a variabilidade das estimativas, os p -valores e os erros padrão ficam estáveis para os dois parâmetros considerados. Para avaliarmos o efeito ao eliminarmos uma ou mais observações sobre os estimadores de máxima verossimilhança, vamos utilizar a razão de mudança. Esta razão de mudança é expressa por

$$RM_j = \left| \frac{\hat{\beta} - \hat{\beta}_j}{\hat{\beta}} \right| \times 100, \quad j \in \{3, 18, 26, 27, 32, 36, 38, 39, 66\},$$

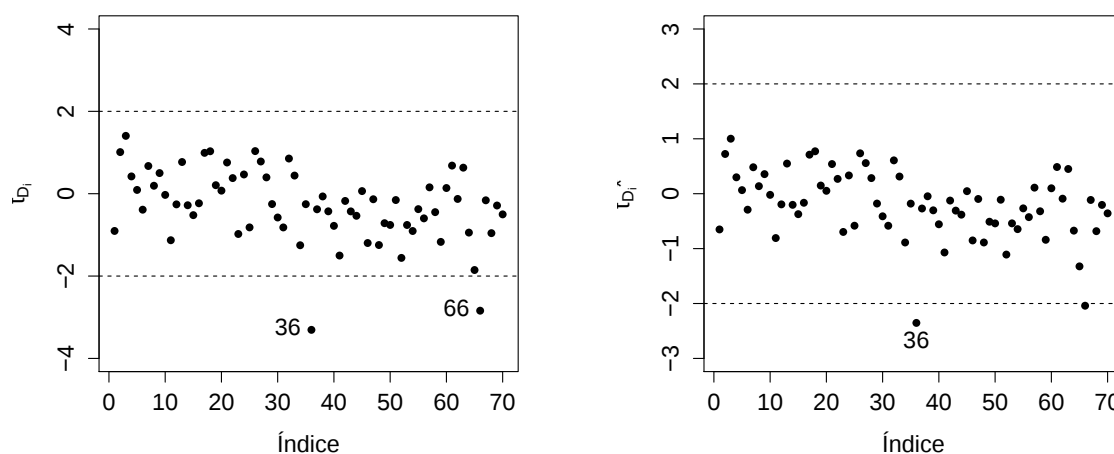
em que $\hat{\beta}_j$ é o estimador de máxima verossimilhança quando se elimina a observação j .

Na Tabela 6.3 são apresentados as razões de mudanças para as pontos considerados. Por tratar-se de um estudo que envolve 70 observações e só dois parâmetros a serem estimados, uma mudança será considerada como relevante se for maior ao 3%. Vemos que em vários casos as mudanças foram maiores do que o esse valor estabelecido. Em geral, e segundo esse valor crítico, as estimativas de β_0 foram mais robustas pois a retirada individual dos pontos não afetou tal parâmetro. Já para o caso do parâmetro β_1 , as observações #18, #26, #32 e #66 causaram mudanças significativas (maiores do que o 3%) nas estimativas de máxima verossimilhanças deste parâmetro. Portanto, deve-se chamar a atenção nas observações #18, #26, #32 e #66.

Tabela 6.3: Razão de mudanças para os pontos sinalizados

RM	$\hat{\beta}_0$	$\hat{\beta}_1$
RM_3	2.45%	2.81%
RM_{18}	2.42%	4.83%
RM_{26}	1.02%	7.46%
RM_{27}	0.14%	1.95%
RM_{32}	0.80%	5.63%
RM_{36}	0.28%	2.38%
RM_{38}	0.13%	0.30%
RM_{39}	0.25%	1.22%
RM_{66}	1.71%	3.06%

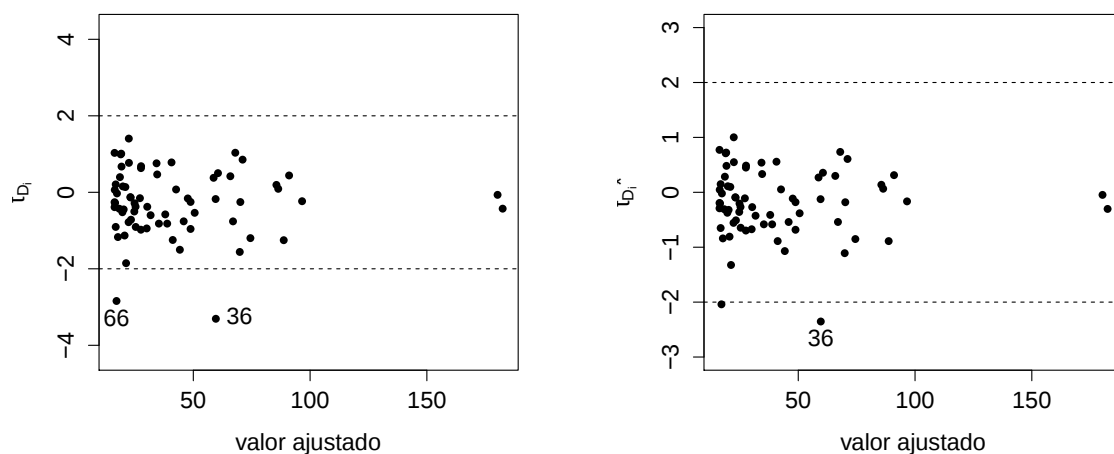
Concluímos assim que os métodos de diagnóstico e influência para os MSPNLGs, desenvolvidos nesta dissertação, se mostraram eficientes para detetar pontos que devam receber um cuidado especial por tratar-se de pontos que podem afetar de forma significativas as estimativas de máxima verossimilhança do modelo de regressão adotado.



(a) Resíduos de componente de desvio

(b) Resíduos de Comp. de desvio padronizados

Figura 6.2: Gráficos dos resíduos contra os índices o para o modelo Delta Binomial ajustado aos dados dos peixes da Tabela C.1



(a) Resíduos de componente de desvio

(b) Resíduos de Comp. de desvio padronizados

Figura 6.3: Gráficos dos resíduos contra os valores ajustados para o modelo Delta Binomial ajustado aos dados dos peixes da Tabela C.1

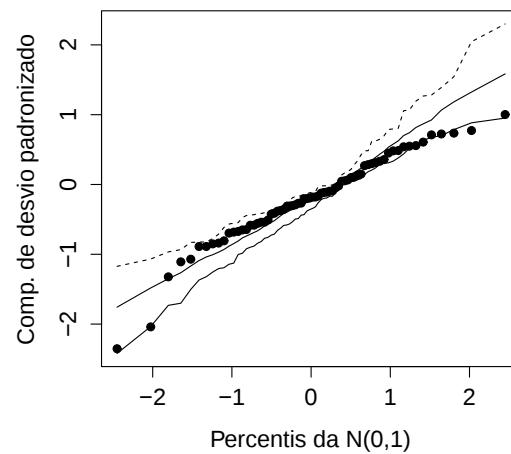
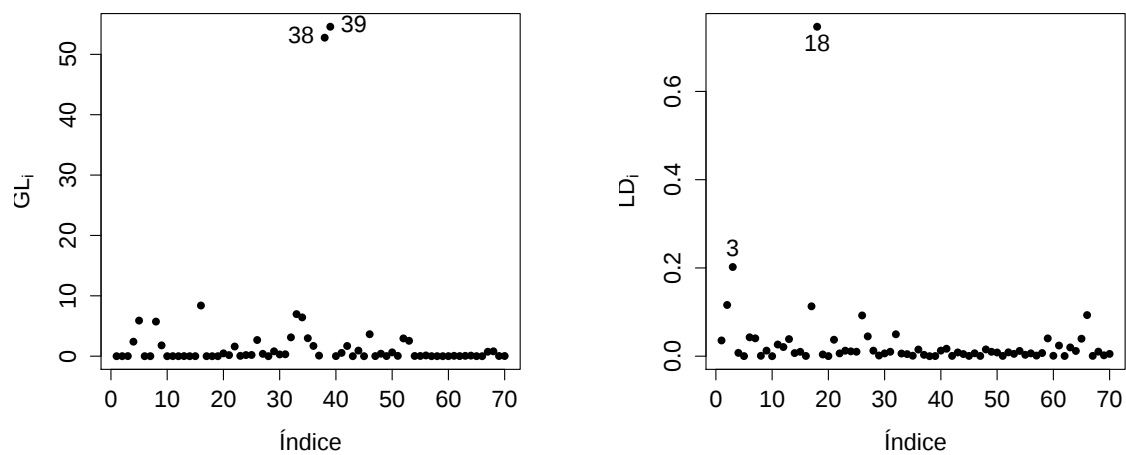


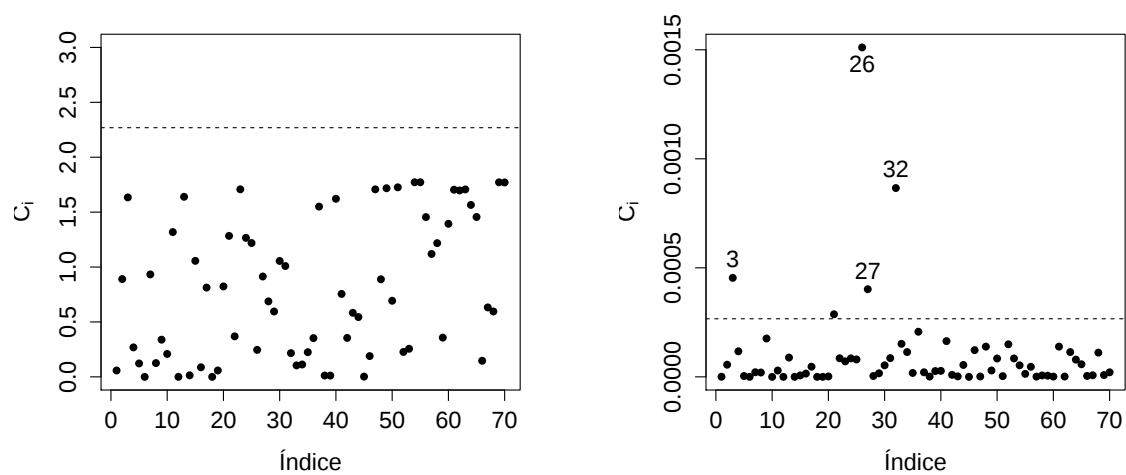
Figura 6.4: Gráfico normal de probabilidades com envelope para o resíduo de componente de desvio para o modelo DB ajustado aos dados dos peixes da Tabela C.1



(a) Alavancagem generalizada

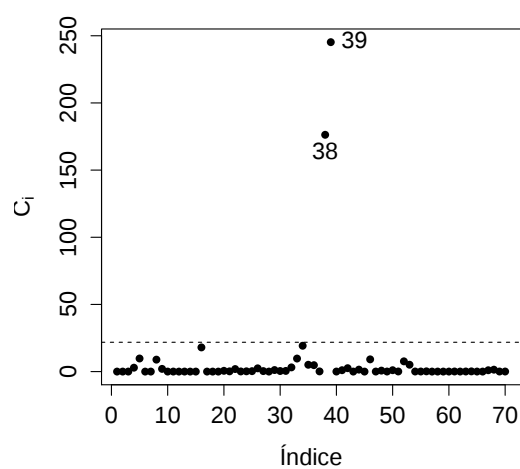
(b) Influência global

Figura 6.5: Gráficos da Influência global para o caso do banco de dados



(a) Perturbação aditiva na resposta

(b) Perturbação de casos ponderados



(c) Perturbação aditiva no preditor

Figura 6.6: Gráficos dos índices de C_i sob o modelo Delta Binomial ajustado aos dados dos peixes da Tabela C.1

Nesta dissertação foram estudados os modelos em séries de potências não lineares generalizados (MSPNLGs) recentemente propostos em [Cordeiro et al. \(2009\)](#). Tais modelos forecem uma alternativa muito atrativa para modelagem de dados discretos uma vez que os MSPNLGs contém, como submodelos particulares, alguns dos modelos de regressão discreta mais importantes e mais utilizados na área, como por exemplo: Poisson generalizada, Binomial negativa generalizada e Geeta- m .

MSPNLGs tem sido estudados anteriormente. Em [Silva \(2010\)](#) foi desenvolvida uma correção de viés de segunda ordem para os estimadores de máxima verossimilhança na classe MPSNLG, além foi implementada a técnica de correção de Barlett para a estatística da razão de verossimilhanças. Em [Lima \(2012\)](#) foram apresentadas versões corrigidas tipo-Bartlett da estatística escore nos MSPNLGs.

Podemos listar conclusões principais deste trabalho, baseado nos MSPNLGs, como segue:

- a. O estudo dos tamanhos dos testes indicou que o teste da razão de verossimilhanças original (LR) é, em geral, liberal, conduzindo a taxas de rejeição da hipótese nula com uma probabilidade maior que o nível de significância nominal do teste. Os testes da razão de verossimilhanças corrigidos através de correção de Bartlett (LR^* e $\widetilde{LR}$), introduzidos

inicialmente em [Silva \(2010\)](#) e estudados novamente neste trabalho, trazem a taxa de rejeição empírica para um valor mais próximo do nível nominal. Por outro lado, neste trabalho foram introduzidas duas estatísticas de teste corrigidas baseadas na estatística original LR . A primeira delas, a estatística LR_{boot} , é baseada na metodologia bootstrap, vide [Efron \(1979\)](#), e consiste na estimação do percentil de referência da distribuição empírica da estatística LR . O teste baseado na estatística LR_{boot} traz a taxa de rejeição empírica para um valor mais próximo do nível nominal, porém, não corrige totalmente a tendência liberal do teste da razão de verossimilhanças em rejeitar demasiadamente a hipótese nula. A segunda estatística proposta neste trabalho baseada na estatística original LR , a estatística de razão de verossimilhança Bartlett-bootstrap, LR_{boot}^* , apresentada em [Rocke \(1989\)](#), é obtida estimando a media da distribuição de LR . O teste LR_{boot}^* mostrou-se eficiente corrigindo a tendência do teste original de rejeitar com demasiada frequência a hipótese nula, exibindo taxas de rejeição menores que o nível nominal do teste. Comparando os testes baseados nas estatísticas $LR, \widetilde{LR}, LR^*, LR_{boot}$ e LR_{boot}^* , vemos que os testes $\widetilde{LR}, LR^*$ e LR_{boot} apresentaram taxas de rejeição bem semelhantes e mais próximas do nível nominal enquanto o teste LR_{boot} corrige o teste original sendo ligeiramente menos preciso do que os outros testes modificados.

- b. Em relação aos estudos de simulação realizados com o objetivo de analisar a influência do número de parâmetros de perturbação no desempenho dos testes, concluímos que o número de parâmetros de perturbação tem impacto considerável sobre a aproximação por χ^2 para a distribuição da estatística da razão de verossimilhanças (LR). É importante ressaltar que o impacto é bem menos marcante no caso dos testes corrigidos $LR^*, \widetilde{LR}, LR_{boot}, LR_{boot}^*$. Para estes testes, as taxas de rejeição permanecem mais estáveis do que para o teste baseado na estatística LR .
- c. Nas simulações envolvendo o poder dos testes, realizadas com o quantil estimado da distribuição empírica de LR para tirar o efeito associado a liberalidade do teste original, é observado que os testes $LR^*, \widetilde{LR}, LR_{boot}$ e LR_{boot}^* são similares entre eles, em termos de poder, mas são ligeiramente menos poderosos que o teste original LR .
- d. A estatística LR_{boot}^* , proposta originalmente neste trabalho nos MSPNLGs, é uma opção muito atraente para realizar testes de hipótese nestes modelos uma vez que a sua obtenção não envolve a álgebra dispendiosa das versões corrigidas via Bartlett. Além disso, baseados nos resultados obtidos podemos dizer que o desempenho desta estatística corrigida é comparável com o desempenho obtido pelas versões corrigidas da estatística de razão de verossimilhanças, via Bartlett, em termos que em termos de proximidade com a taxa nominal. Além disso, LR_{boot}^* não mostrou perda considerável de poder.

- e. O Capítulo 4 apresenta os resultados obtidos das técnicas de diagnóstico como análise dos resíduos, alavancagem generalizada, influência global e influência local em vários cenários de perturbação. Foram realizadas simulações de Monte Carlo para avaliar os métodos de detecção propostos em dois cenários não lineares envolvendo distribuições pertencentes aos MSPNLGs. Os resultados e a análise gráfica dos resultados mostraram-se eficientes para detectar pontos de alta alavancagem, assim como observações aberrantes e influentes entre o conjunto de dados simulados. Tal capítulo é a principal contribuição teórica desta dissertação. Portanto, este trabalho constitui-se pioneiro na exploração deste tipo de técnicas nesta classe de modelos vindo a preencher parcialmente a necessidade deste tipo de estudos para modelos discretos contidos nos MSPNLGs.

Finalmente, todos os resultados obtidos nesta dissertação foram aplicados à um conjunto de dados, previamente estudado em [Barbour & Brown \(1974\)](#), posteriormente em [Rigby et al. \(2008\)](#) e finalmente em [Cordeiro et al. \(2009\)](#), estudando uma população de peixe em lagos de diferentes comprimentos. Os procedimentos feitos mostraram-se úteis na hora de realizar testes de hipóteses e achar pontos que pesam desmedidamente nas estimativas de máxima verossimilhança dos parâmetros do modelos de regressão adotado.

- ATKINSON, A. (1981). Two graphical displays for outlying and influential observations in regression. *Biometrika* 68(1), 13–20.
- BARBOUR, C. D. & BROWN, J. (1974). Fish species diversity in lakes. *American Naturalist* 108(962), 473–489.
- BARNDORFF-NIELSEN, O. & HALL, P. (1988). On the level-error after Bartlett adjustment of the likelihood ratio statistic. *Biometrika* 75(2), 374–378.
- BARTLETT, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences* 160(901), 268–282.
- BATES, D. & WATTS, D. (1988). *Nonlinear Regression Analysis and Its Applications*. New York.
- BAYER, F. M. & CRIBARI-NETO, F. (2013). Bartlett corrections in beta regression models. *Journal of Statistical Planning and Inference* 143(3), 531–547.
- BECKMAN, R. J., NACHTSHEIM, C. J. & COOK, R. D. (1987). Diagnostics for mixed-model analysis of variance. *Technometrics* 29(4), 413–426.
- BELSEY, D. A., KUH, E., & WELSCH, R.E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. John Wiley.
- BORSSOI, J. A., DE BASTIANI, F., URIBE-OPAZO, M. A. & GALEA, M. (2011). Local influence of explanatory variables in gaussian spatial linear models. *Chilean Journal of Statistics* 2(2), 29–38.

- CAMERON, A. C. & TRIVEDI, H. K. (1998). *Regression analysis of count data*. Number 53. Cambridge university press.
- CONSUL, P. (1990). New class of location-parameter discrete probability distributions and their characterizations. *Communications in Statistics-Theory and Methods* 19(12), 4653–4666.
- COOK, R. D. (1977). Detection of influential observation in linear regression. *Technometrics* 19(1), 15–18.
- COOK, R. D. (1986). Assessment of local influence. *Journal of the Royal Statistical Society. Series B (Methodological)* 48(2), 133–169.
- COOK, R. D. & WEISBERG, S. (1982). Residuals and influence in regression.
- CORDEIRO, G. M. (1982). *Improved likelihood ratio statistics for generalized linear models*. Ph. D. thesis, Imperial College of Science and Technology-University of London.
- CORDEIRO, G. M. (1983). Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistical Society. Series B (Methodological)* 45(3), 404–413.
- CORDEIRO, G. M. (1987). On the corrections to the likelihood ratio statistics. *Biometrika* 74(2), 265–274.
- CORDEIRO, G. M. (2004). Corrected likelihood ratio tests in symmetric nonlinear regression models. *Journal of Statistical Computation and Simulation* 74(8), 609–620.
- CORDEIRO, G. M., ANDRADE, M. G. & DE CASTRO, M. (2009). Power series generalized nonlinear models. *Computational Statistics & Data Analysis* 53(4), 1155–1166.
- CORDEIRO, G. M. & CRIBARI-NETO, F. (2014). *An Introduction to Bartlett Correction and Bias Reduction*. Springer.
- CORDEIRO, G. M., CRIBARI-NETO, F., AUBIN, E. C. Q. & FERRARI, S. L. P. (1995). Bartlett corrections for one-parameter exponential family models. *Journal of Statistical Computation and Simulation* 53(3-4), 211–231.
- CORDEIRO, G. M. & FERRARI, S. L. P. (1991). A modified score test statistic having chi-squared distribution to order $n - 1$. *Biometrika* 78(3), 573–582.
- CORDEIRO, G. M., FERRARI, S. L. P. & PAULA, G. A. (1993). Improved score tests for generalized linear models. *Journal of the Royal Statistical Society. Series B (Methodological)* 55(3), 661–674.

- CORDEIRO, G. M. & MONTENEGRO, L. C. C. (2001). Bartlett corrected likelihood ratio tests in location-scale nonlinear models. *Communications in Statistics-Theory and Methods* 30(7), 1353–1372.
- CORDEIRO, G. M. & PAULA, G. A. (1989). Improved likelihood ratio statistics for exponential family nonlinear models. *Biometrika* 76(1), 93–100.
- COX, D. R. & SNELL, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society. Series B (Methodological)*, 248–275.
- CRIBARI-NETO, F. & FERRARI, S. L. P. (1995). Bartlett-corrected tests for heteroskedastic linear models. *Economics Letters* 48(2), 113–118.
- CRIBARI-NETO, F. & ZARKOS, S. (1995). Improved test statistics for multivariate regression. *Economics Letters* 49(2), 113–120.
- CYSNEIROS, A. H. M. A. (1997). Correção de bartlett e tipo-bartlett em modelos lineares generalizados. *Instituto de matemática e estatística. Universidade de São Paulo (USP) Dissertação de mestrado*.
- CYSNEIROS, A. H. M. A. & FERRARI, S. L. P. (2006). An improved likelihood ratio test for varying dispersion in exponential family nonlinear models. *Statistics & probability letters* 76(3), 255–265.
- DA SILVA, R. P. (2013). Aperfeiçoamento de testes em modelos lineares generalizados super-dispersados. Master's thesis, Universidade Federal de Pernambuco.
- DE BASTIANI, F., CYSNEIROS, A. H. M. A., URIBE-OPAZO, M. A., & GALEA, M. (2014). Influence diagnostics in elliptical spatial linear models. *TEST* .(), 1–19.
- DE SOUZA, F. A. & PAULA, G. A. (2002). Theory & methods: Deviance residuals for an angular response. *Australian & New Zealand Journal of Statistics* 44(3), 345–356.
- DOORNIK, J. A. (2006). An object-oriented matrix language–ox. *Timberlake Consultants Press, London* 107(440), 256–259.
- EFRON, B. (1979). Bootstrap methods: another look at the jackknife. *The annals of Statistics* 7(1), 1–26.
- ESCOBAR, L. A. & MEEKER, W. Q. (1992). Assessing influence in regression analysis with censored data. *Biometrics*, 507–528.
- ESPINHEIRA, P. L., FERRARI, S. L. & CRIBARI-NETO, F. (2008). Influence diagnostics in beta regression. *Computational Statistics & Data Analysis* 52(9), 4417–4431.

- FERRARI, S. L. P. & ARELLANO, R. B. (1996). Modified likelihood ratio and score tests in regression models using the t distribution. *Brazilian Journal of Probability and Statistics* 10(15–33).
- FERRARI, S. L. P., CYSNEIROS, A. H. M. A. & CRIBARI-NETO, F. (2004). An improved test for heteroskedasticity using adjusted modified profile likelihood inference. *Journal of statistical planning and inference* 124(2), 423–437.
- FERRARI, S. L. P. & URIBE, M. A. (2001). Corrected likelihood ratio test in class of symmetric linear regression models. *Brazilian Journal of Probability and Statistics* 15(49–67).
- FRYDENBERG, M. & JENSEN, J. L. (1989). Is the ‘improved likelihood ratio statistic’ really improved in the discrete case? *Biometrika* 76(4), 655–661.
- GALEA, M., LEIVA-SÁNCHEZ, V. & PAULA, G. (2004). Influence diagnostics in log-birnbaum-saunders regression models. *Journal of Applied Statistics* 31(9), 1049–1064.
- GALEA, M., PAULA, G. A. & BOLFARINE, H. (1997). Local influence in elliptical linear regression models. *Journal of the Royal Statistical Society: Series D (The Statistician)* 46(1), 71–79.
- GALEA, M., PAULA, G. A. & CYSNEIROS, F. J. A. (2005). On diagnostics in symmetrical nonlinear models. *Statistics & probability letters* 73(4), 459–467.
- GUPTA, R. (1974). Modified power series distribution and some of its applications. *Sankhyā B* 36, 288–298.
- HOAGLIN, D. C. & WELSCH, R. E. (1978). The hat matrix in regression and anova. *The American Statistician* 32(1), 17–22.
- IHAKA, R. & GENTLEMAN, R. (1996). R: a language for data analysis and graphics. *Journal of computational and graphical statistics* 5(3), 299–314.
- KNUTH, D. E. (1984). The texbook, volume a of computers and typesetting.
- LAWLEY, D. (1956). A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika* 43, 295–303.
- LEMONTE, A. J., CORDEIRO, G. M. & MORENO, G. (2012). Bartlett corrections in birnbaum-saunders nonlinear regression models. *Journal of Statistical Computation and Simulation* 82(6), 927–935.
- LEMONTE, A. J., FERRARI, S. L. & CRIBARI-NETO, F. (2010). Improved likelihood inference in birnbaum-saunders regressions. *Computational Statistics & Data Analysis* 54(5), 1307–1316.

- LEMONTE, A. J. & PATRIOTA, A. G. (2011). Influence diagnostics in birnbaum–saunders nonlinear regression models. *Journal of Applied Statistics* 38(5), 871–884.
- LESAFFRE, E. & VERBEKE, G. (1998). Local influence in linear mixed models. *Biometrics*, 570–582.
- LIMA, M. (2012). Aperfeiçoamento de testes nos modelos séries de potência não-lineares generalizados. *Universidade Federal de Pernambuco. Programa de pós graduação em estatística Dissertação de mestrado*.
- LUENBERGER, D. (1973). *Introduction to linear and nonlinear programming*, Volume 28. Addison-Wesley Reading, MA.
- MCCULLAGH, P. (1987). *Tensor methods in statistics*, Volume 161. Chapman and Hall London.
- NELDER, J. A. & MCCULLAGH, P. (1989). *Generalized linear models*, Volume 2. Chapman and Hall London.
- NOCEDAL, J. & WRIGHT, S. J. (1999). *Numerical Optimization*. Springer Series in Operations Research. Springer.
- NOMA, H. (2011). Confidence intervals for a random-effects meta-analysis based on bartlett-type corrections. *Statistics in medicine* 30(28), 3304–3312.
- PATRIOTA, A. G., LEMONTE, A. J. & DE CASTRO, M. (2010). Influence diagnostics in a multivariate normal regression model with general parameterization. *Statistical Methodology* 7(6), 644–654.
- PAULA, G. A. (2003). *Modelos de regressão com apoio computacional*. Universidade De São Paulo.
- PREGIBON, D. (1981). Logistic regression diagnostics. *The Annals of Statistics*, 705–724.
- RAO, R. C. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. In *Mathematical Proceedings of the Cambridge Philosophical Society*, Volume 44, pp. 50–57. Cambridge Univ Press.
- RIGBY, R. A., STASINOPOULOS, M. D. & AKANTZILIOTOU, C. (2008). A framework for modelling overdispersed count data, including the poisson-shifted generalized inverse gaussian distribution. *Computational Statistics & Data Analysis* 53(2), 381–393.
- ROCKE, D. (1989). Bootstrap Bartlett adjustment in seemingly unrelated regression. *Journal of the American Statistical Association* 84(406), 598–601.

- SEBER, G. & WILD, C. (1989). *Nonlinear Regression*. New York.
- SILVA, P. (2010). Correção de viés e de bartlett em modelos em séries de potência não-lineares generalizados. *Universidade Federal de Pernambuco. Programa de pós graduação em estatística Dissertação de mestrado*.
- ST LAURENT, R. T. & COOK, R. D. (1992). Leverage and superleverage in nonlinear regression. *Journal of the American Statistical Association* 87(420), 985–990.
- TERRELL, G. R. (2002). The gradient statistic. *Computing Science and Statistics* 34, 206–215.
- VARGAS, T. M., FERRARI, S. L. & LEMONTE, A. J. (2014). Improved likelihood inference in generalized linear models. *Computational Statistics & Data Analysis* 74, 110–124.
- VILLEGAS, C., PAULA, G. A., CYSNEIROS, F. J. A. & GALEA, M. (2013). Influence diagnostics in generalized symmetric linear models. *Computational Statistics & Data Analysis* 59, 161–170.
- WEI, B.C. (1998). *Exponential family nonlinear models*, Volume 130. Springer Verlag.
- WEI, B.C., HU, Y.Q. & FUNG, W.K. (1998). Generalized leverage and its applications. *Scandinavian Journal of statistics* 25(1), 25–37.

Obtenção dos cumulantes envolvidos na correção de Bartlett

Considerando o modelo em séries de potência não linear generalizado, apresentado na Seção 2.1, supondo o parâmetro de dispersão ϕ conhecido. Neste apêndice, apresentamos a obtenção de expressões gerais para as quantidades necessárias para o cálculo do fator de correção de Bartlett para a estatística de razão de verossimilhança, apresentado no capítulo 2.

O logaritmo da função de verossimilhança do parâmetro β , dado o vetor de observações $\mathbf{y} = (y_1, \dots, y_n)^\top$, do MSPNLG é dado por

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

em que a função $a(y, \phi)$ é positiva e as funções analíticas $g(\mu_i, \phi)$ e $f(\mu_i, \phi)$ dos parâmetros μ_i e ϕ são positivas, finitas e duas vezes diferenciáveis. Denotando $f_i = f(\mu_i, \phi)$, $g_i = g(\mu_i, \phi)$ e o índice sobrescrito (j) indicando a j -ésima derivada em relação μ , $j = 1, 2, 3$, definimos a

seguir algumas quantidades, que serão de grande valia para a simplificação dos cálculos:

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

para $i = 1, \dots, n$ e $j = 1, 2, 3$. Vale ressaltar que as quantidades acima envolvem derivadas que dependem das formas específicas das funções f , g , h e V nas diversas distribuições pertencentes à família de série de potência. Denotando $\tilde{x}_{ir} = \partial \eta_i / \partial \beta_r$, $\tilde{x}_{irs} = \partial^2 \eta_i / \partial \beta_r \partial \beta_s$ e $\tilde{x}_{irst} = \partial^3 \eta_i / \partial \beta_r \partial \beta_s \partial \beta_t$, temos os seguintes resultados:

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

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

e

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

A.1 Derivadas de $\ell(\beta; y)$

Por simples diferenciação em relação aos componentes do parâmetro β , e usando os índices $i = 1, \dots, n$, junto com $r, s, t, u = 1, \dots, k$, temos que

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

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

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

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

e

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

A.2 Cálculo dos cumulantes

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

$$\begin{aligned}\kappa_{rs} &= \sum_i w_{1i} \tilde{x}_{is} \tilde{x}_{ir}, \\ \kappa_{rst} &= \sum_i \{ q_{i1} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + w_{1i} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs}) \}, \\ \kappa_{rstu} &= \sum_i \{ q_{7i} \tilde{x}_{iu} \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{ir} + q_{1i} [\tilde{x}_{itu} \tilde{x}_{is} \tilde{x}_{ir} + \tilde{x}_{it} \tilde{x}_{isu} \tilde{x}_{ir} + \tilde{x}_{it} \tilde{x}_{is} \tilde{x}_{iru} \\ &\quad + \tilde{x}_{iu} (\tilde{x}_{ist} \tilde{x}_{ir} + \tilde{x}_{is} \tilde{x}_{irt} + \tilde{x}_{it} \tilde{x}_{irs})] + w_{1i} (\tilde{x}_{istu} \tilde{x}_{ir} + \tilde{x}_{ist} \tilde{x}_{iru} + \tilde{x}_{isu} \tilde{x}_{irt} \\ &\quad + \tilde{x}_{is} \tilde{x}_{irtu} + \tilde{x}_{itu} \tilde{x}_{irs} + \tilde{x}_{it} \tilde{x}_{irsu} + \tilde{x}_{iu} \tilde{x}_{irst}) \},\end{aligned}$$

A.3 Derivadas dos cumulantes

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

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

e

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

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

Suponha que o logaritmo da função de verossimilhança para o parâmetro θ seja agora expresso na forma

$$\ell_\delta(\theta; \mathbf{y}) = \sum_{i=1}^n \delta_i \ell_i(\theta; y_i), \quad (\text{B.1})$$

em que $\ell(\theta; y_i)$ denota o logaritmo da função de verossimilhança correspondente à i -ésima observação e δ_i é um tipo de perturbação, definida 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 fica, supondo a estrutura (B.1), dada por

$$\hat{\beta}_\delta = (\tilde{X}^\top \Delta \hat{W} \tilde{X})^{-1} \tilde{X}^\top \Delta \hat{W} \hat{\xi}, \quad (\text{B.2})$$

em que $\hat{\xi} = \tilde{X} \hat{\beta} + \hat{L}(\mathbf{y} - \hat{\mu})$ e Δ é uma matriz de uns com δ na i -ésima posição. Ou seja, $\Delta = \text{diag}\{1, \dots, \delta, \dots, 1\}$. Assim

$$\tilde{X}^\top \Delta \hat{W} \tilde{X} = \tilde{X}^\top \hat{W} [I - (1 - \Delta)] \tilde{X} = \tilde{X}^\top \hat{W} \tilde{X} - (1 - \delta) \tilde{x}_i \hat{\omega}_i \tilde{x}_i^\top.$$

Analogamente, tem-se que $\tilde{X}^\top \hat{W} \hat{\xi} = \tilde{X}^\top \hat{W} \hat{\xi} - (1 - \delta) \tilde{x}_i \hat{\omega}_i \hat{\xi}_i$.

Para calcular $(\tilde{X}^\top \Delta \hat{W} \tilde{X})^{-1}$, 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 $\mathbf{A} = \tilde{X}^\top \hat{W} \tilde{X}$, $\mathbf{U} = -(1 - \delta)\tilde{x}_i \hat{\omega}_i$ e $\mathbf{V}^\top = \tilde{x}_i^\top$.

Logo,

$$\begin{aligned} (\tilde{X}^\top \Delta \hat{W} \tilde{X})^{-1} &= (\tilde{X}^\top \hat{W} \tilde{X})^{-1} - \frac{[(\tilde{X}^\top \hat{W} \tilde{X})^{-1}(-(1 - \delta)\tilde{x}_i \hat{\omega}_i)][\tilde{x}_i^\top (\tilde{X}^\top \hat{W} \tilde{X})]}{1 + \tilde{x}_i^\top (\tilde{X}^\top \hat{W} \tilde{X})^{-1}(-(1 - \delta)\tilde{x}_i \hat{\omega}_i)} \\ &= (\tilde{X}^\top \hat{W} \tilde{X})^{-1} + \frac{[(1 - \delta)(\tilde{X}^\top \hat{W} \tilde{X})^{-1}\tilde{x}_i \hat{\omega}_i \tilde{x}_i^\top (\tilde{X}^\top \hat{W} \tilde{X})^{-1}]}{1 - (1 - \delta)\hat{\omega}_i^{1/2} \tilde{x}_i^\top (\tilde{X}^\top \hat{W} \tilde{X})^{-1} \tilde{x}_i \hat{\omega}_i^{1/2}} \\ &= (\tilde{X}^\top \hat{W} \tilde{X})^{-1} + \frac{[(1 - \delta)(\tilde{X}^\top \hat{W} \tilde{X})^{-1}\tilde{x}_i \hat{\omega}_i \tilde{x}_i^\top (\tilde{X}^\top \hat{W} \tilde{X})^{-1}]}{1 - (1 - \delta)\hat{h}_{ii}}, \end{aligned}$$

em que $\hat{h}_{ii} = \hat{\omega}_i^{1/2} \tilde{x}_i^\top (\tilde{X}^\top \hat{W} \tilde{X})^{-1} \tilde{x}_i \hat{\omega}_i^{1/2}$.

Portanto, inserindo a última expressão obtida em (B.2), fica

$$\hat{\beta}_\delta = \hat{\beta} - \frac{(1 - \delta)\hat{r}_i^\beta (\tilde{X}^\top \hat{W} \tilde{X})^{-1} \tilde{x}_i \hat{\omega}_i^{1/2}}{1 - (1 - \delta)\hat{h}_{ii}},$$

em que $\tilde{x}_i$ e $\hat{\omega}_i^{1/2}$ denotam as i -ésimas linhas de $\tilde{X}$ e $\hat{W}$, $\hat{r}_i^\beta = (y_i - \hat{\mu}_i) / \widehat{\text{Var}}(Y_i)$. O caso $\delta = 0$ significa que o i -ésimo ponto foi excluído, ou seja, $\hat{\beta}_{(i)}$, que fica dado por

$$\hat{\beta}_{(i)} = \hat{\beta} - \frac{\hat{r}_i^\beta}{(1 - \hat{h}_{ii})} (\tilde{X}^\top \hat{W} \tilde{X})^{-1} \tilde{x}_i \hat{\omega}_i^{1/2},$$

em que $\hat{h}_{ii}$ corresponde ao i -ésimo elemento da diagonal principal da matriz

$$\hat{H} = \hat{W}^{1/2} \tilde{X} (\tilde{X}^\top \hat{W} \tilde{X})^{-1} \tilde{X}^\top \hat{W}^{1/2}$$

APÊNDICE C

Banco de dados

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

y	x	y	x	y	x	y	x	y	x
10	1.609	14	3.989	37	9.112	14	3.178	22	4.094
37	3.714	67	11.320	22	8.102	28	9.244	15	4.263
60	5.142	36	3.584	18	8.196	17	0.693	9	2.708
113	10.155	30	0.000	214	10.401	17	10.545	23	4.585
99	10.995	19	1.609	177	11.149	21	5.398	48	6.528
13	0.000	46	8.584	17	11.074	13	8.445	21	5.357
30	3.784	68	7.636	50	10.358	14	5.442	46	6.516
114	10.968	93	9.770	5	9.826	21	8.875	14	6.985
112	9.877	13	6.512	22	7.026	24	6.423	7	4.710
17	2.303	53	7.673	156	12.956	12	10.347	5	2.079
10	4.443	17	7.771	74	12.985	26	10.211	40	9.020
14	0.000	245	10.257	13	5.106	13	6.006	18	9.112
39	5.159	88	8.392	11	8.733	19	5.989	20	5.878
14	1.099	24	3.367	48	9.820	19	7.262	17	5.849