

REFINAMENTOS DE TESTES NA CLASSE DOS MODELOS NÃO-LINEARES SIMÉTRICOS HETEROSCEDÁSTICOS

Mariana Correia de Araújo

Orientadora: Prof^a Dr^a Lourdes Coral Contreras Montenegro
Co-orientadora: Prof^a Dr^a Audrey Helen Mariz de Aquino Cysneiros

Tese submetida como requerimento parcial para obtenção do grau de Doutor em
Estatística pela Universidade Federal de Minas Gerais

Belo Horizonte, outubro de 2015

*Para Geraldo (in memoriam) e Rosinha,
pelo dom da vida e todo amor.*

Agradecimentos

Agradeço a Deus, por segurar minhas mãos nas horas difíceis, não me deixar fraquejar e me fazer perseverar quando tudo parecia estar indo de encontro ao meu objetivo.

A minha mãe, por todo amor e carinho dedicado. Pelas tantas ligações diárias enquanto morávamos longe, pela alegria a cada retorno de férias e por tudo que fez e faz por mim.

Ao meu pai (*in memoriam*), por toda a educação, ensinamentos, amor e carinho. O céu está em festa, tenho certeza.

Às professoras Lourdes Montenegro e Audrey Cysneiros, minhas orientadoras, pela dedicação ao trabalho e confiança em mim depositada. Sou imensamente grata por todos os seus ensinamentos e incentivos.

A Amanda, Laís, Larissa, Manolo, Marcondes, Natália, Romison, Saulo e Tamyris, que mesmo distantes, se fizeram presentes em Belo Horizonte. Contar com o apoio de vocês foi fundamental.

A Fernando e William, pela amizade, por cada conversa, estudo e momentos de lazer que nos acompanha desde o mestrado. Sei que posso contar com vocês para as mais diversas coisas e sou muito grata por isto. A academia me deu dois belos presentes em forma de amigo.

A Fabrícia, Fádua, Isabella, Nívea, Silvana e Talita, pela convivência diária, companhia e ensinamentos. Foi muito enriquecedor tudo que passamos, trocas que levarei pela vida inteira. Obrigada pelo carinho e paciência nos dias difíceis.

A Paulo e Wecley, companheiros do doutorado, pelas discussões construtivas, horas de estudo dispensadas e momentos de descontração.

Aos professores do Programa de Pós-Graduação em Estatística da UFMG, pelo valioso conhecimento repassado.

Aos funcionários do Departamento de Estatística da UFMG, em especial, Jéssica, Kate, Maísa, Rogéria e Rose, pela atenção dispensada.

Aos membros da banca, pelas valiosas contribuições para o enriquecimento do trabalho.

À CAPES, pelo apoio financeiro.

“Você é do tamanho do seu sonho.”

Autor desconhecido

Resumo

Nesta tese abordamos o aperfeiçoamento de testes de hipóteses na classe dos modelos não-lineares simétricos heteroscedásticos. Inicialmente, aprimoramos os testes de heteroscedasticidade baseados nas estatísticas da razão de verossimilhanças e razão de verossimilhanças perfiladas modificadas. A versão modificada da verossimilhança perfilada considerada neste trabalho é a proposta por [Cox e Reid \(1987\)](#). Em seguida, obtemos um fator de correção tipo-Bartlett para o teste de heteroscedasticidade baseado na estatística gradiente. A estatística gradiente foi proposta por [Terrell \(2002\)](#) e tem recebido crescente atenção na literatura, uma vez que além de ter estrutura simples de calcular e implementar, é assintoticamente equivalente às estatísticas da razão de verossimilhanças, Wald e score. Ainda, realizamos um estudo de poder local dos testes da razão de verossimilhanças, Wald, score e gradiente a fim de avaliar (localmente) o poder dos quatro testes. Para cada abordagem realizada, foi desenvolvido um estudo de simulação de Monte Carlo a fim de avaliar o desempenho dos testes sob investigação.

Palavras chave: Correção de Bartlett; Correção tipo-Bartlett; Estatística gradiente; Modelos não-lineares simétricos heteroscedásticos; Poder local; Testes de hipóteses; Verossimilhança perfilada; Verossimilhança perfilada modificada.

Abstract

In this thesis, we deal with improvement for hypotheses tests in heteroscedastic symmetric nonlinear models. First, we derive Bartlett adjustments for likelihood ratio statistics and modified profile likelihood ratio statistics in order to improve likelihood ratio and modified profile likelihood ratio tests, respectively. Next, we calculate a type-Bartlett adjustment to improve the gradient test, a new hypotheses test proposed by [Terrell \(2002\)](#) which is asymptotically equivalent to likelihood ratio, Wald and score tests. We also treat about the local power of likelihood ratio, Wald, score and gradient tests in heteroscedastic symmetric nonlinear models. For each approach, we develop a Monte Carlo simulation study in order to evaluate the performance of the tests in finite-size samples.

Keywords: Bartlett correction; Heteroscedastic symmetric nonlinear regression models; Hypotheses tests; Local power; Modified profile likelihood; Profile likelihood; Type-Bartlett correction.

1	Introdução	1
2	Refinamento dos testes da razão de verossimilhanças e razão de verossimilhanças perfiladas modificadas	6
2.1	Introdução	6
2.2	Especificação do modelo	8
2.3	Correções de Bartlett	12
2.3.1	Correção de Bartlett para a estatística LR	13
2.3.2	Correção de Bartlett para a estatística LR_m	15
2.3.3	Correção de Bartlett bootstrap para a estatística LR	16
2.4	Resultados numéricos	17
2.5	Conclusões	25
3	Refinamento dos testes escore e gradiente	27
3.1	Introdução	27
3.2	Aspectos inferenciais	28
3.3	Correção tipo-Bartlett	30
3.3.1	Correção tipo-Bartlett para a estatística S_r	31
3.3.2	Correção tipo-Bartlett para a estatística S_g	33
3.4	Resultados numéricos	36
3.5	Aplicação	47
3.6	Conclusões	48

4	Poder local dos testes da razão de verossimilhanças, Wald, escore e gradiente	49
4.1	Introdução	49
4.2	Poder local	50
4.3	Comparação entre as funções de poder	54
4.4	Resultados numéricos	56
4.5	Conclusões	62
5	Considerações finais	63
A	Aspectos inferenciais do modelos não-lineares simétricos heteroscedásticos	65
B	Cálculo do fator de correção tipo-Bartlett para a estatística LR	68
C	Cálculo do fator de correção c_m para a estatística LR_m	75
D	Derivadas do logaritmo da função de verossimilhança de $\theta = (\beta, \delta)^\top$	78
E	Cálculo dos termos A_1^g, A_2^g e A_3^g do fator de correção tipo-Bartlett para a estatística S_g	84
F	Poder local: obtenção dos b_{il}'s e ξ e comparação de funções de poder	86
G	Coelhos europeus na Austrália	90

Lista de Figuras

2.1	Discrepância relativa de quantis para o modelo t_5 com $n = 35$, $p = 3$, $k = 3$	22
2.2	Discrepância relativa de quantis para o modelo t_5 com $n = 35$, $p = 5$, $k = 3$	22
2.3	Discrepância relativa de quantis para o modelo exponencial potência com $\kappa = 0, 3$, considerando $n = 35$, $p = 3$, $k = 3$	23
2.4	Discrepância relativa de quantis para o modelo exponencial potência com $\kappa = 0, 3$, considerando $n = 35$, $p = 5$, $k = 3$	23
3.1	Discrepância relativa de quantis para o modelo t_5 com $n = 30$, $p = 3$, $k = 3$	44
3.2	Discrepância relativa de quantis para o modelo t_5 com $n = 30$, $p = 5$, $k = 3$	44
3.3	Discrepância relativa de quantis para o modelo exponencial potência com $\kappa = 0,3$, considerando $n = 30$, $p = 3$, $k = 3$	45
3.4	Discrepância relativa de quantis para o modelo exponencial potência com $\kappa = 0,3$, considerando $n = 30$, $p = 5$, $k = 3$	45
4.1	Curva de poder dos quatro testes para o modelo t_5 com $n = 35$, $p = 3$, $k = 4$	61
4.2	Curva de poder dos quatro testes para o modelo exponencial potência com $\kappa = 0, 3$, $n = 35$, $p = 3$, $k = 4$	61

Lista de Tabelas

2.1	Expressões para $g(z^2)$, $\frac{g'(z^2)}{g(z^2)}$ e s para algumas distribuições simétricas.	10
2.2	Tamanho dos testes para o Modelo t_5 com $p = 3, 5$, $k = 3$ e diversos valores para n	20
2.3	Tamanho dos testes para o Modelo exponencial potência com $\kappa = 0, 3$, $p = 3, 5$, $k = 3$ e diversos valores para n	21
2.4	Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0, 3$, considerando $n = 35$, $k = 3$ e diversos valores para p	24
2.5	Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0, 3$, considerando $n = 35$, $p = 3$ e diversos valores para k	25
3.1	Tamanho dos testes para o Modelo t_5 com $n = 35$, $p = 3, 5$, $k = 3$ e diversos valores para n	39
3.2	Tamanho dos testes para o Modelo exponencial potencia com $\kappa = 0.3$, $p = 3, 5$, $k = 3$ e n	40
3.3	Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0.3$, considerando $n = 35$, $k = 3$ e diversos valores para p	41
3.4	Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0.3$, considerando $n = 35$, $p = 3$ e diversos valores para k	42
3.5	Poder dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0.3$, conside- rando $n = 30$, $p = 3, 5$, $k = 3$ e $\alpha = 10\%$	46
4.1	Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0, 3$, considerando $n = 30$, $k = 3$ e diversos valores para p	57

4.2	Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0,3$, considerando $n = 30$, $p = 3$ e diversos valores para k	58
4.3	Tamanho dos testes para o Modelo t_5 com $p = 3$, $k = 4$ e diversos valores para n	59
4.4	Tamanho dos testes para o Modelo exponencial potência com $\kappa = 0,3$, $p = 3$, $k = 4$ e diversos valores para n	60
G.1	Peso das lentes dos olhos de coelhos europeus (y), em miligramas, e a idade (x), em dias, numa amostra contendo 71 observações. (Wei, 1998, Exemplo 6.8) . . .	90

CAPÍTULO 1

Introdução

Embora muito atrativa, a suposição de normalidade para os erros de modelos de regressão com resposta contínua nem sempre é adequada. A presença de observações extremas no conjunto de dados, por exemplo, interfere nas estimativas dos parâmetros dos modelos normais, isto é, tais estimativas são sensíveis a observações aberrantes. Alternativas robustas ao modelo normal têm sido amplamente estudadas na literatura, entre elas, modelos cujos erros seguem distribuições pertencentes à família simétrica, uma vez que esta classe de distribuições contempla, também, distribuições com caudas mais pesadas que a da normal, e assim, acomodando melhor as observações extremas, reduzindo a influência exercida por tais observações sob as estimativas dos parâmetros do modelo. Além da distribuição normal, pertencem à família de distribuições simétricas as distribuições Cauchy, t -Student, t -Student generalizada, logísticas I e II, logística generalizada, exponencial potência, Kotz, Kotz generalizada, entre outras. Detalhes sobre as distribuições simétricas podem ser obtidos em [Fang et al. \(1990\)](#) e [Fang e Anderson \(1990\)](#).

Podemos encontrar aplicações desta família de distribuições nas mais diversas áreas do conhecimento, tais como engenharia, biologia, medicina, ciências sociais, economia, entre outras. Diversos artigos têm sido desenvolvidos com aplicações da distribuição simétrica, por exemplo, [Lang et al. \(1989\)](#) introduziram o modelo de regressão t -Student como alternativa robusta ao modelo normal, [Cordeiro et al. \(1998\)](#) obtiveram uma correção de viés para os estimadores do modelo não-linear com erros t -Student, sendo estes resultados

estendidos por [Cordeiro et al. \(2000\)](#) para a classe dos modelos não-lineares simétricos, [Cysneiros et al. \(2007\)](#) introduziram os modelos lineares simétricos heteroscedásticos e desenvolveram métodos de influência local e [Cysneiros et al. \(2010b\)](#) aperfeiçoaram o teste escore para esta classe de modelos, [Cysneiros et al. \(2010\)](#) fizeram correção de viés para os estimadores do modelo não-linear simétrico heteroscedástico.

Além da distribuição, é extremamente importante verificar se a suposição de variância constante (homoscedasticidade) dos erros é satisfeita, dado que a violação desse pressuposto altera completamente a estratégia da modelagem. Os testes de heteroscedasticidade comumente utilizados na literatura baseiam-se em resultados assintóticos, sendo os mais empregados os testes da razão de verossimilhanças, escore e Wald, cujas estatísticas de teste são equivalentes até primeira ordem ¹ e têm distribuição assintótica nula χ_k^2 com erro de ordem n^{-1} , sendo k o número de restrições impostas pela hipótese nula e n o tamanho da amostra. Recentemente, [Terrell \(2002\)](#) propôs um novo teste assintótico cuja estatística de teste é chamada estatística gradiente. Esta estatística apresenta uma estrutura simples, a qual não depende da matriz de informação observada, nem da esperada. Além disso, a estatística gradiente é de fácil implementação e sob hipótese nula tem distribuição assintótica χ_k^2 a menos de termos de ordem n^{-1} . Em contrapartida, ao contrário das estatísticas da razão de verossimilhanças e escore, a estatística gradiente não é invariante sob reparametrizações não-lineares ([Terrell, 2002](#)). Diversos trabalhos na literatura têm abordado a estatística gradiente como tema, dentre eles, [Lemonte \(2011\)](#), [Lemonte \(2012\)](#), [Lemonte e Ferrari \(2012\)](#), [Lemonte e Ferrari \(2012b\)](#), [Vargas et al. \(2013\)](#), [Lemonte \(2014\)](#) e [Vargas et al. \(2014\)](#).

Por se basearem em resultados assintóticos, a aproximação da distribuição das estatísticas dos testes da razão de verossimilhanças, escore, Wald e gradiente pela distribuição qui-quadrado de referência pode não ser satisfatória, conduzindo a testes com taxas de rejeição distorcidas quando o tamanho da amostra é pequeno ou até mesmo moderado. Uma estratégia para melhorar a aproximação da distribuição das estatísticas de teste pela distribuição qui-quadrado é modificar tais estatísticas através um fator de correção. Para a estatística da razão de verossimilhanças, [Bartlett \(1937\)](#) propôs um fator de correção a ser multiplicado por tal estatística de modo que a sua versão corrigida apresenta distribuição nula χ_k^2 com erro de ordem n^{-2} . Para a estatística escore, [Cordeiro e Ferrari \(1991\)](#)

¹Sejam $\{a_n\}_{n \geq 1}$ e $\{b_n\}_{n \geq 1}$ sequências de números reais, dizemos que a_n é de ordem menor que b_n , denotando por $a_n = o(b_n)$, se $\lim_{n \rightarrow \infty} a_n/b_n = 0$. Dizemos que a_n é de ordem, *no máximo*, igual a b_n , denotando por $a_n = O(b_n)$, se a razão $|a_n/b_n|$ for limitada para todo n suficientemente grande, isto é, se existirem $K \in \mathbb{R}^+$ e $n_0(K)$ tal que $|a_n/b_n| \leq K$, $\forall n \geq n_0(K)$.

propuseram um fator de correção tipo-Bartlett cujo cálculo envolve a própria estatística escore, derivando assim uma estatística corrigida cuja distribuição nula é χ_k^2 sob erro de aproximação de ordem n^{-2} . Maiores detalhes sobre correções de Bartlett e tipo-Bartlett podem ser encontradas em [Cordeiro e Cribari-Neto \(2014\)](#). Para a estatística gradiente, [Vargas et al. \(2014\)](#) derivaram um fator de correção tipo-Bartlett, fornecendo uma estatística corrigida cuja distribuição assintótica nula é χ_k^2 sob erro de aproximação de ordem n^{-2} .

Quando o modelo em investigação envolve parâmetros de perturbação, a inferência do estudo pode ser baseada através da verossimilhança perfilada. A inferência fundamentada em verossimilhança perfilada trata os parâmetros de perturbação como conhecidos, isto é, na prática, estamos substituindo tais parâmetros por suas respectivas estimativas de máxima verossimilhança. Este procedimento pode introduzir viés na função escore e na informação, veja [Ferrari et al. \(2005\)](#), além disso, o aumento do número de parâmetros de perturbação implica numa menor qualidade das aproximações assintóticas. Para superar tais problemas, algumas modificações para a função de verossimilhança perfilada foram propostas na literatura, entre elas, as propostas por [Barndorff-Nielsen \(1983; 1994\)](#), [Cox e Reid \(1987\)](#), [McCullagh e Tibishirani \(1990\)](#) e [Stern \(1997\)](#), que estão descritas com detalhes em [Severini \(2000, Capítulo 9\)](#) e [Pace e Salvan \(1997, Capítulo 11\)](#). Devido às boas propriedades que se obtém com a ortogonalidade global dos parâmetros do modelo, dentre elas a independência assintótica dos estimadores de máxima verossimilhança dos parâmetros de interesse e perturbação e o menor custo computacional na determinação numérica das estimativas de máxima verossimilhança destes parâmetros, veja [Silva \(2005\)](#), neste trabalho consideraremos a versão proposta por [Cox e Reid \(1987\)](#).

Assim como o teste da razão de verossimilhanças usual, o teste baseado na verossimilhança perfilada modificada tem sua estatística de teste com distribuição nula assintótica χ_k^2 a menos de termos de ordem n^{-1} , sendo esta aproximação insatisfatória quando o tamanho da amostra é pequeno. Visando melhorar a aproximação da distribuição da estatística da razão de verossimilhanças perfiladas modificadas pela distribuição χ^2 de referência, [DiCiccio e Stern \(1994\)](#) propuseram um fator de correção de Bartlett para esta estatística, derivando uma estatística corrigida cuja distribuição nula é χ_k^2 com erro de aproximação de ordem n^{-2} .

Considerando os quatro testes assintóticos equivalentes existentes na literatura, a saber, testes da razão de verossimilhanças, Wald, escore e gradiente, um questionamento pertinente é apontar o teste que seja mais adequado para um determinado estudo em de-

envolvimento. Da literatura sabe-se que, até primeira ordem, os quatro testes têm poder igual sob alternativas de Pitman e até ordem $n^{-1/2}$, o mesmo tamanho, o que implica que a escolha pode ser baseada em um critério que seja determinado através da comparação (local) dos seus poderes até ordem $n^{-1/2}$. Para isto, [Hayakawa \(1975\)](#) desenvolveu expansões assintóticas sob uma sequência de hipóteses alternativas contíguas que convergem para hipótese nula à taxa de $n^{-1/2}$, para as distribuições das estatísticas da razão de verossimilhanças e Wald. [Harris e Peers \(1980\)](#) obtiveram resultado análogo para a estatística escore e recentemente [Lemonte e Ferrari \(2012\)](#) para a estatística gradiente. Assim, o objetivo do estudo do poder até ordem $n^{-1/2}$ é estabelecer condições através das quais podemos comparar localmente os poderes dos quatro testes assintóticos em estudo a partir das distribuições não nulas sob hipóteses alternativas de Pitman de suas respectivas estatísticas de teste (até tal ordem).

Diversos trabalhos têm contemplado os modelos simétricos, refinamentos de testes de hipóteses e estudo de poder local. [Ferrari et al. \(2004\)](#) calcularam um fator de correção de Bartlett para a estatística da razão de verossimilhanças perfiladas modificadas para o modelo normal linear estudado por [Simonoff e Tsai \(1994\)](#). Considerando a classe dos modelos não-lineares simétricos, [Cordeiro \(2004\)](#) obteve um teste da razão de verossimilhanças corrigido e [Cysneiros et al. \(2010b\)](#) um teste escore corrigido. [Lemonte \(2012\)](#) realizou um estudo de poder local dos testes da razão de verossimilhanças, Wald, escore e gradiente na classe dos modelos lineares simétricos.

Nesta direção, esta tese tem três objetivos principais. O primeiro é obter fatores de correção de Bartlett para aprimorar os testes da razão de verossimilhanças e razão de verossimilhanças perfiladas modificadas para a classe dos modelo não-lineares simétricos heteroscedásticos (MNLSH), sendo o segundo teste baseado na verossimilhança perfilada modificada proposta por [Cox e Reid \(1987\)](#), generalizando os resultados obtidos por [Ferrari et al. \(2004\)](#) e [Cordeiro \(2004\)](#). O segundo objetivo é derivar um fator de correção tipo-Bartlett para a estatística gradiente na classe dos MNLSH considerando a metodologia proposta por [Vargas et al. \(2013\)](#). O terceiro objetivo da tese é estender os resultados obtidos por [Lemonte \(2012\)](#) para os modelos não-lineares simétricos heteroscedásticos, isto é, realizar um estudo de poder local dos testes da razão de verossimilhanças, Wald, escore e gradiente na classe dos MNLSH.

Esta tese é dividida em cinco capítulos e sete apêndices. No Capítulo 2, fazemos uma breve introdução ao modelo em estudo e apresentamos alguns aspectos inferenciais do mesmo. Ainda, obtemos fatores de correção de Bartlett para as estatísticas da razão

de verossimilhanças e razão de verossimilhanças perfiladas modificadas. Um estudo de simulação de Monte Carlo é realizado para avaliar o desempenho dos testes corrigidos e não corrigidos em amostras finitas. A efeito de comparação, também são considerados testes cujas respectivas estatísticas de teste são obtidas utilizando a técnica de reamostragem bootstrap (Efron, 1979), a saber, teste da razão de verossimilhanças bootstrap e razão de verossimilhanças Bartlett bootstrap (Rocke, 1989).

No Capítulo 3, derivamos um fator de correção tipo-Bartlett para a estatística gradiente para o teste de heteroscedasticidade na classe dos MNLSH. Um estudo de simulação de Monte Carlo para avaliar o comportamento em amostras finitas do teste gradiente corrigido e não corrigido e diferentes testes disponíveis na literatura é realizado. As avaliações consideraram tamanho e poder dos testes sob diversos cenários. Ainda, aplicamos a metodologia estudada a um conjunto de dados reais.

No Capítulo 4, realizamos um estudo de poder local dos testes da razão de verossimilhanças, Wald, score e gradiente na classe dos MNLSH. Para isto, derivamos expansões assintóticas para a distribuição das estatísticas dos quatro testes supracitados sob hipóteses alternativas de Pitman e comparamos analiticamente os seus poderes até ordem $n^{-1/2}$. Ainda, apresentamos um estudo de simulação de Monte Carlo para avaliar o desempenho dos testes em amostras de tamanho pequeno e moderado. No Capítulo 5, fazemos algumas considerações finais gerais dos resultados obtidos na tese. Nos apêndices são apresentados detalhes técnicos.

É válido observar que esta tese foi escrita de maneira que os Capítulos 2, 3 e 4 possam ser lidos individualmente e em qualquer ordem, assim, algumas notações e resultados básicos são apresentados mais de uma vez.

Refinamento dos testes da razão de verossimilhanças e razão de verossimilhanças perfiladas modificadas

2.1 Introdução

Quando o tamanho da amostra é pequeno ou moderado, sabemos da literatura que o teste da razão de verossimilhanças pode apresentar taxas de rejeição distorcidas, isto é, muito diferentes do esperado. Tal fato se deve à aproximação da distribuição de sua estatística de teste pela distribuição qui-quadrado de referência, que pode não ser satisfatória, uma vez que este teste é baseado em resultados assintóticos. Dessa maneira, se faz necessário o desenvolvimento de procedimentos inferenciais mais acurados, isto é, que nos garanta uma inferência mais precisa.

A fim de melhorar a aproximação da distribuição da estatística da razão de verossimilhanças pela distribuição χ^2 de referência, [Bartlett \(1937\)](#) propôs um fator de correção a ser incorporado a tal estatística de modo a obter um teste aprimorado, proporcionando uma inferência mais confiável quando o tamanho da amostra é pequeno.

Ao realizar inferência em modelos com parâmetros de perturbação, podemos basear nossas investigações em verossimilhança perfilada, que é uma função que só depende do parâmetro de interesse, sendo os parâmetros de perturbação substituídos por estimativas consistentes. O teste da razão de verossimilhanças baseado em verossimilhança perfilada tem sua estatística de teste equivalente à estatística da razão de verossimilhanças usual, com distribuição assintótica nula χ^2 . No mais, um ponto que devemos levar em consi-

deração é a quantidade de parâmetros de perturbação, uma vez que a verossimilhança perfilada pode fornecer estimadores de máxima verossimilhança inconsistentes e ineficientes para problemas que envolvem um grande número destes parâmetros. Para minimizar o efeito do número de parâmetros de perturbação, [Cox e Reid \(1987\)](#) propuseram uma versão modificada para a verossimilhança perfilada, a qual exige ortogonalidade entre os parâmetros de interesse e perturbação.

Como mencionado anteriormente, a estatística da razão de verossimilhanças perfiladas modificadas é assintoticamente equivalente à estatística da razão de verossimilhanças usual, isto é, tem distribuição assintótica nula χ^2 , podendo não ser bem aproximada pela distribuição qui-quadrado de referência quando o tamanho da amostra é pequeno e/ou moderado. Com o objetivo de melhorar a aproximação da distribuição da estatística deste teste pela distribuição χ^2 de referência, [DiCiccio e Stern \(1994\)](#) propuseram um fator de correção de Bartlett derivando o teste da razão de verossimilhanças perfiladas modificadas corrigido, o qual confere uma inferência mais precisa quando o tamanho da amostra não é suficientemente grande.

A metodologia das correções de Bartlett tem recebido bastante atenção na literatura e diversos trabalhos abordando esta temática têm sido propostos. Na classe dos modelos simétricos, [Ferrari e Uribe-Opazo \(2001\)](#) obtiveram um fator de correção de Bartlett para a estatística da razão de verossimilhanças para modelos lineares, sendo estes resultados estendidos para os modelos não-lineares por [Cordeiro \(2004\)](#). [Ferrari et al. \(2004\)](#) obtiveram um fator de correção de Bartlett para a estatística da razão de verossimilhanças perfiladas modificadas para o modelo normal linear heteroscedástico. Uma alternativa numérica à correção de Bartlett analítica para a estatística da razão de verossimilhanças é a correção de Bartlett bootstrap, proposta por [Roche \(1989\)](#). Tal fator de correção é obtido a partir da técnica de reamostragem bootstrap ([Efron, 1979](#)), sendo uma alternativa bastante atrativa, principalmente para casos em que é difícil calcular a correção de Bartlett analítica.

Nesta direção, nosso objetivo é obter fatores de correção de Bartlett para as estatísticas da razão de verossimilhanças e razão de verossimilhanças perfiladas modificadas na classe dos modelos não-lineares simétricos heteroscedásticos, generalizando os resultados obtidos por [Cordeiro \(2004\)](#) e [Ferrari et al. \(2004\)](#). No que segue, iremos definir o modelo não-linear simétrico heteroscedástico (MNLSH) e apresentaremos aspectos inferenciais do mesmo. Na Seção 2.3, abordaremos as correções de Bartlett analíticas e numérica e apresentaremos expressões, em notação matricial, para os fatores de correção de Bartlett

para as estatísticas da razão de verossimilhanças e razão de verossimilhanças perfiladas modificadas para o teste de heteroscedasticidade na classe dos MNLSH. Para avaliar o desempenho dos testes em estudo considerando amostras de tamanho pequeno e moderado, na Seção 2.4 iremos realizar um estudo de simulação de Monte Carlo sob diversos cenários. Ainda para o estudo de simulação, consideraremos, para efeito de comparação, os testes da razão de verossimilhanças bootstrap e razão de verossimilhanças Bartlett bootstrap. Conclusões acerca dos resultados obtidos são apresentadas na Seção 2.5.

2.2 Especificação do modelo

Suponha y_1, \dots, y_n variáveis aleatórias contínuas independentes. Dizemos que y_ℓ , $\ell = 1, \dots, n$, segue distribuição simétrica com $\mu_\ell \in \mathbb{R}$ e $\phi_\ell > 0$ parâmetros de locação e dispersão, respectivamente, se sua função de densidade é escrita como

$$\pi(y_\ell; \mu_\ell, \phi_\ell) = \frac{1}{\sqrt{\phi_\ell}} g(u_\ell), \quad y_\ell \in \mathbb{R}, \quad \ell = 1, \dots, n, \quad (2.1)$$

para alguma função $g(\cdot)$, usualmente denominada geradora de densidade, tal que $g(u_\ell) > 0$, para $u_\ell > 0$ e $\int_0^\infty u_\ell^{-1/2} g(u_\ell) du_\ell = 1$, com $u_\ell = (y_\ell - \mu_\ell)^2 / \phi_\ell$. Denotamos a variável aleatória simétrica por $y_\ell \sim S(\mu_\ell, \phi_\ell, g)$, para $\ell = 1, \dots, n$. Além da normal, a classe de distribuições simétricas contempla as distribuições normal contaminada, t -Student, t -Student generalizada, Kotz, Kotz generalizada, logística tipos I e II, logística generalizada, exponencial potência, entre outras.

Para determinar o modelo não-linear simétrico heteroscedástico, duas estruturas de regressão foram introduzidas à classe de distribuições simétricas (2.1). Assumimos para a resposta média $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^\top$ a seguinte estrutura:

$$\mu_\ell = f(x_\ell; \boldsymbol{\beta}), \quad \ell = 1, \dots, n, \quad (2.2)$$

em que $f(\cdot; \cdot)$ é uma função não-linear, contínua e duplamente diferenciável com respeito aos componentes de $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^\top$, que é um vetor de parâmetros desconhecidos a serem estimados ($p < n$), e $\boldsymbol{x}_\ell = (x_{\ell 1}, \dots, x_{\ell m})^\top$ é o vetor de m variáveis explicativas associadas à ℓ -ésima observação. Ainda, é assumido uma componente sistemática para o vetor de parâmetros de dispersão $\boldsymbol{\phi} = (\phi_1, \dots, \phi_n)^\top$ dado por

$$\phi_\ell = \sigma^2 m(\tau_\ell), \quad \ell = 1, \dots, n, \quad (2.3)$$

com $m(\tau_\ell) > 0$, sendo $\tau_\ell = \mathbf{w}_\ell^\top \boldsymbol{\delta}$, $\mathbf{w}_\ell = (w_{\ell 1}, \dots, w_{\ell k})^\top$ um vetor de variáveis explicativas que pode ter componentes comuns com \mathbf{x}_ℓ , $\boldsymbol{\delta} = (\delta_1, \dots, \delta_k)^\top$ um vetor de parâmetros desconhecidos a serem estimados e $\sigma^2 \in (0, \infty)$ uma constante desconhecida. Para efeito de simplificação, no que segue, denotaremos $m(\tau_\ell) = m_\ell$.

Assim sendo, o modelo não-linear simétrico heteroscedástico é dado por

$$y_\ell = \mu_\ell + \sqrt{\phi_\ell} \epsilon_\ell; \quad \ell = 1, \dots, n, \quad (2.4)$$

em que $\epsilon_\ell \sim S(0, 1, g)$, sendo μ_ℓ e ϕ_ℓ como definidos em (2.2) e (2.3), respectivamente.

A contribuição deste capítulo está centrada em testar a hipótese nula $H_0 : \boldsymbol{\delta} = \boldsymbol{\delta}_0$, em que $\boldsymbol{\delta}_0$ é um vetor de dimensão $k \times 1$ de constantes especificadas tal que $m(\mathbf{w}_\ell^\top \boldsymbol{\delta}_0) = 1$ para todo $\ell = 1, \dots, n$, contra $H_1 : \boldsymbol{\delta} \neq \boldsymbol{\delta}_0$. Dessa forma, o número de parâmetros de interesse é k e o número de parâmetros de perturbação é $p + 1$. O logaritmo da função de verossimilhança do vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top, \sigma^2)^\top$, dado o vetor de observações $\mathbf{y} = (y_1, \dots, y_n)^\top$, do modelo definido em (2.4) é expresso por

$$l(\boldsymbol{\theta}; \mathbf{y}) = -\frac{n}{2} \log \sigma^2 - \frac{1}{2} \sum_{\ell=1}^n \log m_\ell + \sum_{\ell=1}^n t(z_\ell),$$

com $t(z_\ell) = \log g(z_\ell^2)$, sendo $z_\ell = \sqrt{u_\ell} = \frac{(y_\ell - \mu_\ell)}{\sqrt{\phi_\ell}}$, $\ell = 1, \dots, n$. A função escore total de $\boldsymbol{\theta} = (\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top, \sigma^2)^\top$ tem a forma $\mathbf{U}(\boldsymbol{\theta}) = (\mathbf{U}_\delta^\top, \mathbf{U}_\beta^\top, U_{\sigma^2})^\top$ e seus componentes (ver Apêndice A) são expressos, em notação matricial, por

$$\begin{aligned} \mathbf{U}_\delta &= -\frac{1}{2} \boldsymbol{\Psi}^\top (\mathbf{M}\boldsymbol{\nu} - \mathbf{S}\mathbf{M}\mathbf{u}), \\ \mathbf{U}_\beta &= \tilde{\mathbf{X}}^\top \mathbf{S}\boldsymbol{\Phi}(\mathbf{y} - \boldsymbol{\mu}) \text{ e} \\ U_{\sigma^2} &= -\frac{n}{\sigma^2} + \frac{1}{2\sigma^2} (\mathbf{y} - \boldsymbol{\mu})^\top \mathbf{S}\boldsymbol{\Phi}(\mathbf{y} - \boldsymbol{\mu}), \end{aligned}$$

em que $\tilde{\mathbf{X}} = \partial \boldsymbol{\mu} / \partial \boldsymbol{\beta}^\top$, $\mathbf{S} = \text{diag}\{s_1, \dots, s_n\}$, com $s_\ell = \frac{-2g'(z_\ell^2)}{g(z_\ell^2)}$ e $g'(z_\ell^2) = \frac{\partial g(z_\ell^2)}{\partial z_\ell^2}$, $\boldsymbol{\Phi} = \text{diag}\{1/\phi_1, \dots, 1/\phi_n\}$, $\boldsymbol{\Psi}$ é uma matriz $n \times k$ com a ℓ -ésima linha dada por $\partial m_\ell / \partial \boldsymbol{\delta}^\top$, $\mathbf{M} = \text{diag}\{1/m_1, \dots, 1/m_n\}$, $\mathbf{u} = (u_1, \dots, u_n)^\top$ e $\boldsymbol{\nu}$ é um vetor $n \times 1$ de uns. Algumas das distribuições simétricas são listadas na Tabela 2.1, em que também incluímos expressões para $g(z^2)$ e s .

A matriz de informação de Fisher para o modelo resultante (ver Apêndice A) é expressa da forma

$$I = -E \left(\frac{\partial^2 l}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right) = \begin{pmatrix} \left(\frac{1 - \delta_{(0,1,0,0,2)}}{4} \right) \boldsymbol{\Psi}^\top \mathbf{M}^{(2)} \boldsymbol{\Psi} & \mathbf{0} & \left(\frac{1 - \delta_{(0,1,0,0,2)}}{4} \right) \boldsymbol{\Psi}^\top \boldsymbol{\Phi} \boldsymbol{\nu} \\ \mathbf{0} & -\delta_{(0,1,0,0,0)} \tilde{\mathbf{X}}^\top \boldsymbol{\Phi} \tilde{\mathbf{X}} & \mathbf{0} \\ \left(\frac{1 - \delta_{(0,1,0,0,2)}}{4} \right) \boldsymbol{\nu}^\top \boldsymbol{\Phi} \boldsymbol{\Psi} & \mathbf{0} & (1 - \delta_{(0,1,0,0,2)}) \frac{n}{4\sigma^4} \end{pmatrix},$$

sendo $\mathbf{M}^{(2)} = \mathbf{M} \odot \mathbf{M}$, em que \odot denota o produto Hadamard, isto é, produto elemento a elemento (Rao, 1973, pág. 30), $\delta_{(a,b,c,d,e)} = E\{t(z_\ell)^{(1)a}t(z_\ell)^{(2)b}t(z_\ell)^{(3)c}t(z_\ell)^{(4)d}z_\ell^e\}$ para $a, b, c, d, e \in \{1, 2, 3, 4\}$ e $t(z_\ell)^{(k)} = \frac{\partial^k t(z_\ell)}{\partial z_\ell^k}$, para $k = 1, 2, 3, 4$ e $\ell = 1, \dots, n$. Os valores dos δ 's correspondentes a algumas distribuições simétricas podem ser encontrados em Uribe-Opazo et al. (2008)

Tabela 2.1: Expressões para $g(z^2)$, $\frac{g'(z^2)}{g(z^2)}$ e s para algumas distribuições simétricas.

Distribuições	$g(z^2)$	$\frac{g'(z^2)}{g(z^2)}$	s
1. Normal	$\frac{\exp\{-\frac{1}{2}z^2\}}{\sqrt{2\pi}}$	$-\frac{1}{2}$	1
2. Cauchy	$\frac{1}{\pi}(1+z^2)^{-1}$	$-\frac{1}{1+z^2}$	$\frac{2}{1+z^2}$
3. t -Student	$\frac{\nu^{\frac{\nu}{2}}[\nu+z^2]^{-\frac{\nu+1}{2}}}{B(1/2, \nu/2)}$	$-\frac{\nu+1}{2(\nu+z^2)}$	$\frac{\nu+1}{(\nu+z^2)}$
4. Logística tipo I	$c \frac{\exp\{-z^2\}}{(1+\exp\{-z^2\})^2}$, $c \approx 1,484300029$	$-\frac{1-\exp\{-z^2\}}{1+\exp\{-z^2\}}$	$\frac{2(1-\exp\{-z^2\})}{1+\exp\{-z^2\}}$
5. Logística tipo II	$\frac{\exp\{-z\}}{(1+\exp\{-z\})^2}$	$-\frac{1}{2} \frac{(\exp\{ z\} -1)}{(z \exp\{ z\}+1)}$	$\frac{(\exp\{ z\} -1)}{(z \exp\{ z\}+1)}$
6. Exponencial potência	$c(k) \exp\left\{-\frac{1}{2}z^{2/(1+k)}\right\}$, $c(k)^{-1} = \Gamma\left(1 + \frac{1+k}{2}\right)2^{1+(1+k)/2}$	$-\frac{z^{2(-k/(1+k))}}{2(1+k)}$	$\frac{z^{2(-k/(1+k))}}{(1+k)}$

Na presença de parâmetros de perturbação, podemos realizar inferências com base na função de verossimilhança perfilada, sendo tais parâmetros são substituídos por suas respectivas estimativas de máxima verossimilhança para valores fixos dos parâmetros de interesse, resultando em uma função de verossimilhança que depende apenas dos parâmetros de interesse. O logaritmo da função de verossimilhança perfilada do modelo resultante é dado por

$$l_p(\boldsymbol{\delta}) = l_p(\boldsymbol{\delta}; \mathbf{y}) = l(\boldsymbol{\delta}, \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\sigma}_{\boldsymbol{\delta}}^2; \mathbf{y}), \quad (2.5)$$

sendo $\hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}$ e $\hat{\sigma}_{\boldsymbol{\delta}}^2$ os estimadores de máxima verossimilhança de $\boldsymbol{\beta}$ e σ^2 dado $\boldsymbol{\delta}$, respectivamente. Supondo válidas as condições gerais de regularidade, $\hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}$ e $\hat{\sigma}_{\boldsymbol{\delta}}^2$ são, respectivamente, soluções das equações $\mathbf{U}_{\boldsymbol{\beta}} = \mathbf{0}$ e $U_{\sigma^2} = 0$, as quais não podem ser obtidas em forma fechada. Dado isto, os estimadores de máxima verossimilhança de $\boldsymbol{\beta}$ e σ^2 restritos a $\boldsymbol{\delta}$ são obtidos computacionalmente a partir de técnicas iterativas de maximização restrita. Maiores detalhes destas técnicas podem ser obtidas em Nocedal e Wright (1999, Capítulo 18).

A obtenção do estimador de máxima verossimilhança de $\boldsymbol{\delta}$ pode ser feita maximizando o logaritmo da função de verossimilhança dado em (2.5) sujeita às restrições $\mathbf{U}_\beta = \mathbf{0}$ e $U_{\sigma^2} = 0$.

Para testar a hipótese nula de interesse, isto é, $H_0 : \boldsymbol{\delta} = \boldsymbol{\delta}_0$, contra a hipótese alternativa $H_1 : \boldsymbol{\delta} \neq \boldsymbol{\delta}_0$, sendo $\boldsymbol{\delta}_0$ um vetor $k \times 1$ de constantes especificadas, a estatística da razão de verossimilhanças (LR) é dada por

$$LR = 2\{l_p(\hat{\boldsymbol{\delta}}) - l_p(\boldsymbol{\delta}_0)\},$$

sendo $\hat{\boldsymbol{\delta}}$ o estimador de máxima verossimilhança de $\boldsymbol{\delta}$. Assintoticamente e sob a hipótese nula, LR tem distribuição χ_k^2 , sendo k o número de restrições impostas por H_0 , sob erro de ordem n^{-1} .

Ao substituir os parâmetros de perturbação por suas respectivas estimativas de máxima verossimilhança estamos, de certo modo, os tratando como conhecidos. Por esta razão, a função de verossimilhança perfilada pode apresentar viés na função escore e na informação (Ferrari et al., 2005). Ainda, este procedimento pode fornecer estimadores de máxima verossimilhança inconsistentes e ineficientes para problemas com um número grande de parâmetros de perturbação. Cox e Reid (1987) propuseram uma versão modificada da função de verossimilhança perfilada com o intuito de atenuar o efeito do número de parâmetros de perturbação, entretanto, esta versão requer ortogonalidade entre os parâmetros de interesse e os de perturbação, o que em nosso caso de estudo implica que $\boldsymbol{\delta}$ deve ser ortogonal aos demais parâmetros. Como podemos observar na matriz de informação de Fisher, $(\boldsymbol{\delta}^\top, \sigma^2)^\top$ é ortogonal a $\boldsymbol{\beta}$ e para satisfazer tal restrição é preciso que a transformação $(\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top, \sigma^2) \rightarrow (\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top, \gamma)$ nos conduza à ortogonalidade entre $\boldsymbol{\delta}^\top$ e γ , ou seja, é necessário e suficiente que $E(-\partial l / \partial \delta_a \partial \gamma) = 0$, $a = 1, \dots, k$. Seguindo a metodologia de Cox e Reid (1987, Eq. 4), temos que a transformação desejada é obtida a partir do sistema de equações diferenciais

$$\frac{n}{2\sigma^4} \frac{\partial \sigma^2}{\partial \delta_a} = -\frac{1}{2\sigma^2} \sum_{\ell=1}^n \frac{\partial m_\ell}{\partial \delta_a} \frac{1}{m_\ell}$$

que tem como solução (Simonoff e Tsai, 1994)

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

Considerando o modelo reparametrizado, o logaritmo da função de verossimilhança perfilada modificada (Cox e Reid, 1987) é dado por

$$l_{CR}^*(\boldsymbol{\delta}) = l_p^*(\boldsymbol{\delta}; \mathbf{y}) - \frac{1}{2} \log\{\det[j^*(\boldsymbol{\delta}; \hat{\boldsymbol{\beta}}_\delta, \hat{\gamma}_\delta)]\},$$

em que

$$l_p^*(\boldsymbol{\delta}; \mathbf{y}) = l^*(\boldsymbol{\delta}, \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\gamma}_{\boldsymbol{\delta}}; \mathbf{y}), \quad \text{com } \hat{\gamma}_{\boldsymbol{\delta}} = \hat{\sigma}_{\boldsymbol{\delta}}^2 \left(\prod_{\ell=1}^n m_{\ell} \right)^{1/n},$$

é o logaritmo da função de verossimilhança perfilada para $\boldsymbol{\delta}$, sendo

$$l^*(\boldsymbol{\theta}^*; \mathbf{y}) = -\frac{n}{2} \log \gamma + \sum_{\ell=1}^n t(z_{\ell})$$

o logaritmo da função de verossimilhança para o vetor de parâmetros $\boldsymbol{\theta}^* = (\boldsymbol{\beta}^{\top}, \boldsymbol{\delta}^{\top}, \gamma)^{\top}$ e $j^*(\boldsymbol{\delta}; \boldsymbol{\beta}, \gamma)$ é o bloco da matriz de informação observada para os parâmetros de perturbação $(\boldsymbol{\beta}^{\top}, \gamma)^{\top}$ avaliada em $(\boldsymbol{\delta}, \hat{\boldsymbol{\beta}}_{\boldsymbol{\delta}}, \hat{\gamma}_{\boldsymbol{\delta}})$, expresso por

$$j^*(\boldsymbol{\delta}; \hat{\boldsymbol{\beta}}, \hat{\gamma}) = - \begin{pmatrix} j_{\beta\beta}^* & \mathbf{0} \\ \mathbf{0} & j_{\gamma\gamma}^* \end{pmatrix},$$

em que $j_{\beta\beta}^*$ é uma matriz quadrada de ordem p cujas entradas são dadas por

$$j_{\beta\beta}^* = -\frac{\partial^2 l^*(\boldsymbol{\theta}^*; \mathbf{y})}{\partial \beta_j \partial \beta_l} = -\frac{1}{\gamma} \sum_{\ell=1}^n t(z_{\ell})^{(2)} q_{\ell}(j, l)_{\ell} + \frac{1}{\gamma^{1/2}} \sum_{\ell=1}^n t(z_{\ell})^{(1)} q_{\ell}^{1/2}(jl)_{\ell}$$

e

$$j_{\gamma\gamma}^* = -\frac{\partial^2 l^*(\boldsymbol{\theta}^*; \mathbf{y})}{\partial \gamma^2} = -\frac{n}{2\gamma^2} - \frac{3}{4\gamma^2} \sum_{\ell=1}^n t(z_{\ell})^{(1)} z_{\ell} - \frac{1}{4\gamma^2} \sum_{\ell=1}^n t(z_{\ell})^{(2)} z_{\ell}^2$$

é um escalar, com $q_{\ell} = \frac{(\prod_{s=1}^n m_s)^{1/n}}{m_{\ell}}$, $(j, l)_{\ell} = (\partial \mu_{\ell} / \partial \beta_j)(\partial \mu_{\ell} / \partial \beta_l)$ e $(jl)_{\ell} = \partial^2 \mu_{\ell} / \partial \beta_j \partial \beta_l$.

A estatística da razão de verossimilhanças perfiladas modificadas (LR_m) para o teste de H_0 contra H_1 é dada por

$$LR_m = 2\{l_{CR}(\hat{\boldsymbol{\delta}}) - l_{CR}(\boldsymbol{\delta}_0)\}, \quad (2.6)$$

sendo $\hat{\boldsymbol{\delta}}$ o estimador de máxima verossimilhança de $\boldsymbol{\delta}$. Sob H_0 , a estatística LR_m tem distribuição assintótica χ_k^2 , com erro de aproximação de ordem n^{-1} .

2.3 Correções de Bartlett

Quando a distribuição nula exata da estatística de teste é desconhecida ou difícil de ser obtida, é comum na literatura o uso de testes assintóticos. Os testes de razão de verossimilhanças usual e verossimilhanças perfiladas modificadas são baseados em grandes amostras e suas respectivas estatísticas de teste, LR e LR_m , têm distribuição assintótica

nula χ_k^2 , sendo k o número de restrições impostas por H_0 , sob um erro de aproximação de ordem n^{-1} . No entanto, quando o tamanho da amostra não é suficientemente grande, a aproximação da distribuição das estatísticas LR e LR_m pela distribuição χ^2 pode não ser satisfatória, conduzindo a taxas de rejeição bastante distorcidas e podendo resultar numa tomada de decisão equivocada.

Uma maneira de melhorar a aproximação das distribuições das estatísticas LR e LR_m pela distribuição χ^2 é introduzir fatores de correção de Bartlett (Bartlett, 1937; DiCiccio e Stern, 1994) à essas estatísticas de modo que o erro de aproximação das distribuições das estatísticas corrigidas pela distribuição χ^2 de referência seja menor que o erro de aproximação para as respectivas versões não corrigidas. Os fatores de correção de Bartlett não dependem de um modelo paramétrico particular, sendo bastante gerais, o que implica que suas expressões devem ser obtidas para cada problema de interesse.

Uma outra alternativa para atenuar a distorção das taxas de rejeição do teste da razão de verossimilhanças quando lidamos com amostras pequenas é aplicar a técnica bootstrap (Efron, 1979). O teste da razão de verossimilhanças bootstrap segue a metodologia descrita por Efron e Tibshirani (1993), que baseada na estatística LR possibilita encontrar a distribuição empírica da mesma, a partir da amostra observada $\mathbf{y} = (y_1, \dots, y_n)^\top$. Ainda, podemos considerar o teste da razão de verossimilhanças Bartlett bootstrap proposto por Roche (1989) como alternativa numérica do teste da razão de verossimilhanças corrigido via correção de Bartlett.

2.3.1 Correção de Bartlett para a estatística LR

Sabe-se que para grandes amostras e sob hipótese nula, a estatística LR tem distribuição qui-quadrado com erro de aproximação de ordem n^{-1} . Bartlett (1937) propôs multiplicar a estatística LR por um fator de correção $(1+c/k)^{-1}$ derivando uma estatística corrigida LR^* expressa por

$$LR^* = \frac{LR}{1 + c/k},$$

em que c é uma constante de ordem n^{-1} que pode ser estimada sob H_0 e escrita em função de momentos das derivadas do logaritmo da função de verossimilhança (Lawley, 1956). Em particular, $P(LR^* \leq x) = P(\chi_k^2 \leq x) + O(n^{-2})$ sob H_0 . Maiores detalhes sobre correção de Bartlett podem ser encontrados em Cordeiro e Cribari-Neto (2014).

Para o teste de $H_0 : \boldsymbol{\delta} = \boldsymbol{\delta}_0 \times H_1 : \boldsymbol{\delta} \neq \boldsymbol{\delta}_0$ na classe dos MNLSH considerando heteroscedasticidade com efeitos multiplicativos, isto é, $m_\ell = \exp(\boldsymbol{\omega}_\ell^\top \boldsymbol{\delta})$, temos que a

constante c do fator de correção de Bartlett para a estatística LR é dada por

$$c = \epsilon_k(\boldsymbol{\delta}) + \epsilon_{p,k}(\boldsymbol{\beta}, \boldsymbol{\delta}) + \epsilon_{p,k}(\boldsymbol{\delta}, \boldsymbol{\gamma}) + \epsilon_{p,k}(\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\gamma}), \quad (2.7)$$

sendo

$$\begin{aligned} \epsilon_k(\boldsymbol{\delta}) &= \frac{M_4}{4} \text{tr}(\mathbf{H}_d^2) + \frac{M_3^2}{6} \boldsymbol{\iota}^\top \mathbf{H}^{(3)} \boldsymbol{\iota} + \frac{M_3^2}{4} \boldsymbol{\iota}^\top \mathbf{H} \mathbf{H}_d \mathbf{H} \boldsymbol{\iota}, \\ \epsilon_{p,k}(\boldsymbol{\beta}, \boldsymbol{\delta}) &= -\frac{M_7}{4\delta_{(0,1,0,0,0)}} \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} - \frac{M_8}{4\delta_{(0,1,0,0,0)}} \boldsymbol{\iota} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} \\ &+ \frac{\delta_{(0,0,1,0,1)}}{2\delta_{(0,1,0,0,0)}} \boldsymbol{\iota} \mathbf{Q} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} + \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} \\ &- \frac{M_7}{4\delta_{(0,1,0,0,0)}} \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} - \frac{M_8}{4\delta_{(0,1,0,0,0)}} \boldsymbol{\iota} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} \\ &+ \left(\frac{M_{10}^2}{2(\delta_{(0,1,0,0,0)})^2} + \frac{M_{10}}{\delta_{(0,1,0,0,0)}} \right) \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{Z}_{\beta} \odot \mathbf{H} \odot \mathbf{Z}_{\beta} \mathbf{Q} \boldsymbol{\iota} \\ &+ \frac{M_{10}^2}{4(\delta_{(0,1,0,0,0)})^2} \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{Z}_{\beta d} \mathbf{H} \mathbf{Z}_{\beta d} \mathbf{Q} \boldsymbol{\iota} - \frac{M_3 M_{10}}{2\delta_{(0,1,0,0,0)}} \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{Z}_{\beta d} \mathbf{H} \mathbf{H}_d \boldsymbol{\iota}, \\ \epsilon_{p,k}(\boldsymbol{\delta}, \boldsymbol{\gamma}) &= -\frac{M_9 M_{11}}{2} \text{tr}(\mathbf{H}_d) + M_2 M_{11} \text{tr}(\mathbf{H}_d) - \frac{M_1^2 M_{11}}{2} \boldsymbol{\iota}^\top \mathbf{H}^{(2)} \boldsymbol{\iota} \\ &- \frac{M_1^2 M_{11}}{4} [\text{tr}(\mathbf{H}_d)]^2 + \left(\frac{M_1 M_6}{4} + \frac{M_1 (M_6 - 4M_5)}{4} \right) M_{11}^2 \text{tr}(\mathbf{H}_d) \text{ e} \\ \epsilon_{p,k}(\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\gamma}) &= \frac{M_1 M_{10} M_{11}}{4\delta_{(0,1,0,0,0)}} [(\boldsymbol{\iota}^\top \mathbf{W} \mathbf{Z}_{\beta d} \boldsymbol{\iota}) \odot (\boldsymbol{\iota}^\top \mathbf{H}_d \boldsymbol{\iota}) + \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota}], \end{aligned}$$

em que $\mathbf{Q} = \text{diag}(q_1, \dots, q_n)$, com $q_\ell = \exp\{-(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})^\top \boldsymbol{\delta}\}$ e $\bar{\boldsymbol{\omega}} = (\bar{\omega}_1, \dots, \bar{\omega}_k)^\top$, $\mathbf{W} = \partial \boldsymbol{\tau} / \partial \boldsymbol{\delta}$, $\mathbf{H} = \{h_{\ell s}\} = -(\mathbf{W} - \bar{\mathbf{W}})[(\mathbf{W} - \bar{\mathbf{W}})^\top \mathbf{V}(\mathbf{W} - \bar{\mathbf{W}})]^{-1}(\mathbf{W} - \bar{\mathbf{W}})^\top$, com $(\mathbf{W} - \bar{\mathbf{W}}) = (\mathbf{w}_1 - \bar{\mathbf{w}}, \dots, \mathbf{w}_n - \bar{\mathbf{w}})^\top$ e \mathbf{V} matriz diagonal de ordem n com entradas $v_\ell = (1 - \delta_{(0,1,0,0,2)})/4$, $\ell, s = 1, \dots, n$, $\mathbf{H}^{(2)} = (h_{\ell s}^2)$, $\mathbf{H}^{(3)} = (h_{\ell s}^3)$, $\mathbf{Z}_{\beta} = \tilde{\mathbf{X}}(\tilde{\mathbf{X}}^\top \mathbf{Q} \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^\top$, $\boldsymbol{\iota}$ vetor $n \times 1$ de uns, os subscritos d indicam que apenas os elementos da diagonal principal das matrizes foram considerados. Ainda, os escalares M_1 a M_{11} expressos por

$$\begin{aligned} M_1 &= -\frac{1}{8} \{ \delta_{(0,0,1,0,3)} + 3\delta_{(0,1,0,0,2)} + \delta_{(1,0,0,0,1)} \}, \\ M_2 &= -M_1, \quad M_3 = M_1, \\ M_4 &= \frac{1}{16} \{ \delta_{(0,0,0,1,4)} + 6\delta_{(0,0,1,0,3)} + 7\delta_{(0,1,0,0,2)} + \delta_{(1,0,0,0,1)} \}, \\ M_5 &= -\frac{n}{2} \{ 2 + 3\delta_{(1,0,0,0,1)} + \delta_{(0,1,0,0,2)} \}, \\ M_6 &= -\frac{n}{8} \{ 8 + 15\delta_{(1,0,0,0,1)} + 9\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)} \}, \end{aligned}$$

$$\begin{aligned}
M_7 &= \frac{1}{4}\{\delta_{(0,0,0,1,2)} + 3\delta_{(0,0,1,0,1)}\}, \\
M_8 &= \frac{1}{2}\{\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}\}, \\
M_9 &= \frac{1}{16}\{\delta_{(0,0,0,1,4)} + 8\delta_{(0,0,1,0,3)} + 13\delta_{(0,1,0,0,2)} + 3\delta_{(1,0,0,0,1)}\}, \\
M_{10} &= -M_8 \text{ e} \\
M_{11} &= \frac{4}{n\{2 + \delta_{(0,1,0,0,2)} + 3\delta_{(1,0,0,0,1)}\}}.
\end{aligned}$$

Podemos observar que a constante c do fator de correção envolve apenas operações simples de matrizes e podem ser facilmente implementados em pacotes de computação simbólica e linguagens de programação que permite execução de operações simples de álgebra linear, tais como, `0x` e `R`. Detalhes sobre a obtenção da constante c são apresentados no Apêndice B.

2.3.2 Correção de Bartlett para a estatística LR_m

A estatística LR_m , assim como a estatística da razão de verossimilhanças usual, tem distribuição nula assintótica χ_k^2 sob erro de aproximação de ordem n^{-1} . DiCiccio e Stern (1994) propuseram uma correção de Bartlett para esta estatística, reduzindo o erro de aproximação para ordem n^{-2} . A estatística corrigida é dada por

$$LR_m^* = \frac{LR_m}{1 + c_m/k},$$

sendo c_m uma constante de ordem n^{-1} tal que, sob H_0 , $E(LR_m^*) = k + O(n^{-3/2})$. A equação geral para c_m é definida em DiCiccio e Stern (1994, Eq. 25). Baseado nessa expressão, Ferrari et al. (2004, Eq. 5) obtiveram uma equação para o cálculo de c_m que pode ser empregada em qualquer classe de modelos que usa a partição do vetor de parâmetros tal como feita neste capítulo. No Apêndice C, obtemos com detalhes c_m para o teste de H_0 que, considerando heteroscedasticidade com efeitos multiplicativos, na classe dos MNLSH, é escrito em notação matricial como

$$\begin{aligned}
c_m &= \frac{1}{4}M_4\text{tr}(\mathbf{H}_d^2) - \frac{1}{4}M_1^2M_{11}[\text{tr}(\mathbf{H}_d)]^2 + \frac{1}{4}M_1^2\boldsymbol{\iota}^\top \mathbf{H}_d \mathbf{H} \mathbf{H}_d \boldsymbol{\iota} + \frac{1}{6}M_1^2\boldsymbol{\iota}^\top \mathbf{H}^{(3)}\boldsymbol{\iota} \\
&- M_1M_{11}\text{tr}(\mathbf{H}_d) - M_1M_5M_{11}^2\text{tr}(\mathbf{H}_d) - \frac{1}{2}M_1^2M_{11}\boldsymbol{\iota}^\top \mathbf{H}^{(2)}\boldsymbol{\iota}.
\end{aligned} \tag{2.8}$$

A constante c_m envolve apenas operações simples de matrizes, tal como a constante c do fator de correção da estatística LR . Ainda, o fator de correção c_m depende apenas da matriz de covariáveis \mathbf{W} , do número de parâmetros desconhecidos em ϕ_ℓ e do número

de observações, assim, não dependendo de parâmetros desconhecidos ou do número de parâmetros de perturbação. Além disso, pode-se observar que a não linearidade do modelo não exerce influência sobre o fator de correção c_m para a estatística da razão de verossimilhanças perfiladas modificadas para o teste de heteroscedasticidade.

2.3.3 Correção de Bartlett bootstrap para a estatística LR

Como alternativa aos testes assintóticos, pode-se realizar inferências baseadas em testes com valores críticos (p -valor) obtidos através da técnica bootstrap (Efron, 1979). O teste da razão de verossimilhanças bootstrap apresenta inferência confiável e não envolve cálculos complexos, no mais, é computacionalmente custoso. Considerando a técnica bootstrap (Efron, 1979), Roche (1989) propôs uma forma numérica alternativa ao cálculo do fator de correção de Bartlett para a estatística LR , derivando uma estatística da razão de verossimilhanças bootstrap (LR_{boot}^*) que é obtida como segue. Inicialmente geramos B amostras bootstrap (y_1^*, \dots, y_B^*) a partir do modelo assumido considerando H_0 verdadeira e substituindo os parâmetros por suas respectivas estimativas restritas calculadas usando a amostra original (y_1, \dots, y_n). Depois calculamos o valor da estatística da razão de verossimilhanças para cada pseudo-amostra bootstrap. Denotamos por LR_{boot}^b , sendo $b = 1, \dots, B$, a b -ésima amostra bootstrap. A estatística da razão de verossimilhanças Bartlett bootstrap é calculada por

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

sendo k o número de restrições impostas por H_0 e $\overline{LR}_{boot}^* = \frac{1}{B} \sum_{b=1}^B LR_{boot}^b$. Sob H_0 , a estatística LR_{boot}^* segue uma distribuição χ_k^2 , aproximadamente.

É valido salientar que a estatística da razão de verossimilhanças bootstrap (LR_{boot}) não segue distribuição χ^2 , sendo o teste baseado nesta estatística realizado como descrito a seguir:

- Para o nível de significância fixado (α), calcule o percentil $1 - \alpha$ da distribuição empírica de LR_{boot} , que é estimado pelo valor $\hat{q}_{(1-\alpha)}$ tal que

$$\frac{\#\{LR_{boot} \leq \hat{q}_{(1-\alpha)}\}}{B} = 1 - \alpha,$$

com $\#$ denotando a cardinalidade do conjunto. Rejeite a hipótese nula se $LR > \hat{q}_{1-\alpha}$. De forma alternativa, a regra de decisão pode ser escrita com base no p -valor

bootstrap dado por

$$p^* = \frac{\#\{LR_{boot}^b \geq LR\}}{B}.$$

Assim, rejeitamos H_0 para p^* menor que o nível de significância considerado.

Trabalhos recentes na literatura realizam inferência baseada nos testes da razão de verossimilhanças bootstrap e Bartlett bootstrap (Bayer e Cribari-Neto, 2013; Cribari-Neto e Queiroz, 2014). Uma vantagem de utilizar a correção de Bartlett bootstrap ao invés da técnica bootstrap usual é devido à sua eficiência computacional. Para obter o valor crítico do teste utilizando a correção de Bartlett bootstrap é necessário um número de amostras bootstrap menor do que as necessárias quando é utilizada a técnica bootstrap usual, o que implica que a correção de Bartlett bootstrap é computacionalmente mais eficiente do que a técnica bootstrap usual, ver [Rocke \(1989\)](#).

2.4 Resultados numéricos

Realizamos estudos de simulação de Monte Carlo a fim de comparar os desempenhos dos testes da razão de verossimilhanças usual e as demais versões apresentadas neste capítulo. A saber, consideramos os testes baseados nas seguintes estatísticas: razão de verossimilhanças (LR), sua versão corrigida via Bartlett (LR^*), razão de verossimilhanças perfiladas modificadas (LR_m), sua versão corrigida via Bartlett (LR_m^*), razão de verossimilhanças bootstrap (LR_{boot}) e razão de verossimilhanças Bartlett bootstrap (LR_{boot}^*). Avaliamos os desempenhos dos testes segundo a proximidade das probabilidades de rejeição da hipótese nula quando esta é verdadeira (probabilidade do erro tipo I) aos respectivos níveis nominais dos testes para cenários com um número fixo de parâmetros de interesse e de perturbação, variando o tamanho amostral e estudamos separadamente o efeito do número de parâmetros de perturbação e de interesse. Ainda, apresentamos gráficos de discrepância de quantis para alguns dos cenários simulados.

Baseamos o estudo da simulação no modelo de regressão

$$y_\ell = \beta_0 + \exp\{\beta_1 x_{\ell 1}\} + \sum_{s=2}^{p-1} \beta_s x_{\ell s} + \epsilon_\ell, \quad \ell = 1, \dots, n,$$

sendo ϵ_ℓ variáveis aleatórias independentes em que $\epsilon_\ell \sim S(0, \sigma^2 \exp\{\boldsymbol{\omega}_\ell^\top \boldsymbol{\delta}\}, g)$. As hipóteses testadas foram $H_0 : \boldsymbol{\delta} = \mathbf{0}$, sendo $\boldsymbol{\delta}$ um vetor k -dimensional, assim indicando que o

modelo é homoscedástico contra a hipótese alternativa $H_1 : \delta_i \neq 0$ para pelo menos um i , $i = 1, \dots, k$.

Para o estudo de simulação foram consideradas as distribuições simétricas t -Student com 5 graus de liberdade (ν) e exponencial potência com parâmetro de forma $\kappa = 0,3$. Lucas (1997) e Villegas et al. (2013) abordam aspectos de robustez do modelo t -Student para quando fixamos os graus de liberdade e apontam que tal modelo acomoda melhor observações aberrantes quando os graus de liberdade são fixados, por esta razão neste trabalho consideramos ν fixo. A escolha dos graus de liberdade e do parâmetro de forma das respectivas distribuições tiveram base nos coeficientes de curtose apresentados pelas distribuições para os valores dos graus de liberdade e parâmetro de forma adotados. Como nosso interesse é trabalhar com distribuições simétricas com caudas mais pesadas que as da normal (coeficiente de curtose igual a 3), escolhemos $\nu = 5$ para a t -Student e $\kappa = 0,3$ para a exponencial potência pois para estes respectivos valores dos graus de liberdade e parâmetros de forma tais distribuições apresentam curtose iguais a 9 e 3,67, respectivamente, isto é, maiores que a curtose da distribuição normal. Para os parâmetros de regressão assumimos $\beta_0 = \dots = \beta_{p-1} = 1$, $\sigma^2 = 1$, $\delta_1 = 0,1$; $\delta_2 = 0,3$; $\delta_3 = 0,5$ e $\delta_4 = \delta_5 = 1,0$. As covariáveis $x_{\ell_1}, \dots, x_{\ell_{p-1}}$ e $w_{\ell_1}, \dots, w_{\ell_k}$ foram mantidas fixas e geradas como amostras aleatórias da distribuição $U(0,1)$. O número de réplicas de Monte Carlo e bootstrap foi fixado em 10.000 e 500, respectivamente, e os níveis nominais considerados foram $\alpha = 1\%$, 5% e 10% . As simulações foram realizadas usando a linguagem de programação matricial **0x** (Doornik, 2006) e os gráficos de discrepância de quantis foram feitos usando o software estatístico **R**.

Considerando cada tamanho amostral e nível nominal estipulado, calculamos as taxas de rejeição para cada teste, isto é, estimamos via simulação de Monte Carlo $\mathbb{P}(LR \geq \chi_{(\alpha;k)}^2)$, $\mathbb{P}(LR^* \geq \chi_{(\alpha;k)}^2)$, $\mathbb{P}(LR_m \geq \chi_{(\alpha;k)}^2)$, $\mathbb{P}(LR_m^* \geq \chi_{(\alpha;k)}^2)$, $\mathbb{P}(LR_{boot}^* \geq \chi_{(\alpha;k)}^2)$, com $\chi_{(\alpha;k)}^2$ sendo o percentil $(1 - \alpha)$ da distribuição χ_k^2 . Para o teste bootstrap, a taxa de rejeição foi obtida a partir do cálculo da probabilidade $\mathbb{P}(LR \geq \hat{q}_{(1-\alpha)})$, em que $\hat{q}_{(1-\alpha)}$ é o percentil $(1 - \alpha)$ estimado da distribuição empírica de LR_{boot} . Nas Tabelas apresentadas todas as entradas são porcentagens.

As Tabelas 2.2 e 2.3 apresentam as taxas de rejeição nulas dos diferentes testes para os modelos t_5 , exponencial potência, respectivamente, para $p = 3,5$ e $k = 3$ quando o tamanho da amostra aumenta. Como podemos observar nestas Tabelas, o teste da razão de verossimilhanças é consideravelmente liberal, como exemplo, na Tabela 2.2 para $p = 3$, $\alpha = 5\%$ e considerando todos os tamanhos amostrais ($n = 30, 40, 50, 100$), as taxas

de rejeição do teste LR são iguais a 13,6%, 11,9%, 10,7% e 9,5%, respectivamente. A versão corrigida para o teste da razão verossimilhanças atenua a distorção do teste usual, embora ainda apresente taxas de rejeição maiores que os níveis nominais considerados, a exemplo, considerando o mesmo cenário anterior, as taxas de rejeição do teste LR^* para os quatro tamanhos amostrais são 7,0%, 6,6%, 5,9% e 5,4%, respectivamente. De forma geral, para ambos os testes, conforme o tamanho da amostra aumenta, as distorções dos testes diminuem.

Os resultados de simulação para o modelo t_5 apresentados na Tabela 2.2 indicam que os testes corrigidos LR_m^* e LR_{boot}^* apresentam os melhores desempenhos, sendo o teste LR_m^* o de melhor desempenho, seguido do teste LR_{boot}^* . Por exemplo, para $p = 3$ e $\alpha = 5\%$, as taxas de rejeição nulas do teste LR_m^* para os quatro tamanhos amostrais são 5%, 4,9%, 5%, 5,3% e as taxas correspondentes para o teste LR_{boot}^* são 5,3%, 5,4%, 5%, 4,4%. De forma geral, o bom desempenho do teste LR_m^* pode ser observado em todos os cenários. Na Tabela 2.3 são apresentados os resultados para o modelo exponencial potência, em que podemos observar que os testes corrigidos LR_m^* e LR_{boot}^* apresentam melhores desempenhos outra vez, sendo o teste LR_{boot}^* o de melhor performance, seguido do teste LR_m^* . Por exemplo, para $p = 5$ e $\alpha = 10\%$, as taxas de rejeição nulas do teste LR_{boot}^* para os quatro tamanhos amostrais foram 10%, 9,9%, 10%, 9,9%, enquanto que para o teste LR_m^* foram 7,2%, 9,1%, 9,9%, 10,6%.

Na Tabela 2.4 avaliamos o efeito do número de parâmetros de perturbação no desempenho dos testes, fixando o tamanho amostral ($n = 35$), o número de parâmetros de interesse ($k = 3$) e variando o número de parâmetros de perturbação ($p = 2, 3, 4, 5$) para os modelos t_5 e exponencial potência. Para ambos os modelos, os testes usuais e suas versões corrigidas são bastante distorcidos e conforme o número de parâmetros de perturbação aumenta, a distorção dos testes LR e LR^* aumenta também. Em contrapartida, o aumento do número de parâmetros de perturbação foi indiferente para os demais testes, tendo os testes corrigidos LR_m^* e LR_{boot}^* desempenhos melhores, destacando-se, de forma geral, o teste LR_{boot}^* . Por exemplo, para o modelo t_5 considerando $p = 4$ e $\alpha = 5\%$, temos as taxas de rejeição nulas dos testes iguais a 15,8% (LR), 8,8% (LR^*), 4,5% (LR_m), 4,7% (LR_m^*), 5,1% (LR_{boot}) e 5,0% (LR_{boot}^*). Considerando o mesmo cenário, temos para o modelo exponencial potência as seguintes taxas de rejeição nulas: 14,7% (LR), 8,8% (LR^*), 3,4% (LR_m), 4,1% (LR_m^*), 5,3% (LR_{boot}) e 5,1% (LR_{boot}^*).

O efeito do aumento do número de parâmetros de interesse no desempenho dos testes é avaliado na Tabela G.1. Para isto, fixamos o tamanho amostral ($n = 35$) e o

Tabela 2.2: Tamanho dos testes para o Modelo t_5 com $p = 3, 5$, $k = 3$ e diversos valores para n .

α	Teste	$p = 3$				$p = 5$			
		n				n			
		30	40	50	100	30	40	50	100
$\alpha = 10\%$	LR	21,2	19,4	18,3	16,3	33,0	25,5	19,7	14,4
	LR^*	13,2	12,5	11,4	10,7	21,1	15,9	13,8	11,6
	LR_m	9,2	9,4	9,4	10,0	8,8	10,2	9,9	9,9
	LR_m^*	9,6	9,7	9,6	10,3	9,3	10,4	10,1	10,1
	LR_{boot}	10,4	11,1	10,2	9,6	10,6	10,0	10,2	10,5
	LR_{boot}^*	10,4	10,9	10,2	9,5	10,5	10,0	10,0	10,5
$\alpha = 5\%$	LR	13,6	11,9	10,7	9,5	22,6	16,8	12,3	7,7
	LR^*	7,0	6,6	5,9	5,4	12,9	8,8	7,4	5,8
	LR_m	4,7	4,7	4,9	5,1	4,2	5,0	4,5	5,0
	LR_m^*	5,0	4,9	5,0	5,3	4,5	5,2	4,6	5,2
	LR_{boot}	5,2	5,4	5,1	4,6	5,5	5,0	5,1	5,3
	LR_{boot}^*	5,3	5,4	5,0	4,4	5,2	4,9	5,1	5,2
$\alpha = 1\%$	LR	4,7	4,0	3,4	2,6	9,5	6,5	3,8	1,9
	LR^*	1,6	1,4	1,2	1,1	3,9	2,4	1,9	1,1
	LR_m	0,9	0,9	1,0	1,0	1,0	0,9	1,0	1,0
	LR_m^*	1,0	0,9	1,0	1,0	1,0	1,0	1,0	1,0
	LR_{boot}	1,2	1,2	1,2	1,1	1,3	1,3	1,1	1,2
	LR_{boot}^*	1,1	1,0	1,1	0,9	1,1	1,2	1,0	1,0

número de parâmetros de perturbação ($p=3$), variando o número de parâmetros de interesse ($k=2,3,4,5$). O desempenho dos testes foi semelhante ao apresentado na Tabela 2.4, tendo o teste LR_{boot}^* apresentado taxas mais próximas aos níveis nominais adotados quando comparado aos demais testes corrigidos, principalmente para o modelo exponencial potência. Por exemplo, considerando o cenário em que $k = 3$ e $\alpha = 1\%$, temos as taxas de rejeição dos testes para o modelo exponencial potência iguais a 3,9% (LR), 1,1% (LR^*), 0,7% (LR_m), 0,9% (LR_m^*), 1,2% (LR_{boot}) e 1,0% (LR_{boot}^*). Considerando o mesmo cenário, para o modelo t_5 temos taxas de rejeição iguais a 4% (LR), 1,8% (LR^*), 0,8% (LR_m), 0,8% (LR_m^*), 1,1% (LR_{boot}) e 0,9% (LR_{boot}^*).

Os resultados numéricos apresentados nas Tabelas 2.2-G.1 evidenciam a superioridade no desempenho dos testes corrigidos perante os não corrigidos em amostras de tamanho pequeno e moderado. De modo geral, o teste Bartlett bootstrap teve melhor desempenho,

Tabela 2.3: Tamanho dos testes para o Modelo exponencial potência com $\kappa = 0, 3$, $p = 3, 5$, $k = 3$ e diversos valores para n .

α	Teste	$p = 3$				$p = 5$			
		n				n			
		30	40	50	100	30	40	50	100
$\alpha = 10\%$	LR	22,6	17,2	14,9	12,1	32,0	22,2	20,2	14,2
	LR^*	14,9	12,8	11,6	10,4	22,4	18,6	16,2	13,7
	LR_m	7,7	8,9	9,3	9,7	6,1	8,3	9,1	10,2
	LR_m^*	8,8	9,7	10,2	10,0	7,2	9,1	9,9	10,6
	LR_{boot}	10,7	10,5	10,0	10,0	10,0	10,1	10,0	10,0
	LR_{boot}^*	10,6	10,4	10,0	9,9	10,0	9,9	10,0	9,9
$\alpha = 5\%$	LR	14,1	10,1	8,6	6,4	22,8	14,2	12,0	7,8
	LR^*	8,5	6,9	6,1	5,5	14,0	10,8	9,3	7,4
	LR_m	3,5	4,4	4,7	4,8	2,6	3,8	4,2	4,7
	LR_m^*	4,2	5,0	5,2	5,1	3,3	4,3	4,8	4,9
	LR_{boot}	5,4	5,1	5,2	5,0	5,3	5,4	5,1	5,0
	LR_{boot}^*	5,5	4,9	4,9	5,1	5,1	5,1	5,1	5,0
$\alpha = 1\%$	LR	4,7	3,0	2,4	1,0	9,6	4,8	3,6	1,9
	LR^*	2,2	1,5	1,4	1,1	4,7	3,2	2,8	1,9
	LR_m	0,4	0,8	1,2	1,0	0,4	0,6	0,7	0,9
	LR_m^*	0,6	1,0	1,3	1,1	0,5	0,7	0,8	1,0
	LR_{boot}	1,3	1,2	1,3	1,1	1,3	1,1	1,3	1,1
	LR_{boot}^*	1,2	1,0	1,0	1,1	1,1	1,0	1,1	1,0

seguido do teste da razão de verossimilhanças perfiladas modificadas corrigido. O teste da razão de verossimilhanças corrigido atenua as distorções do teste usual, embora ainda apresente taxas de rejeição um pouco distorcidas, sobretudo quando aumentamos o número de parâmetros nos modelos.

Nas Figuras 2.1 - 2.4 construímos o gráfico do quociente da diferença entre o quantil exato (estimado por simulação) e o quantil assintótico (χ^2) pelo quantil assintótico para os modelos t_5 (Figuras 2.1 e 2.2) e exponencial potência (Figuras 2.3 e 2.4) considerando $n = 35$, $p = 3, 5$ e $k = 3$. Consideramos apenas os testes corrigidos, uma vez que os testes baseados nas estatísticas LR e LR_m apresentaram comportamento liberal e conservativo, respectivamente, indicando que tais estatísticas não têm suas respectivas distribuições bem aproximadas da distribuição χ^2 de referência. Dos gráficos, quanto mais próxima a curva de discrepância estiver da ordenada nula, melhor é a aproximação da distribuição

Figura 2.1: Discrepância relativa de quantis para o modelo t_5 com $n = 35$, $p = 3$, $k = 3$

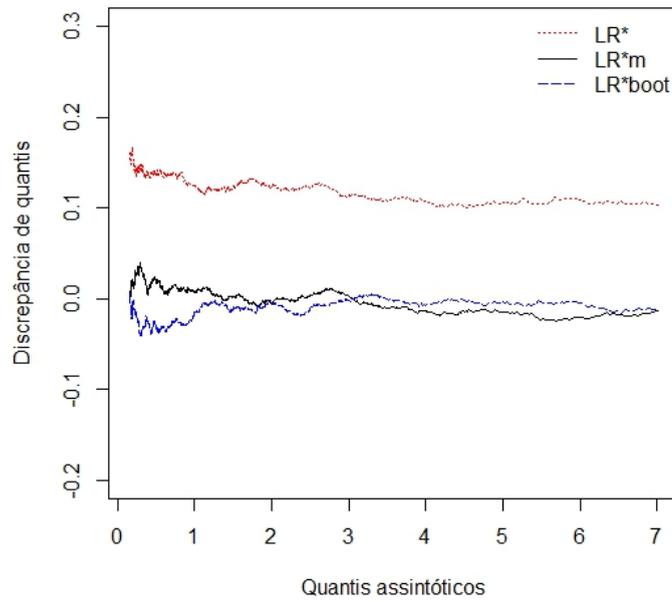


Figura 2.2: Discrepância relativa de quantis para o modelo t_5 com $n = 35$, $p = 5$, $k = 3$

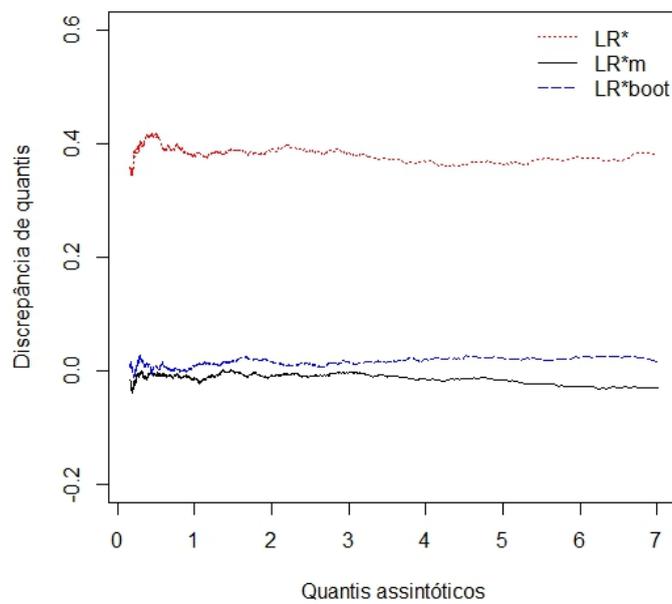


Figura 2.3: Discrepância relativa de quantis para o modelo exponencial potência com $\kappa = 0,3$, considerando $n = 35$, $p = 3$, $k = 3$

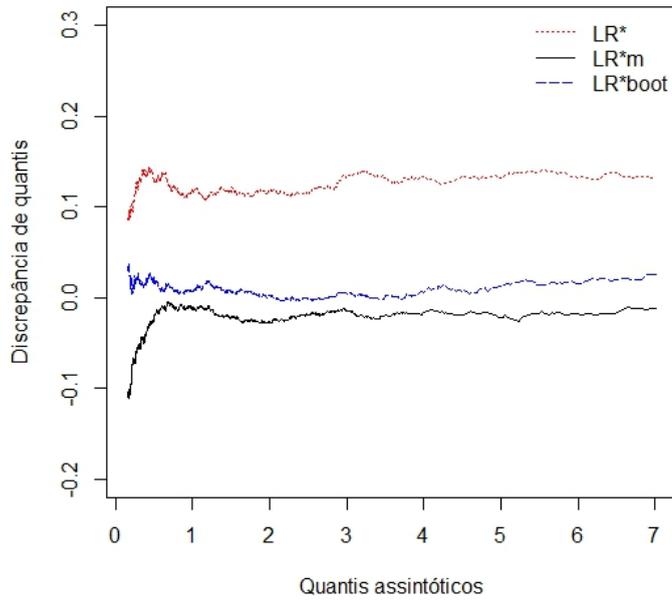


Figura 2.4: Discrepância relativa de quantis para o modelo exponencial potência com $\kappa = 0,3$, considerando $n = 35$, $p = 5$, $k = 3$

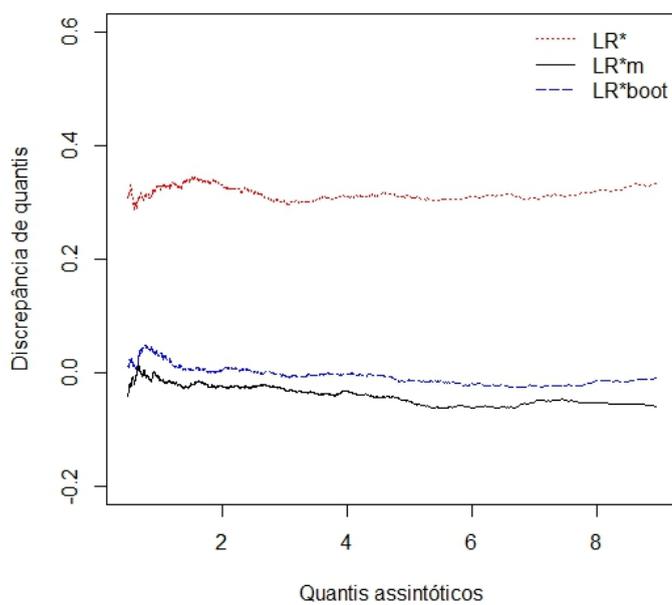


Tabela 2.4: Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0,3$, considerando $n = 35$, $k = 3$ e diversos valores para p .

α	<i>Teste</i>	Modelo t_5				Modelo exponencial potência			
		p				p			
		2	3	4	5	2	3	4	5
10%	<i>LR</i>	16,3	19,4	24,5	29,1	15,4	19,0	23,0	28,6
	<i>LR*</i>	11,4	12,9	15,4	17,9	12,3	13,7	15,8	18,8
	<i>LR_m</i>	9,4	9,4	9,7	9,0	8,5	8,3	7,6	7,5
	<i>LR_m*</i>	9,7	9,7	10,0	9,3	9,4	9,6	8,5	8,7
	<i>LR_{boot}</i>	9,6	9,8	10,4	10,3	10,4	9,8	10,4	9,5
	<i>LR_{boot}*</i>	9,6	9,8	10,2	10,0	10,2	9,7	10,3	9,4
5%	<i>LR</i>	9,3	11,9	15,8	19,4	8,9	11,5	14,7	19,1
	<i>LR*</i>	5,9	7,0	8,8	10,5	6,4	7,5	8,8	11,4
	<i>LR_m</i>	4,5	4,7	4,5	4,3	4,2	3,8	3,4	3,6
	<i>LR_m*</i>	4,8	4,9	4,7	4,5	5,0	4,6	4,1	4,3
	<i>LR_{boot}</i>	4,9	4,9	5,1	5,1	5,6	5,0	5,3	5,1
	<i>LR_{boot}*</i>	4,8	4,9	5,0	4,8	5,5	4,9	5,1	5,0
1%	<i>LR</i>	2,8	4,0	5,8	7,5	2,7	3,9	5,3	7,9
	<i>LR*</i>	1,2	1,8	2,3	3,2	1,5	1,1	2,4	3,8
	<i>LR_m</i>	0,7	0,9	1,0	0,9	0,8	0,7	0,6	0,6
	<i>LR_m*</i>	0,8	0,9	1,0	1,0	1,0	0,9	0,7	0,7
	<i>LR_{boot}</i>	0,9	1,1	1,2	1,2	1,0	1,2	1,4	1,2
	<i>LR_{boot}*</i>	0,8	0,9	1,0	1,0	0,9	1,1	1,2	1,0

nula da estatística de teste pela distribuição χ^2 de referência. Para ambos modelos é notório que a distribuição nula da estatística da razão de verossimilhanças corrigida não é bem aproximada pela distribuição χ^2 de referência em todos os cenários considerados, uma vez que as curvas de discrepância estão bem acima da ordenada nula. Em contrapartida, a curva de discrepância das estatísticas da razão de verossimilhanças perfiladas modificadas corrigida e Bartlett bootstrap se aproximaram bem da ordenada nula, indicando boa aproximação de suas respectivas distribuições nula pela distribuição χ^2 de referência, o que foi refletido no desempenho dos testes apresentados nas Tabelas 2.2-G.1.

Tabela 2.5: Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0,3$, considerando $n = 35$, $p = 3$ e diversos valores para k .

α	<i>Teste</i>	Modelo t_5				Modelo exponencial potência			
		k				k			
		2	3	4	5	2	3	4	5
10%	<i>LR</i>	16,0	19,5	21,5	23,9	16,4	18,9	20,0	21,8
	<i>LR*</i>	11,4	12,3	13,7	15,2	12,0	13,7	14,1	15,4
	<i>LR_m</i>	9,1	9,1	9,7	10,2	9,0	8,3	8,0	7,3
	<i>LR_m*</i>	9,4	9,4	10,0	10,3	10,0	9,5	9,0	8,5
	<i>LR_{boot}</i>	9,9	9,8	10,4	10,5	10,2	9,8	10,4	10,7
	<i>LR_{boot}*</i>	9,8	9,9	10,2	10,4	10,2	9,7	10,2	10,8
5%	<i>LR</i>	9,2	11,8	13,0	15,0	9,6	11,5	12,1	13,4
	<i>LR*</i>	6,0	7,0	7,2	8,6	6,2	7,5	7,7	8,9
	<i>LR_m</i>	4,7	4,4	4,5	5,4	4,8	3,8	3,7	3,5
	<i>LR_m*</i>	5,0	4,6	4,7	5,4	5,0	4,6	4,4	4,2
	<i>LR_{boot}</i>	5,0	4,9	5,1	5,3	5,2	5,0	5,3	5,6
	<i>LR_{boot}*</i>	4,9	4,9	5,0	5,2	5,1	4,9	5,0	5,6
1%	<i>LR</i>	2,9	4,0	4,1	5,6	2,9	3,9	4,0	4,6
	<i>LR*</i>	1,4	1,8	1,7	2,2	1,5	1,1	2,0	2,5
	<i>LR_m</i>	0,9	0,8	0,9	1,0	0,8	0,7	0,7	0,6
	<i>LR_m*</i>	1,0	0,8	1,0	1,0	1,1	0,9	0,9	0,8
	<i>LR_{boot}</i>	1,1	1,1	1,2	1,3	1,2	1,2	1,4	1,3
	<i>LR_{boot}*</i>	0,9	0,9	1,0	1,2	1,0	1,0	1,2	1,3

2.5 Conclusões

Para amostras de tamanho pequeno e até moderado, a distribuição da estatística da razão de verossimilhanças não é bem aproximada da distribuição χ^2 limite, como consequência, o teste apresenta taxas de rejeição bastante distorcidas, sendo assim importante desenvolver estratégias de inferência mais acuradas para quando o tamanho da amostra não for grande o suficiente.

Neste capítulo apresentamos correções de Bartlett para os testes da razão de verossimilhanças baseados na verossimilhança usual e verossimilhança perfilada modificada. Nosso trabalho generaliza os resultados apresentados por [Cordeiro \(2004\)](#), que obteve um fator de correção de Bartlett para a estatística da razão de verossimilhanças para o modelo linear simétrico e [Ferrari et al. \(2004\)](#), que obteve um teste da razão de verossimilhanças

perfiladas modificadas para o modelo normal linear. Afim de comparar o desempenho dos testes, também foram considerados os testes da razão de verossimilhanças bootstrap e Bartlett bootstrap.

Os resultados numéricos mostraram que o teste da razão de verossimilhanças usual é bastante liberal e sua versão corrigida via Bartlett, embora atenuie tal tendência, ainda apresenta taxas de rejeição distorcidas conforme o número de parâmetros no modelo aumenta. O teste da razão de verossimilhanças perfiladas modificadas apresenta desempenho superior ao teste usual e não sofre efeito do aumento do número de parâmetros no modelo. Em particular, pode-se observar também a superioridade do desempenho dos testes da razão de verossimilhanças perfiladas modificadas corrigido e Bartlett bootstrap em amostras de tamanho pequeno e moderado, dessa maneira, em aplicações práticas, sugerimos o uso destes testes.

Refinamento dos testes escore e gradiente

3.1 Introdução

Assim como o teste da razão de verossimilhanças, os testes escore e gradiente também podem ser aprimorados, derivando testes cuja distribuição de suas respectivas estatísticas é melhor aproximada da distribuição qui-quadrado de referência. Para obter os testes escore e gradiente melhorados, são incorporados fatores de correção, denominados fatores de correção tipo-Bartlett, às estatísticas destes testes. Os fatores de correção para as estatísticas escore e gradiente foram obtidos, respectivamente, por [Cordeiro e Ferrari \(1991\)](#) e [Vargas et al. \(2013\)](#).

Diversos trabalhos na literatura têm abordado esta temática, dentre eles, [Cysneiros et al. \(2010b\)](#) derivaram um fator de correção tipo-Bartlett para a estatística escore na classe dos modelos não-lineares simétricos, tendo seus resultados posteriormente generalizados por [Nascimento \(2010\)](#), que aprimorou o teste escore para os modelos não-lineares simétricos heteroscedásticos e avaliou numericamente apenas os testes sob os parâmetros de locação. [Vargas et al. \(2014\)](#) obtiveram um fator de correção tipo-Bartlett para a estatística gradiente na classe dos modelos lineares generalizados. Nesta direção, iremos obter um fator de correção tipo-Bartlett para a estatística gradiente para o teste de heteroscedasticidade na classe dos modelos não-lineares simétricos heteroscedásticos (MNLSH). Este trabalho é pioneiro no sentido de aprimorar o teste gradiente para modelos com erros seguindo distribuição simétrica. No mais, diferentemente das estatísticas da razão de verossimilhanças e escore, a estatística gradiente não é invariante sob reparametriza-

ções não-lineares (Terrell, 2002). Desse modo, neste capítulo não iremos utilizar a mesma parametrização do Capítulo 2. Aqui, nosso vetor de parâmetros será $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$, onde faremos testes para um subconjunto $\boldsymbol{\delta}_1$ de $\boldsymbol{\delta}$. No que segue, apresentamos o modelo não-linear simétrico heteroscedástico seguindo a nova parametrização e seus aspectos inferenciais. Na Seção 3.3, discorreremos sobre as correções tipo-Bartlett para as estatísticas escore e gradiente, obtendo expressões em notação matricial para os fatores de correção tipo-Bartlett para o teste de heteroscedasticidade na classe dos MNLSH. Ainda, consideramos estatísticas de teste alternativas à estatística escore corrigida. Com o intuito avaliar o desempenho dos testes sob investigação, na Seção 3.4 será realizado um estudo de simulação de Monte Carlo sob diversos cenários. Para efeito de comparação, iremos considerar no estudo de simulação os testes da razão de verossimilhanças usual e corrigido, sendo o segundo obtido para a classe dos MNLSH, considerando a parametrização adotada neste capítulo, por Brito (2009). Na Seção 3.5 apresentamos uma aplicação a dados reais. Considerações sobre os resultados obtidos são apresentados na Seção 3.6.

3.2 Aspectos inferenciais

Sejam y_1, \dots, y_n variáveis aleatórias contínuas independentes simétricas (2.1). O modelo de interesse é definido por

$$y_\ell = \mu_\ell + \sqrt{\phi_\ell^*} \epsilon_\ell; \quad \epsilon_\ell \sim S(0, 1, g), \quad \ell = 1, \dots, n, \quad (3.1)$$

sendo μ_ℓ como definido em (2.2) e ϕ_ℓ^* definido por

$$\phi_\ell^* = m(\tau_\ell^*), \quad \ell = 1, \dots, n, \quad (3.2)$$

com $m(\tau_\ell^*) > 0$, sendo $\tau_\ell^* = \mathbf{w}_\ell^{*\top} \boldsymbol{\delta}$, $\mathbf{w}_\ell^* = (1, w_{\ell 1}, \dots, w_{\ell k-1})^\top$ um vetor de variáveis explicativas e $\boldsymbol{\delta} = (\delta_0, \delta_1, \dots, \delta_{k-1})^\top$ um vetor de parâmetros desconhecidos a serem estimados.

O logaritmo da função de verossimilhança do vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top)^\top$, dado os valores observados y_1, \dots, y_n , do modelo definido em (3.1) é expresso por

$$l(\boldsymbol{\theta}; \mathbf{y}) = -\frac{1}{2} \sum_{\ell=1}^n \log(\phi_\ell^*) + \sum_{\ell=1}^n t(z_\ell),$$

com $t(z_\ell)$ e z_ℓ tal como definidos na Seção 2.2. Ainda, assumimos $l(\boldsymbol{\theta}; \mathbf{y})$ regular com respeito às derivadas dos componentes de $\boldsymbol{\beta}$ e $\boldsymbol{\delta}$ até quarta ordem (Cox e Hinkley, 1974, Cap. 9).

As funções escore de $\boldsymbol{\delta}$ e $\boldsymbol{\beta}$ são dadas em notação matricial, respectivamente, por

$$\mathbf{U}_\delta = \frac{1}{2} \mathbf{W}^\top \boldsymbol{\Lambda} (\mathbf{S} \mathbf{F}_1 \mathbf{u} - \mathbf{F}_1 \boldsymbol{\nu}) \quad \text{e} \quad \mathbf{U}_\beta = \tilde{\mathbf{X}}^\top \mathbf{S} \boldsymbol{\Lambda} (\mathbf{y} - \boldsymbol{\mu}),$$

em que $\tilde{\mathbf{X}} = \partial \boldsymbol{\mu} / \partial \boldsymbol{\beta}$, $\mathbf{W} = \partial \boldsymbol{\tau}^* / \partial \boldsymbol{\delta}$, $\boldsymbol{\Lambda} = \text{diag}\{1/\phi_1^*, \dots, 1/\phi_n^*\}$, $\mathbf{S} = \text{diag}\{s_1, \dots, s_n\}$, com $s_\ell = -\frac{2g'(u_\ell)}{g(u_\ell)}$, $\mathbf{F}_1 = \text{diag}\{h'_1, \dots, h'_n\}$, com $h'_\ell = \partial \phi_\ell^* / \partial \tau_\ell$, para $\ell = 1, \dots, n$, $\mathbf{u} = (u_1, \dots, u_n)^\top$ e $\boldsymbol{\nu}$ um vetor de $n \times 1$ de uns. Características de algumas distribuições simétricas são apresentadas na Tabela 2.1.

A matriz de informação de Fisher para $\boldsymbol{\theta}$ tem estrutura bloco-diagonal, $\mathbf{K}_\theta = \text{diag}\{\mathbf{K}_\delta, \mathbf{K}_\beta\}$, onde

$$\mathbf{K}_\delta = \mathbf{W}^\top \mathbf{V} \mathbf{W} \quad \text{e} \quad \mathbf{K}_\beta = -\delta_{(0,1,0,0,0)} \tilde{\mathbf{X}}^\top \boldsymbol{\Lambda}^{-1} \tilde{\mathbf{X}},$$

com $\mathbf{V} = \text{diag}\{v_1, \dots, v_n\}$, sendo $v_\ell = ((1 - \delta_{(0,1,0,0,0)})h'_\ell{}^2)/4\phi_\ell^{*2}$ e $\delta_{(a,b,c,d,e)} = E\{t(z_\ell)^{(1)a}, t(z_\ell)^{(2)b}, t(z_\ell)^{(3)c}, t(z_\ell)^{(4)d} z_\ell^e\}$ para $a, b, c, d, e \in \{1, 2, 3, 4\}$ e $t(z_\ell)^{(k)} = \partial^k t(z_\ell) / \partial z_\ell^k$, para $k = 1, 2, 3, 4$ e $\ell = 1, \dots, n$. Os valores dos δ 's variam dentre as distribuições simétricas, em Uribe-Opazo et al. (2008) podem ser encontrados estes valores. Os parâmetros $\boldsymbol{\delta}$ e $\boldsymbol{\beta}$ são globalmente ortogonais (Cox e Reid, 1987) e seus respectivos estimadores de máxima verossimilhança são assintoticamente independentes. As estimativas de máxima verossimilhança de $\hat{\boldsymbol{\delta}}$ e $\hat{\boldsymbol{\beta}}$ podem ser obtidas iterativamente a partir do método scoring-Fisher, (ver Cysneiros et al. (2010)).

Considere a partição $\boldsymbol{\delta} = (\delta_0, \boldsymbol{\delta}_1)^\top$, em que δ_0 é um escalar e $\boldsymbol{\delta}_1 = (\delta_1, \dots, \delta_{k-1})^\top$. Estamos interessados em testar a hipótese nula $H_0 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_1^{(0)}$ contra a hipótese alternativa $H_1 : \boldsymbol{\delta}_1 \neq \boldsymbol{\delta}_1^{(0)}$, sendo $\boldsymbol{\delta}_1^{(0)}$ um vetor de dimensão $(k-1) \times 1$ de constantes especificadas tal que $m(\mathbf{w}_\ell^{*\top} \boldsymbol{\delta}_1^{(0)}) = 1$, ou seja, sob H_0 temos $\phi_\ell^* = m(\delta_0)$, indicando que o modelo é homoscedástico.

A partição em $\boldsymbol{\delta}$ induz as partições na matriz \mathbf{W} , função escore \mathbf{U}_δ e matriz informação de Fisher \mathbf{K}_δ expressas, respectivamente, por: $\mathbf{W} = (\mathbf{W}_0, \mathbf{W}_1)$, sendo $\mathbf{W}_0 = \mathbf{1}$ um vetor $n \times 1$ de uns e $\mathbf{W}_1 = \partial \boldsymbol{\tau}^* / \partial \boldsymbol{\delta}_1$, $\mathbf{U}_\delta = (\mathbf{U}_{\delta_0}^\top, \mathbf{U}_{\delta_1}^\top)^\top$, com $\mathbf{U}_{\delta_0} = \frac{1}{2} \mathbf{W}_0^\top \boldsymbol{\Lambda} (\mathbf{S} \mathbf{F}_1 \mathbf{u} - \mathbf{F}_1 \boldsymbol{\nu})$, $\mathbf{U}_{\delta_1} = \frac{1}{2} \mathbf{W}_1^\top \boldsymbol{\Lambda} (\mathbf{S} \mathbf{F}_1 \mathbf{u} - \mathbf{F}_1 \boldsymbol{\nu})$ e

$$\mathbf{K}_\delta = \begin{bmatrix} K_{\delta_0 \delta_0} & \mathbf{K}_{\delta_0 \delta_1} \\ \mathbf{K}_{\delta_1 \delta_0} & \mathbf{K}_{\delta_1 \delta_1} \end{bmatrix}$$

com $K_{\delta_0 \delta_0} = \mathbf{W}_0^\top \mathbf{V} \mathbf{W}_0$, $\mathbf{K}_{\delta_0 \delta_1}^\top = \mathbf{K}_{\delta_1 \delta_0} = \mathbf{W}_1^\top \mathbf{V} \mathbf{W}_0$ e $\mathbf{K}_{\delta_1 \delta_1} = \mathbf{W}_1^\top \mathbf{V} \mathbf{W}_1$.

A inversa da matriz de informação de Fisher é dada por $\mathbf{K}_\theta^{-1} = \text{diag}\{\mathbf{K}_\delta^{-1}, \mathbf{K}_\beta^{-1}\}$. Utilizando a fórmula de inversa de matrizes simétricas particionadas (Rao, 1973, pág. 33),

após alguma álgebra, chegamos a

$$\mathbf{K}_\delta^{-1} = \begin{bmatrix} K^{\delta_0\delta_0} & \mathbf{K}^{\delta_0\delta_1} \\ \mathbf{K}^{\delta_1\delta_0} & \mathbf{K}^{\delta_1\delta_1} \end{bmatrix},$$

sendo $\mathbf{K}^{\delta_1\delta_1} = (\mathbf{R}^\top \mathbf{V} \mathbf{R})^{-1}$, $\mathbf{K}^{\delta_1\delta_0} = \mathbf{K}^{\delta_0\delta_1\top} = (\mathbf{R}^\top \mathbf{V} \mathbf{R})^{-1} \mathbf{C}^\top$ e $K^{\delta_0\delta_0} = (\mathbf{W}_0^\top \mathbf{V} \mathbf{W}_0)^{-1} + \mathbf{C}(\mathbf{R}^\top \mathbf{V} \mathbf{R})^{-1} \mathbf{C}^\top$, com

$$\mathbf{R} = \mathbf{W}_1 - \mathbf{W}_0 \mathbf{C} \text{ e } \mathbf{C} = (\mathbf{W}_0^\top \mathbf{V} \mathbf{W}_0)^{-1} (\mathbf{W}_0^{-1} \mathbf{V} \mathbf{W}_1).$$

A fim de testar H_0 contra H_1 na classe dos MNLSH, consideremos os testes da razão de verossimilhanças, escore e gradiente, cujas estatísticas de teste são dadas, respectivamente, por

$$\begin{aligned} LR &= 2\{l(\hat{\boldsymbol{\delta}}_1, \hat{\delta}_0, \hat{\boldsymbol{\beta}}) - l(\boldsymbol{\delta}_1^{(0)}, \tilde{\delta}_0, \tilde{\boldsymbol{\beta}})\}, \\ S_r &= \frac{1}{4}[\mathbf{W}_1 \tilde{\boldsymbol{\Lambda}}(\tilde{\mathbf{S}} \tilde{\mathbf{F}}_1 \tilde{\mathbf{u}} - \tilde{\mathbf{F}}_1 \boldsymbol{\iota})]^\top (\tilde{\mathbf{R}}^\top \tilde{\mathbf{V}} \tilde{\mathbf{R}})^{-1} [\mathbf{W}_1 \tilde{\boldsymbol{\Lambda}}(\tilde{\mathbf{S}} \tilde{\mathbf{F}}_1 \tilde{\mathbf{u}} - \tilde{\mathbf{F}}_1 \boldsymbol{\iota})] \text{ e} \\ S_g &= \frac{1}{2}[\mathbf{W}_1 \tilde{\boldsymbol{\Lambda}}(\tilde{\mathbf{S}} \tilde{\mathbf{F}}_1 \tilde{\mathbf{u}} - \tilde{\mathbf{F}}_1 \boldsymbol{\iota})]^\top (\hat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}_1^{(0)}), \end{aligned}$$

em que $\hat{\boldsymbol{\delta}}_1$, $\hat{\delta}_0$, $\hat{\boldsymbol{\beta}}$ denotam os estimadores de máxima verossimilhança irrestritos de $\boldsymbol{\delta}_1$, δ_0 e $\boldsymbol{\beta}$, respectivamente, e $\tilde{\delta}_0$, $\tilde{\boldsymbol{\beta}}$ os estimadores restritos de δ_0 e $\boldsymbol{\beta}$, respectivamente. Os sinais ‘ $\hat{\cdot}$ ’ e ‘ $\tilde{\cdot}$ ’ denotam que as quantidades foram avaliadas, respectivamente, sob os estimadores de máxima verossimilhança irrestrito e restrito dos parâmetros. Sob H_0 e para grandes amostras, as três estatísticas, LR , S_r e S_g , são equivalentes e têm distribuição χ_{k-1}^2 , em que $(k-1)$ é o número de restrições impostas por H_0 , sob um erro de ordem n^{-1} (Sen e Singer, 1993; Lemonte e Ferrari, 2012).

3.3 Correção tipo-Bartlett

Como vimos no Capítulo 2, Bartlett (1937) propôs um fator de correção que ao ser incorporado à estatística LR melhora a aproximação da distribuição da estatística pela χ^2 de referência, isto é, o erro de aproximação é reduzido. A fim de aprimorar o teste escore, Cordeiro e Ferrari (1991) propuseram um fator de correção para a estatística desse teste, produzindo uma estatística corrigida cuja distribuição nula é χ^2 com erro de aproximação de ordem n^{-2} . Esse fator de correção é um polinômio na própria estatística e por tal razão não é um fator de Bartlett genuíno, sendo denominado na literatura por fator tipo-Bartlett. Recentemente Vargas et al. (2013) obtiveram um fator de correção tipo-Bartlett

para a estatística gradiente considerando os resultados de [Cordeiro e Ferrari \(1991\)](#). A estatística gradiente corrigida tem distribuição nula χ^2 , agora com erro de aproximação de ordem n^{-2} , ([Vargas et al., 2013](#)).

3.3.1 Correção tipo-Bartlett para a estatística S_r

Como mencionado anteriormente, a estatística escore (S_r) tem distribuição nula assintótica χ_{k-1}^2 , com $(k-1)$ sendo o número de restrições impostas por H_0 . Além disso, para pequenas amostras a aproximação da distribuição de S_r pela distribuição qui-quadrado de referência pode não ser satisfatória, conduzindo a testes distorcidos. Para contornar tal problema, [Cordeiro e Ferrari \(1991\)](#) propuseram um fator de correção tipo-Bartlett que quando incorporado à estatística S_r produz uma estatística corrigida cuja distribuição é melhor aproximada pela distribuição χ^2 do que a versão usual.

A estatística escore corrigida (S_r^*) é definida por

$$S_r^* = S_r \{1 - (c + bS_r + aS_r^2)\},$$

com os coeficientes

$$a = \frac{A_3}{12(k-1)((k-1)+2)((k-1)+4)}, \quad b = \frac{A_2 - 2A_3}{12(k-1)((k-1)+21)} \quad \text{e} \quad c = \frac{A_1 - A_2 + A_3}{12(k-1)}$$

de ordem n^{-1} . As quantidades A_1, A_2 e A_3 (ver [Cordeiro, 1999](#)) são funções de momentos das derivadas do logaritmo da função de verossimilhança do modelo sob investigação e dependem das hipóteses testadas. A estatística S_r^* tem distribuição nula $\chi_{(k-1)}^2$ sob erro de aproximação de ordem n^{-2} . No caso de S_r^* envolver parâmetros desconhecidos, estes devem ser substituídos pelos respectivos estimadores de máxima verossimilhança e isto não afeta a ordem de aproximação da distribuição S_r^* pela distribuição χ^2 ([Cordeiro e Ferrari, 1991](#)).

A estatística S_r^* pode ser uma função não-monótona de S_r . Em virtude disto, algumas alternativas à esta estatística foram propostas na literatura. [Kakisawa \(1996\)](#) sugeriu uma transformação monótona envolvendo a estatística S_r e os coeficientes a, b e c . A transformação é dada por $S_{r_2}^* = S_r^* + P(S_r)$, com

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

um polinômio de grau 5 em S_r . Ainda, [Cordeiro et al. \(1998\)](#) obtiveram uma estatística modificada ($S_{r_1}^*$) para melhorar a estatística escore que é função monótona de S_r . A

estatística alternativa $S_{r_1}^*$ é dada por

$$S_{r_1}^* = \begin{cases} \sqrt{\frac{\pi}{3a}} \exp\left(\frac{b^2}{3a} - c\right) \times \\ \left\{ \Phi\left(\sqrt{6a}S_r + \sqrt{\frac{2}{3}b}\right) - \Phi\left(\sqrt{\frac{2}{3}b}\right) \right\} & \text{se } a > 0, \\ \frac{1}{2b} \exp(-c) \{1 - \exp(-2bS_r)\} & \text{se } a = 0 \text{ e } b \neq 0 \end{cases}$$

Se $a = b = 0$, S_r^* é uma transformação monótona de S_r e não há necessidade de definir uma estatística alternativa. As três estatísticas S_r^* , $S_{r_1}^*$ e $S_{r_2}^*$ são equivalentes até segunda ordem, isto é, elas diferem em termos de ordem $O_p(n^{-3/2})$ (Cordeiro, 2003).

Para testar $H_0 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_1^{(0)}$ contra $H_1 : \boldsymbol{\delta}_1 \neq \boldsymbol{\delta}_1^{(0)}$ na classe dos MNLSH considerando heteroscedasticidade com efeitos multiplicativos, isto é, $m(\tau_\ell^*) = \exp(\tau_\ell^*)$, temos que os termos A_1 , A_2 e A_3 são dados por:

$$\begin{aligned} A_1 &= 3Q_2^2 \boldsymbol{\nu}^\top \boldsymbol{\Lambda} \mathbf{Z}_{\beta_d} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\beta_d} \boldsymbol{\Lambda} \boldsymbol{\nu} + 3Q_1 Q_2 \boldsymbol{\nu}^\top \boldsymbol{\Lambda} \mathbf{Z}_{\beta_d} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\delta_{0d}} \boldsymbol{\nu} \\ &+ 3Q_1^2 \boldsymbol{\nu}^\top \mathbf{Z}_{\delta_{0d}} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\delta_{0d}} \boldsymbol{\nu} + 3Q_1 Q_2 \boldsymbol{\nu}^\top \boldsymbol{\Lambda} \mathbf{Z}_{\beta_d} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\delta_{0d}} \boldsymbol{\nu} \\ &+ 12Q_1 Q_2 \boldsymbol{\nu}^\top \boldsymbol{\Lambda} \mathbf{Z}_{\beta_d} \mathbf{Z}_{\delta_0} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \boldsymbol{\nu} + 12Q_1^2 \boldsymbol{\nu}^\top \mathbf{Z}_{\delta_{0d}} \mathbf{Z}_{\delta_0} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \boldsymbol{\nu} \\ &+ 6Q_2 (2\delta_{(0,1,0,0,0)} + 3Q_2) \boldsymbol{\nu}^\top \boldsymbol{\Lambda} \mathbf{Z}_{\beta}^{(2)} \odot (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \boldsymbol{\Lambda} \boldsymbol{\nu} + 18Q_1^2 \boldsymbol{\nu}^\top \mathbf{Z}_{\delta_0}^{(2)} \odot (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \boldsymbol{\nu} \\ &- 6(N_1 - 3N_2 + N_5) \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{Z}_{\delta_{0d}}\} - 6(N_3 - N_4 + N_6) \text{tr}\{\boldsymbol{\Lambda} \mathbf{Z}_{\beta_d} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d\}, \end{aligned}$$

$$\begin{aligned} A_2 &= -12Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{Z}_{\delta_0} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \boldsymbol{\nu} - 12Q_1 Q_2 \boldsymbol{\nu}^\top \boldsymbol{\Lambda} \mathbf{Z}_{\beta_d} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \boldsymbol{\nu} \\ &- 12Q_1^2 \boldsymbol{\nu}^\top \mathbf{Z}_{\delta_{0d}} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \boldsymbol{\nu} - 24Q_1^2 \boldsymbol{\nu}^\top \mathbf{Z}_{\delta_0} \odot (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})^{(2)} \boldsymbol{\nu} \\ &+ 3(N_1 - 3N_2) \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d^{(2)}\} \text{ e} \end{aligned}$$

$$A_3 = 12Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \boldsymbol{\nu} + 8Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})^{(3)} \boldsymbol{\nu},$$

com $\mathbf{Z}_\beta = \tilde{\mathbf{X}} (\tilde{\mathbf{X}}^\top \boldsymbol{\Lambda} \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^\top$, $\mathbf{Z}_\beta^{(2)} = \mathbf{Z}_\beta \odot \mathbf{Z}_\beta$, $\mathbf{Z}_{\beta_d} = \text{diag}\{z_{\beta_{11}}, \dots, z_{\beta_{nn}}\}$, $\mathbf{Z}_\delta = \mathbf{W} (\mathbf{W}^\top \mathbf{V} \mathbf{W})^{-1} \mathbf{W}^\top$, $\mathbf{Z}_{\delta_d} = \text{diag}\{z_{\delta_{11}}, \dots, z_{\delta_{nn}}\}$, $\mathbf{Z}_{\delta_0} = \mathbf{W}_0 (\mathbf{W}_0^\top \mathbf{V} \mathbf{W}_0)^{-1} \mathbf{W}_0^\top$, $\mathbf{Z}_{\delta_{0d}} = \text{diag}\{z_{0_{11}}, \dots, z_{0_{nn}}\}$, $(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d = \mathbf{Z}_{\delta_d} - \mathbf{Z}_{\delta_{0d}}$, \odot denota o produto Hadamard, isto é, produto elemento a elemento (Rao, 1973, pág. 30), e $Q_1, Q_2, N_1, N_2, N_3, N_4, N_5$ são escalares

iguais a

$$\begin{aligned}
Q_1 &= \frac{1}{8}\{1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\}, \\
Q_2 &= -\frac{1}{2}\{\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}\}, \\
N_1 &= \frac{1}{16}\{1 + 4\delta_{(1,0,0,0,1)} + 6\delta_{(2,0,0,0,2)} + 4\delta_{(3,0,0,0,3)} + \delta_{(4,0,0,0,4)}\}, \\
N_2 &= \frac{1}{16}\{1 + 4\delta_{(1,0,0,0,1)} + 2\delta_{(2,0,0,0,2)} + 4\delta_{(1,0,0,0,1)}\delta_{(2,0,0,0,2)} + 4\delta_{(1,0,0,0,1)}^2 + \delta_{(2,0,0,0,2)}^2\}, \\
N_3 &= \frac{1}{4}\{\delta_{(2,0,0,0,0)} + 2\delta_{(3,0,0,0,1)} + \delta_{(4,0,0,0,2)}\}, \\
N_4 &= \frac{\delta_{(2,0,0,0,0)}}{4}\{1 + 2\delta_{(1,0,0,0,1)} + \delta_{(2,0,0,0,2)}\}, \\
N_5 &= \frac{1}{16}\{2\delta_{(1,1,0,0,3)} - \delta_{(0,1,0,0,2)} + \delta_{(3,0,0,0,3)} + 2\delta_{(2,0,0,0,2)} + \delta_{(2,1,0,0,4)} + (\delta_{(0,1,0,0,1)})^2\} e \\
N_6 &= \frac{1}{4}\{2\delta_{(1,1,0,0,1)} + \delta_{(2,1,0,0,2)} + \delta_{(0,1,0,0,0)}\delta_{(0,1,0,0,2)}\}.
\end{aligned}$$

Os termos A_1 , A_2 e A_3 foram calculadas em Nascimento (2010) e os δ 's necessários são apresentados em Uribe-Opazo et al. (2008). Os momentos das derivadas do logaritmo da função de verossimilhança necessários ao cálculo de A_1 , A_2 e A_3 são apresentadas no Apêndice D.

3.3.2 Correção tipo-Bartlett para a estatística S_g

O teste gradiente proposto por Terrell (2002) aparece na literatura como alternativa aos testes assintóticos clássicos. Além de ter sua estatística de teste (S_g) assintoticamente equivalente às dos testes da razão de verossimilhanças, score e Wald, a estatística gradiente é muito simples de ser computada, não envolvendo nem a matriz de informação observada, nem a esperada.

Assim como as estatísticas de teste LR e S_r , para pequenas amostras a estatística S_g pode apresentar distribuição mal aproximada pela distribuição χ^2 de referência. Para superar esta dificuldade, Vargas et al. (2013) propuseram um fator de correção para a estatística S_g , de modo que sua versão corrigida (S_g^*) apresenta distribuição nula $\chi_{(k-1)}^2$, em que $(k-1)$ é o número de restrições impostas por H_0 , sob erro de aproximação de ordem n^{-2} .

Antes de definir S_g^* , consideremos algumas notações adicionais: Dado o vetor de parâmetros desconhecidos $\boldsymbol{\theta} = (\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top)^\top$, com $\boldsymbol{\delta} = (\delta_0, \boldsymbol{\delta}_1)^\top$, sendo $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^\top$, δ_0 um escalar e $\boldsymbol{\delta}_1 = (\delta_1, \dots, \delta_{k-1})^\top$, seja $\kappa_{rs} = E[\partial^2 l(\boldsymbol{\theta}) / \partial \theta_r \partial \theta_s]$, $\kappa_{rst} = E[\partial^3 l(\boldsymbol{\theta}) / \partial \theta_r \partial \theta_s \partial \theta_t]$,

etc, $(\kappa_{rs})_t = \partial\kappa_{rs}/\partial\theta_t^*$, $(\kappa_{rs})_{tu} = \partial^2\kappa_{rs}/\partial\theta_t^*\partial\theta_u^*$, etc. O teste que estamos interessados em realizar é

$$H_0 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_1^{(0)} \times H_1 : \boldsymbol{\delta}_1 \neq \boldsymbol{\delta}_1^{(0)}, \quad (3.3)$$

em que $\boldsymbol{\delta}_1^{(0)}$ é um vetor de dimensão $(k-1) \times 1$ de constantes especificadas. A matriz de informação de Fisher de $\boldsymbol{\theta}$ é dada por

$$\mathbf{K}_\theta = \begin{bmatrix} \mathbf{K}_\delta & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_\beta \end{bmatrix}, \quad \text{com} \quad \mathbf{K}_\delta = \begin{bmatrix} K_{\delta_0\delta_0} & \mathbf{K}_{\delta_0\delta_1} \\ \mathbf{K}_{\delta_1\delta_0} & \mathbf{K}_{\delta_1\delta_1} \end{bmatrix}.$$

Seja \mathbf{K}_θ^{-1} a notação para a inversa de \mathbf{K}_θ . Definamos as matrizes

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_\delta & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_\beta^{-1} \end{bmatrix}, \quad \text{com} \quad \mathbf{A}_\delta = \begin{bmatrix} K_{\delta_0\delta_0}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

e $\mathbf{M} = \mathbf{K}_\theta^{-1} - \mathbf{A}$.

A estatística gradiente corrigida (S_g^*) é definida como (Vargas, 2013, Capítulo 3)

$$S_g^* = S_g \{1 - (c_g + b_g S_g + a_g S_g^2)\},$$

sendo

$$a_g = \frac{A_3^g}{12(k-1)((k-1)+2)((k-1)+4)}, \quad b_g = \frac{A_2^g - 2A_3^g}{12(k-1)((k-1)+21)},$$

$$c_g = \frac{A_1^g - A_2^g + A_3^g}{12(k-1)}.$$

Os termos A_1^g , A_2^g e A_3^g são funções dos momentos das derivadas do logaritmo da função de verossimilhança do modelo sob investigação, sendo expressos, respectivamente, por

$$\begin{aligned} A_1^g &= 3 \sum' \kappa_{jrs} \kappa_{klu} [m^{ju} a^{lu} (m^{sk} + 2a^{sk}) + a^{jr} m^{sk} a^{lu} + 2m^{jk} a^{rl} a^{su}] \\ &- 12 \sum' (\kappa_{jr})_s (\kappa_{kl})_u (\kappa^{sj} \kappa^{rk} \kappa^{lu} + a^{sj} a^{rk} a^{lu} + \kappa^{sk} \kappa^{lj} \kappa^{ru} + a^{sk} a^{lj} a^{ru}) \\ &- 6 \sum' \kappa_{jrs} (\kappa_{kl})_u [(a^{su} - \kappa^{su}) (\kappa^{jk} \kappa^{lr} - a^{jk} a^{lr}) + m^{jr} (a^{sk} a^{lu} + \kappa^{su} \kappa^{lu}) \\ &+ 2a^{rs} (\kappa^{jk} \kappa^{lu} - a^{jk} a^{lu}) + 2a^{rk} a^{ls} m^{ju}] \\ &+ 6 \sum' \kappa_{jr su} m^{jr} a^{su} - 6 \sum' \kappa_{jr s}^{(u)} [m^{jr} (a^{su} - \kappa^{su}) + 2m^{ju} a^{rs}] \\ &+ 12 \sum' \kappa_{rs}^{(ju)} (\kappa^{jr} \kappa^{su} - a^{jr} a^{su}), \end{aligned}$$

$$\begin{aligned}
A_2^g &= -3 \sum' \kappa_{jrs} \kappa_{klu} [m^{jr} m^{sk} a^{lu} + m^{jr} a^{sk} m^{lu} + 2m^{jk} m^{rl} a^{su}] \\
&+ \frac{1}{4} (3m^{jr} m^{sk} m^{lu} + 2m^{jk} m^{rl} m^{su}) \\
&+ 6 \sum' \kappa_{trs} (\kappa_{kv})_u [m^{su} (\kappa^{tk} \kappa^{vr} - a^{tk} a^{vr}) + m^{tr} (\kappa^{sk} \kappa^{vu} - a^{sk} a^{vu})] \\
&+ 6 \sum' (\kappa_{jrs})_u m^{jr} m^{su} - 3 \sum' \kappa_{jr su} m^{jr} m^{su} \text{ e} \\
A_3^g &= \frac{1}{4} \sum' \kappa_{jrs} \kappa_{klu} (3m^{jr} m^{sk} m^{lu} + 2m^{jk} m^{rl} m^{su}),
\end{aligned}$$

com j, r, s, k, l, u variando sob todos os parâmetros do modelo, sendo $\kappa^{rs}, a^{rs}, m^{rs}$ os elementos (rs) da matrizes $\mathbf{K}_\theta^{-1}, \mathbf{A}, \mathbf{M}$, respectivamente, e os somatórios \sum' denotando que a soma foi realizada sob todos os parâmetros de θ . Para a classe dos MNLSH, os momentos das derivadas do logaritmo da função de verossimilhança necessárias ao cálculo de A_1^g, A_2^g e A_3^g são dados no Apêndice D.

Para testar H_0 contra H_1 tal como definido em (3.3) na classe dos MNLSH considerando heteroscedasticidade com efeitos multiplicativos, isto é, $m(\tau_\ell^*) = \exp(\tau_\ell^*)$, os termos A_1^g, A_2^g e A_3^g do fator de correção tipo-Bartlett para a estatística S_g são dados por

$$\begin{aligned}
A_1^g &= 12\delta_{(0,1,0,0,0)} Q_2 \boldsymbol{\nu}^\top \boldsymbol{\Lambda} \mathbf{Z}_\beta^{(2)} \odot (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \boldsymbol{\Lambda} \boldsymbol{\nu} + 3Q_2^2 \boldsymbol{\nu}^\top \boldsymbol{\Lambda} \mathbf{Z}_{\beta_d} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\beta_d} \boldsymbol{\Lambda} \boldsymbol{\nu} \\
&+ 6Q_2^2 \boldsymbol{\nu}^\top \boldsymbol{\Lambda} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \odot \mathbf{Z}_\beta^{(2)} \boldsymbol{\Lambda} \boldsymbol{\nu} + 3Q_1 Q_2 \boldsymbol{\nu}^\top \boldsymbol{\Lambda} \mathbf{Z}_{\beta_d} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\delta_{0d}} \boldsymbol{\nu} \\
&+ 3Q_1 Q_2 \boldsymbol{\nu}^\top \mathbf{Z}_{\delta_{0d}} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\beta_d} \boldsymbol{\Lambda} \boldsymbol{\nu} + 3Q_1 Q_2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\beta_d} \boldsymbol{\Lambda} \boldsymbol{\nu} \\
&+ 6Q_1 Q_2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{Z}_{\delta_0} \mathbf{Z}_{\beta_d} \boldsymbol{\Lambda} \boldsymbol{\nu} + 3Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\delta_{0d}} \boldsymbol{\nu} \\
&+ 6Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{Z}_{\delta_0} \mathbf{Z}_{\delta_{0d}} \boldsymbol{\nu} + 3Q_1^2 \boldsymbol{\nu}^\top \mathbf{Z}_{\delta_{0d}} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\delta_{0d}} \boldsymbol{\nu} \\
&+ 6Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \odot \mathbf{Z}_{\delta_0}^{(2)} \boldsymbol{\nu} + 6Q_3 \text{tr}\{\mathbf{Z}_{\delta_{0d}} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d\} \\
&- 12Q_5 \text{tr}\{\boldsymbol{\Lambda} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{Z}_{\beta_d}\} + 6Q_4 \text{tr}\{\boldsymbol{\Lambda} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{Z}_{\beta_d}\},
\end{aligned}$$

$$\begin{aligned}
A_2^g &= -3Q_1 Q_3 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\beta_d} \boldsymbol{\Lambda} \boldsymbol{\nu} - 3Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) \mathbf{Z}_{\delta_{0d}} \boldsymbol{\nu} \\
&- 3Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{Z}_{\delta_0} (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \boldsymbol{\nu} - 6Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})^{(2)} \odot \mathbf{Z}_{\delta_0} \boldsymbol{\nu} \\
&- \frac{9}{4} Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \boldsymbol{\nu} - \frac{3}{2} Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})^{(3)} \boldsymbol{\nu} \\
&- 3Q_3 \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d^{(2)}\} \text{ e}
\end{aligned}$$

$$A_3^g = \frac{3}{4} Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \boldsymbol{\nu} + \frac{1}{2} Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})^{(3)} \boldsymbol{\nu},$$

com Q_3, Q_4, Q_5 escalares iguais a

$$\begin{aligned} Q_3 &= \frac{1}{16} \{7\delta_{(0,1,0,0,2)} - 1 + 6\delta_{(0,0,1,0,3)} + \delta_{(0,0,0,1,4)}\}, \\ Q_4 &= \frac{1}{4} \{\delta_{(0,0,0,1,2)} + 5\delta_{(0,0,1,0,1)} + 4\delta_{(0,1,0,0,0)}\} \text{ e} \\ Q_5 &= -Q_2. \end{aligned}$$

Os termos A_1^g, A_2^g e A_3^g do fator de correção tipo-Bartlett para a estatística S_g são calculados com detalhes no Apêndice E, sendo os valores dos δ 's necessários ao cálculo apresentados em [Uribe-Opazo et al. \(2008\)](#).

3.4 Resultados numéricos

Considerando os testes baseados nas estatísticas da razão de verossimilhanças (LR), sua versão corrigida via Bartlett (LR^*), score (S_r), sua versão corrigida via correção tipo-Bartlett (S_r^*), as versões modificadas propostas por [Cordeiro et al. \(1998\)](#) ($S_{r_1}^*$) e por [Kakisawa \(1996\)](#) ($S_{r_2}^*$), gradiente (S_g) e sua versão corrigida via correção tipo-Bartlett (S_g^*), realizamos um estudo de simulação de Monte Carlo onde avaliamos os seus respectivos desempenhos segundo à proximidade entre as probabilidades de rejeição da hipótese nula quando esta era assumida verdadeira (probabilidade do erro tipo I) e os respectivos níveis nominais. Para realizar este estudo, consideramos diversos cenários, a saber: fixamos o número de parâmetros de interesse e de perturbação, variando o tamanho amostral e estudamos separadamente o efeito do número de parâmetros de perturbação e interesse fixado o tamanho amostral. Ainda, apresentamos, para alguns cenários simulados, os gráficos de discrepância de quantis e avaliamos os poderes dos testes em estudo.

Para o estudo de simulação, consideramos o modelo não-linear simétrico heteroscedástico

$$y_\ell = \beta_0 + \exp\{\beta_1 x_{\ell 1}\} + \sum_{s=2}^{p-1} \beta_s x_{\ell s} + \epsilon_\ell, \quad \ell = 1, \dots, n,$$

sendo ϵ_ℓ variáveis aleatórias independentes tais que $\epsilon_\ell \sim S(0, \exp\{\mathbf{w}_\ell^{*\top} \boldsymbol{\delta}\}, g)$. O teste de interesse é $H_0 : \boldsymbol{\delta}_1 = \mathbf{0}$, sendo $\boldsymbol{\delta}_1$ um vetor de dimensão $k - 1$, contra $H_1 : \delta_i \neq 0$ para pelo menos um $i, i = 1, \dots, k - 1$. Sob H_0 , temos que $\exp\{\mathbf{w}_\ell^\top \boldsymbol{\delta}\} = \exp\{\delta_0\}$, indicando que o modelo é homoscedástico.

As distribuições simétricas consideradas no estudo de simulação foram a t -Student com 5 graus de liberdade (ν) e exponencial potência com parâmetro de forma $\kappa = 0, 3$.

A escolha dos graus de liberdade e do valor do parâmetro de forma seguiram o objetivo do estudo, que é trabalhar com distribuições com caudas mais pesadas que as da normal (curtose igual a 3). Assim, escolhemos $\nu = 5$ e $\kappa = 0,3$ devido o coeficiente de curtose apresentado pelas distribuições t -Student e exponencial potência para tais valores adotados para os graus de liberdade e parâmetro de forma, que eram, respectivamente iguais a 9 e 3,67, isto é, maiores que o da distribuição normal.

Para os parâmetros do modelo, assumimos $\beta_0 = \dots = \beta_{p-1} = 1$, $\delta_1 = 0,1$; $\delta_2 = 0,3$; $\delta_3 = 0,5$ e $\delta_4 = \delta_5 = \delta_6 = 1$. As covariáveis $x_{\ell 1}, \dots, x_{\ell p-1}$ e $w_{\ell 2}, \dots, w_{\ell k-1}$ foram geradas da distribuição $U(0,1)$. Fixamos o número de réplicas de Monte Carlo em 10.000 e consideramos os níveis nominais $\alpha = 1\%$, 5% e 10% . Para as simulações, usamos a linguagem de programação matricial **Ox** (Doornik, 2006). Os gráficos de discrepância de quantis foram feitos usando o software estatístico **R**.

Nas Tabelas a seguir, apresetamos as taxas de rejeição dos testes estudados para cada tamanho amostral e nível de significância estipulado. As taxas de rejeição são obtidas via simulação de Monte Carlo, onde estimamos $\mathbb{P}(T \geq \chi_{\alpha; k-1}^2)$, com $T = LR, LR^*, S_r, S_r^*, S_{r_1}^*, S_{r_2}^*, S_g, S_g^*$ e $\chi_{\alpha; k-1}^2$ sendo o percentil $(1 - \alpha)$ da distribuição χ_{k-1}^2 . Todas as entradas das Tabelas são porcentagem.

Nas Tabelas 3.1 e 3.2 avaliamos o comportamento dos testes para os modelos t_5 e exponencial potência, respectivamente, quando o tamanho da amostra aumenta, fixando o número de parâmetros ($p = 3, 5$ e $k = 3$). Em todos os cenários estudados pode-se observar que o teste da razão de verossimilhanças usual (LR) é bastante liberal, isto é, rejeita mais do que deveria. Quando o tamanho da amostra (n) aumenta, as taxas de rejeição se tornam próximas dos níveis nominais considerados, embora o teste ainda apresente caráter liberal, por exemplo, quando $p = 3$ e $\alpha = 10\%$, temos para o modelo t_5 as taxas de rejeição de LR para $n = 20$ e 50 , iguais a 26,7% e 14%, respectivamente. O teste baseado na versão corrigida (LR^*) apresenta comportamento semelhante, embora as taxas sejam mais próximas aos níveis nominais adotados, evidenciando a eficácia da correção. Por exemplo, ainda para o modelo t_5 , $p = 3$ e $\alpha = 10\%$, as taxas de rejeição do teste LR^* para $n = 20$ e 50 são, respectivamente, 14,7% e 11,3%. O teste escore (S_r) apresentou taxas de rejeição bastante próximas aos níveis nominais adotados, independente do tamanho amostral. De maneira geral, o teste S_r tem caráter ligeiramente liberal, principalmente para o caso em que $p = 5$, sendo tal tendência atenuada por sua versão corrigida (S_r^*). As versões modificadas do teste S_r , os testes $S_{r_1}^*$ e $S_{r_2}^*$, apresentam comportamento bastante similar ao teste S_r^* . Como exemplo, para o modelo exponencial potência, quando $p = 5$ e $\alpha = 5\%$,

temos as taxas de rejeição dos testes para $n = 30$ iguais a 5,6% (S_r), 4,9% (S_r^* , $S_{r_1}^*$ e $S_{r_2}^*$) e para $n = 50$, 5,3% (S_r), 4,9% (S_r^* , $S_{r_1}^*$ e $S_{r_2}^*$). O teste gradiente (S_g) apresenta comportamento semelhante ao teste LR , no mais, com taxas menos distorcidas. O teste baseado na sua versão corrigida (S_g^*) atenua o comportamento extremamente liberal do teste S_g , aproximando a taxa de rejeição do teste aos níveis nominais adotados. Para o modelo t_5 com $p = 3$ e $\alpha = 1\%$, quando $n = 20$, temos os tamanhos dos testes dados por 5,3% (S_g) e 1,6% (S_g^*) e para $n = 40$, temos 2,1% (S_g) e 1,1% (S_g^*). Considerando o mesmo cenário para o modelo exponencial potência, temos para $n = 20$ as taxas iguais a 5,1% (S_g) e 1,3% (S_g^*) e para $n = 40$, 2,1% (S_g) e 1,2% (S_g^*).

Na Tabela 3.3 avaliamos o efeito do número de parâmetros de perturbação no desempenho dos testes, para isto, é fixado o tamanho amostral ($n = 35$), o número de parâmetros δ ($k = 3$) e variando o número de parâmetros de perturbação ($p = 3, 4, 5, 6$). A partir dos resultados obtidos podemos observar que o teste LR é bastante sensível ao aumento do número de parâmetros de perturbação, isto é, conforme p aumenta, as taxas de rejeição deste teste se tornam mais distorcidas, o deixando extremamente liberal. Observamos ainda que a tendência liberal do teste LR é atenuada por sua versão corrigida LR^* , a ver, para o modelo t_5 e $\alpha = 10\%$, temos as taxas de rejeição para $k = 3$ iguais a 16,9% (LR), 11,9% (LR^*) e quando $k = 5$, iguais a 23,1% (LR), 13,5% (LR^*), e no modelo exponencial potência para $k = 3$, 15,8% (LR), 12,4% (LR^*) e para $k = 5$, 22,9% (LR), 16,2% (LR^*). No que diz respeito ao comportamento dos demais testes conforme o número de parâmetros de perturbação aumenta, com exceção do teste S_g , que teve comportamento semelhante ao teste LR , nenhum deles se apresentou sensível ao aumento do número de parâmetros de perturbação. Ainda, observa-se que o teste S_r , dentre os três testes não corrigidos, apresenta melhor desempenho, com taxas de rejeição bem próximas aos níveis nominais adotados, embora ligeiramente acima dos níveis nominais adotados. O teste S_r^* corrige a tendência liberal do teste S_r , aproximando ainda mais as taxas de rejeição dos níveis nominais considerados. Observa-se também que os testes $S_{r_1}^*$ e $S_{r_2}^*$ têm desempenhos semelhantes ao teste S_r^* . A exemplo, para $\alpha = 10\%$ e $k = 5$, temos as taxas de rejeição dos testes iguais a 11,7% (S_r), 10,4% (S_r^* , $S_{r_1}^*$, $S_{r_2}^*$) para o modelo t_5 , e 10,5% (S_r^*), 10,1% (S_r^* , $S_{r_1}^*$, $S_{r_2}^*$) para o modelo exponencial potência. Como mencionado anteriormente, o teste gradiente sofre influência do número de parâmetros de perturbação, isto é, conforme p aumenta, também aumenta a taxa de rejeição do teste S_g . Além disso, tal teste tem comportamento extremamente liberal, tal como o teste LR . O teste gradiente corrigido S_g^* atenua a tendência liberal do teste S_g , como era esperado, e, de modo

Tabela 3.1: Tamanho dos testes para o Modelo t_5 com $n = 35$, $p = 3, 5$, $k = 3$ e diversos valores para n .

α	<i>Teste</i>	$p = 3$				$p = 5$			
		n				n			
		20	30	40	50	20	30	40	50
$\alpha = 10\%$	<i>LR</i>	26,7	18,4	15,8	14,0	41,2	28,9	19,0	19,2
	<i>LR*</i>	14,7	12,0	11,7	11,3	23,2	16,7	12,9	12,4
	<i>S_r</i>	11,0	10,4	10,8	10,3	13,7	11,9	10,8	11,0
	<i>S_r*</i>	10,0	9,8	10,2	9,8	11,3	10,5	10,0	10,2
	<i>S_{r1}*</i>	10,1	9,8	10,2	9,8	11,4	10,5	10,1	10,2
	<i>S_{r2}*</i>	10,1	9,8	10,2	9,8	11,4	10,5	10,1	10,2
	<i>S_g</i>	24,1	17,1	15,2	13,5	38,1	27,1	18,4	18,7
	<i>S_g*</i>	10,6	10,5	9,9	9,8	13,1	13,2	10,6	11,4
$\alpha = 5\%$	<i>LR</i>	17,7	11,0	9,9	7,7	31,7	19,8	11,5	11,4
	<i>LR*</i>	8,2	6,4	6,1	5,7	15,4	9,6	6,5	6,3
	<i>S_r</i>	5,7	5,0	5,5	5,0	7,5	6,1	5,0	5,4
	<i>S_r*</i>	5,5	4,8	5,3	4,8	6,2	5,1	4,9	5,2
	<i>S_{r1}*</i>	5,5	4,8	5,3	4,8	6,3	5,1	4,9	5,2
	<i>S_{r2}*</i>	5,5	4,8	5,3	4,8	6,3	5,1	4,9	5,2
	<i>S_g</i>	15,2	9,7	8,8	7,3	28,3	17,9	10,9	11,0
	<i>S_g*</i>	5,6	5,4	4,9	5,0	8,0	6,9	5,6	6,1
$\alpha = 1\%$	<i>LR</i>	6,5	3,2	2,6	2,0	17,4	7,7	3,7	3,7
	<i>LR*</i>	2,0	1,3	1,2	1,3	5,5	2,6	1,5	1,3
	<i>S_r</i>	0,9	0,8	0,9	1,8	2,8	2,0	1,7	2,1
	<i>S_r*</i>	1,1	0,8	0,9	1,1	1,2	0,6	0,8	1,4
	<i>S_{r1}*</i>	1,1	0,8	0,9	1,1	1,3	0,8	0,8	1,4
	<i>S_{r2}*</i>	1,1	0,8	0,9	1,1	1,3	0,7	0,8	1,4
	<i>S_g</i>	5,3	2,4	2,1	1,7	15,2	6,3	3,3	3,3
	<i>S_g*</i>	1,6	1,0	1,1	1,0	2,8	1,7	1,3	1,5

geral, não sofre influência do número de parâmetros de perturbação, apresentando taxas próximas aos níveis nominais adotados independentemente do valor de p . Por exemplo, para $p = 3$ e $\alpha = 5\%$, temos as taxas de rejeição para o modelo t_5 iguais a 9,5% (S_g), 6,1% (S_g^*), e para o modelo exponencial potência 8,6% (S_g), 5,8% (S_g^*), enquanto que para $p = 5$ e $\alpha = 5\%$, temos os tamanhos dos testes para o modelo t_5 iguais a 14% (S_g), 6,4% (S_g^*), e para o modelo exponencial potência 13,3% (S_g) e 6% (S_g^*).

A fim de estudar o efeito do número de parâmetros de interesse nos testes, fixamos o

Tabela 3.2: Tamanho dos testes para o Modelo exponencial potencia com $\kappa = 0.3$, $p = 3, 5$, $k = 3$ e n .

α	Teste	$p = 3$				$p = 5$			
		n				n			
		20	30	40	50	20	30	40	50
$\alpha = 10\%$	LR	25,2	18,6	14,4	13,8	37,7	26,7	20,3	16,6
	LR^*	16,7	13,9	12,6	10,9	24,4	18,2	14,3	12,3
	S_r	9,7	9,9	10,4	9,7	12,4	10,8	10,3	10,6
	S_r^*	10,3	10,2	10,2	9,7	11,3	10,2	10,0	10,1
	$S_{r_1}^*$	10,3	10,2	10,3	9,7	11,3	10,2	10,0	10,2
	$S_{r_2}^*$	10,3	10,2	10,3	9,7	11,3	10,2	10,0	10,2
	S_g	23,0	18,0	13,9	13,5	34,9	24,6	19,2	16,2
	S_g^*	11,3	11,4	9,9	10,2	10,8	12,1	10,6	10,2
$\alpha = 5\%$	LR	16,3	11,1	8,2	7,3	27,9	18,0	12,5	9,7
	LR^*	9,6	7,7	6,6	5,8	16,2	10,9	7,8	6,8
	S_R	4,8	5,2	5,4	5,0	6,9	5,6	5,2	5,3
	S_r^*	5,4	5,5	4,8	4,9	6,1	4,9	5,0	4,9
	$S_{r_1}^*$	5,4	5,5	4,9	4,9	6,2	4,9	5,0	4,9
	$S_{r_2}^*$	5,4	5,5	4,9	4,9	6,2	4,9	5,0	4,9
	S_g	14,6	10,5	7,8	7,1	24,5	15,8	11,4	9,3
	S_g^*	6,2	6,0	5,0	5,1	6,0	6,4	5,4	5,1
$\alpha = 1\%$	LR	6,2	3,4	2,3	2,1	13,9	7,0	3,9	2,7
	LR^*	2,9	1,9	1,8	1,2	5,9	3,1	2,2	1,5
	S_r	0,9	1,4	1,2	1,0	1,9	1,0	1,2	1,1
	S_r^*	1,2	1,4	0,5	0,8	1,7	0,7	1,0	0,8
	$S_{r_1}^*$	1,2	1,4	0,8	1,5	1,7	0,7	1,0	0,8
	$S_{r_2}^*$	1,2	1,4	0,8	1,5	1,7	0,7	1,0	0,8
	S_g	5,1	3,2	2,1	1,9	11,8	5,6	3,3	2,5
	S_g^*	1,3	1,2	1,2	1,2	1,5	1,5	1,3	1,0

tamanho amostral ($n = 35$) e o número de parâmetros de perturbação ($p = 3$), variando os valores de k em $k = 3, 4, 5, 6$. Na Tabela 3.4 apresentamos os resultados referentes a esse estudo. Pode-se observar que, assim como para os parâmetros de perturbação, conforme aumenta-se o valor de k , o teste LR se torna mais distorcido, isto é, suas taxas de rejeição se distanciam cada vez mais dos níveis nominais adotados. Ainda, observa-se que o comportamento liberal do teste LR é atenuado por sua versão corrigida LR^* , que aproxima as taxas de rejeição aos níveis nominais adotados, embora ainda apresente

Tabela 3.3: Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0.3$, considerando $n = 35$, $k = 3$ e diversos valores para p .

α	<i>Teste</i>	Modelo t_5				Modelo exponencial potência			
		p				p			
		3	4	5	6	3	4	5	6
10%	<i>LR</i>	16,9	20,8	23,1	29,1	15,8	19,0	22,9	26,9
	<i>LR*</i>	11,9	13,3	13,5	16,9	12,4	14,0	16,2	17,9
	<i>S_r</i>	10,8	11,5	11,7	11,1	9,6	10,8	10,5	10,9
	<i>S_r*</i>	10,3	10,6	10,4	10,0	9,7	10,7	10,1	10,5
	<i>S_{r1}*</i>	10,3	10,6	10,4	10,0	9,7	10,7	10,1	10,5
	<i>S_{r2}*</i>	10,3	10,6	10,4	10,0	9,7	10,7	10,1	10,5
	<i>S_g</i>	16,3	20,1	22,1	28,3	15,4	18,7	21,3	26,4
	<i>S_g*</i>	11,4	11,6	11,9	15,3	10,6	11,2	11,6	14,4
5%	<i>LR</i>	10,1	13,0	15,0	20,2	8,8	11,7	14,5	18,0
	<i>LR*</i>	6,2	7,2	7,3	10,0	6,6	7,9	9,0	11,0
	<i>S_r</i>	5,3	5,6	5,7	5,7	5,0	5,6	5,4	5,8
	<i>S_r*</i>	5,3	5,2	5,0	5,3	4,8	5,2	5,0	5,9
	<i>S_{r1}*</i>	5,3	5,2	5,0	5,3	4,8	5,2	5,0	5,9
	<i>S_{r2}*</i>	5,3	5,2	5,0	5,3	4,8	5,2	5,0	5,9
	<i>S_g</i>	9,5	12,5	14,0	19,6	8,6	11,3	13,3	17,6
	<i>S_g*</i>	6,1	6,2	6,4	9,0	5,8	5,9	6,0	8,1
1%	<i>LR</i>	3,0	4,2	5,5	8,7	2,9	3,8	5,1	7,4
	<i>LR*</i>	1,2	1,7	2,0	3,2	1,6	2,0	2,5	3,2
	<i>S_r</i>	0,9	1,2	1,2	1,3	1,1	1,1	1,3	1,4
	<i>S_r*</i>	1,1	1,1	1,1	1,4	0,9	0,7	1,1	1,8
	<i>S_{r1}*</i>	1,1	1,1	1,1	1,4	0,9	0,8	1,1	1,9
	<i>S_{r2}*</i>	1,1	1,1	1,1	1,4	0,9	0,8	1,1	1,8
	<i>S_g</i>	2,6	4,0	4,6	8,2	2,6	3,7	4,1	6,7
	<i>S_g*</i>	1,4	1,4	1,4	2,9	1,4	1,5	1,3	2,0

caráter ligeiramente liberal. A exemplo, para $\alpha = 5\%$ e $k = 3$, o modelo t_5 apresenta as taxas de rejeição para os testes *LR* e *LR** iguais a 10,1% e 6,2%, respectivamente, e o modelo exponencial potência 8,8% e 6,6%, respectivamente. Aumentando o valor de k para $k = 5$, temos para o modelo t_5 , 12,1% (*LR*) e 5,9% (*LR**), e para o modelo exponencial potência, 11,3% (*LR*) e 6,5% (*LR**). O teste escore apresenta taxas bastante próximas dos níveis nominais considerados, sendo, de forma geral, ligeiramente liberal. A sua versão corrigida tente a atenuar tal comportamento, trazendo as taxas para valores

Tabela 3.4: Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0.3$, considerando $n = 35$, $p = 3$ e diversos valores para k .

α	<i>Teste</i>	Modelo t_5				Modelo exponencial potência			
		k				k			
		3	4	5	6	3	4	5	6
10%	<i>LR</i>	16,9	19,4	20,5	22,7	15,8	17,3	19,1	21,5
	<i>LR*</i>	11,9	11,8	10,9	11,9	12,4	12,1	12,2	13,9
	S_r	10,8	10,6	10,2	10,9	9,6	9,7	9,5	10,5
	S_r^*	10,3	10,1	9,7	10,2	9,7	10,0	9,3	9,8
	$S_{r_1}^*$	10,3	10,1	9,7	10,2	9,7	10,0	9,3	9,8
	$S_{r_2}^*$	10,3	10,1	9,7	10,2	9,7	10,0	9,3	9,8
	S_g	16,3	18,6	18,7	21,0	15,4	16,6	17,8	20,1
	S_g^*	11,4	11,9	10,7	13,2	10,6	11,2	10,9	12,9
5%	<i>LR</i>	10,1	11,7	12,1	14,4	8,8	10,0	11,3	13,5
	<i>LR*</i>	6,2	6,4	5,9	6,3	6,6	6,3	6,5	7,7
	S_r	5,3	5,3	5,3	5,2	5,0	5,2	5,2	5,6
	S_r^*	5,3	5,3	5,2	4,9	4,8	5,3	4,6	4,2
	$S_{r_1}^*$	5,3	5,3	5,2	4,9	4,8	5,3	4,7	4,3
	$S_{r_2}^*$	5,3	5,3	5,2	4,9	4,8	5,3	4,7	4,3
	S_g	9,5	10,8	10,7	12,9	8,6	9,7	10,4	12,2
	S_g^*	6,1	6,2	5,9	7,5	5,8	5,8	5,6	7,1
1%	<i>LR</i>	3,0	3,6	3,8	5,0	2,9	3,1	3,4	4,2
	<i>LR*</i>	1,2	1,4	1,2	1,3	1,6	1,6	1,6	1,6
	S_r	0,9	1,1	0,9	1,1	1,1	1,3	1,2	1,5
	S_r^*	1,0	1,2	1,0	1,1	0,9	1,2	0,6	0,3
	$S_{r_1}^*$	1,0	1,2	1,0	1,1	0,9	1,3	0,7	0,6
	$S_{r_2}^*$	1,0	1,2	1,0	1,1	0,9	1,3	0,7	0,5
	S_g	2,6	3,0	3,0	4,0	2,6	2,8	2,9	3,5
	S_g^*	1,4	1,5	1,3	2,0	1,4	1,4	1,2	1,6

ainda mais próximos dos níveis nominais adotados. Ainda, as versões modificadas dos testes *escore*, $S_{r_1}^*$ e $S_{r_2}^*$, apresentam comportamento semelhante ao do teste S_r^* . Por exemplo, para $\alpha = 5\%$ e $k = 5$, temos para o modelo t_5 as taxas de rejeição dos testes iguais a 5,3% (S_r) e 5% (S_r^* , $S_{r_1}^*$, $S_{r_2}^*$) e para o modelo exponencial potência, 5,2% (S_r), 4,6% (S_r^*) e 4,7% ($S_{r_1}^*$, $S_{r_2}^*$). O teste gradiente tem comportamento semelhante ao do teste *LR*, no mais, apresenta taxas de rejeição menores que as do teste *LR*. O teste gradiente corrigido, como era esperado, aprimora o teste S_g , trazendo as taxas de rejeição para valores próxi-

mos aos níveis nominais adotados. A saber, para $\alpha = 1\%$ e $k = 4$, temos para o modelo t_5 as taxas de rejeição dos testes S_g e S_g^* iguais a 3,0% e 1,5%, respectivamente, e para o modelo exponencial potência, 2,8% e 1,4%, respectivamente.

Nas Figuras 3.1-3.4 construímos o gráfico de discrepância relativa de quantis para os modelos t_5 (Figuras 3.1-3.2) e exponencial potência (Figuras 3.3-3.4) considerando o cenário em que $n = 30$, $p = 3, 5$ e $k = 3$. A discrepância relativa é definida como

$$\frac{ST(1 - \alpha) - \chi_{k-1}^2(1 - \alpha)}{\chi_{k-1}^2(1 - \alpha)},$$

onde $ST(1 - \alpha)$ denota o quantil amostral de ordem $(1 - \alpha)$ do conjunto de valores simulados da estatística de teste ST ($LR, LR^*, S_r, S_r^*, S_{r_1}^*, S_{r_2}^*, S_g$ ou S_g^* , conforme o caso) e χ_{k-1}^2 denota o correspondente quantil da distribuição χ_{k-1}^2 . Sendo assim, quanto mais próxima da ordenada nula a curva de discrepância estiver, melhor aproximada está a distribuição nula da estatística de teste pela distribuição assintótica χ^2 de referência.

Analisando as Figuras 3.1-3.4, podemos observar que as estatísticas LR, LR^* e S_g não têm distribuição nula bem aproximada pela distribuição χ^2 de referência, uma vez que suas curvas de discrepância estão bastante acima da ordenada nula em todos os cenários, o que ratifica a tendência liberal dos testes baseados nestas estatísticas já observada na simulação para o tamanho dos testes. Ainda, podemos observar que, de forma geral, as estatísticas $S_r, S_r^*, S_{r_1}^*, S_{r_2}^*$ e S_g^* têm curvas de discrepância bastante próximas à ordenada nula e para ambos os modelos considerando o cenário em que $p = 3$, Figuras 3.1 e 3.3, as estatísticas $S_r^*, S_{r_1}^*, S_{r_2}^*$ e S_g^* apresentam curvas melhor aproximadas da ordenada nula e praticamente sobrepostas, mostrando que a aproximação da distribuição nula de tais estatísticas pela distribuição qui-quadrado de referência é bastante similar. Considerando o cenário em que $p = 5$, Figuras 3.2 e 3.4, podemos observar que as estatísticas $S_r^*, S_{r_1}^*$ e $S_{r_2}^*$ têm a curva de discrepância melhor aproximada da ordenada nula, com destaque para o modelo t_5 , Figura 3.2. Ainda para este cenário, a estatística S_g^* apresenta curva de discrepância próxima, porém ligeiramente acima, da ordenada nula, confirmando a tendência ligeiramente liberal do teste baseado em tal estatística.

Para o estudo do poder, consideramos apenas as estatísticas que apresentaram taxas de rejeição próximas aos níveis nominais considerados e distribuição nula com boa aproximação pela distribuição χ^2 de referência, isto é, as estatísticas $S_r, S_r^*, S_{r_1}^*, S_{r_2}^*$ e S_g^* . Os resultados foram obtidos a partir da hipótese alternativa $H_1 : \delta_2 = \delta_3 = \delta \neq 0$, consideramos para os modelos t_5 e exponencial potência os cenários em que $n = 30$, $p = 3, 5$, $k = 3$, $\alpha = 10\%$ e diferentes valores para δ , com δ variando entre 0,1 e 5,0 e são apresentados

Figura 3.1: Discrepância relativa de quantis para o modelo t_5 com $n = 30$, $p = 3$, $k = 3$

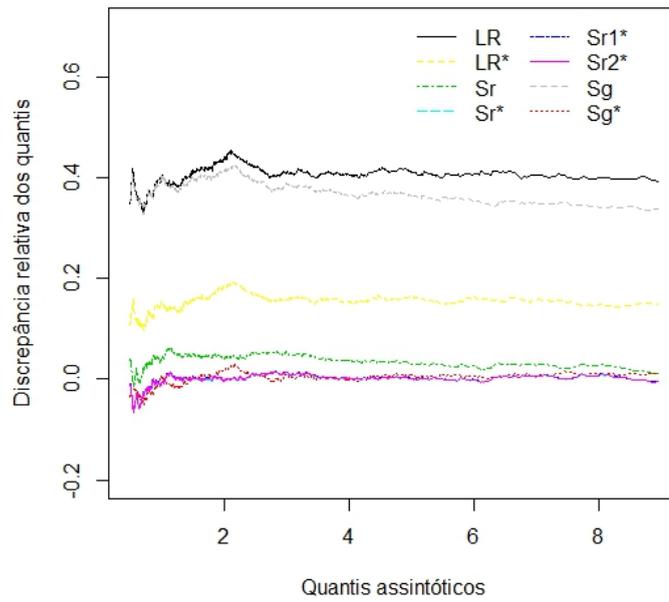


Figura 3.2: Discrepância relativa de quantis para o modelo t_5 com $n = 30$, $p = 5$, $k = 3$

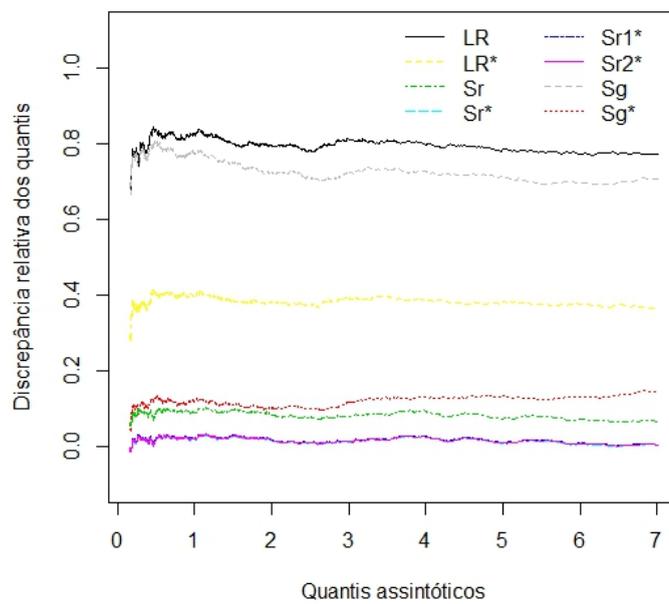


Figura 3.3: Discrepância relativa de quantis para o modelo exponencial potência com $\kappa = 0.3$, considerando $n = 30$, $p = 3$, $k = 3$

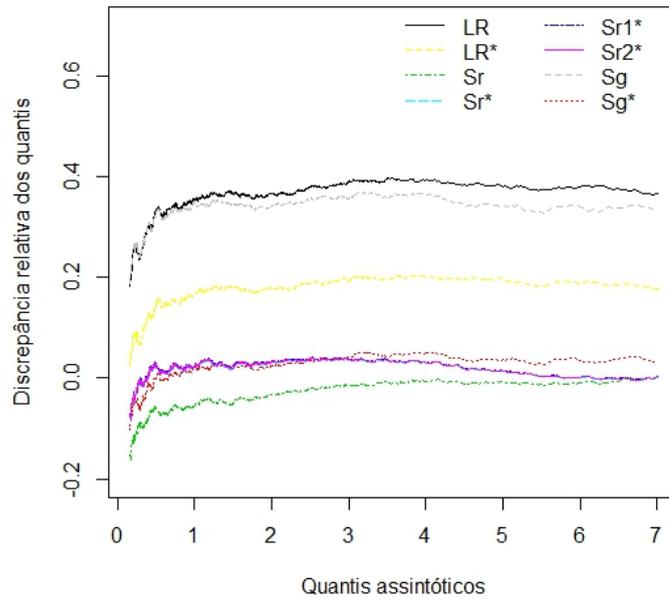
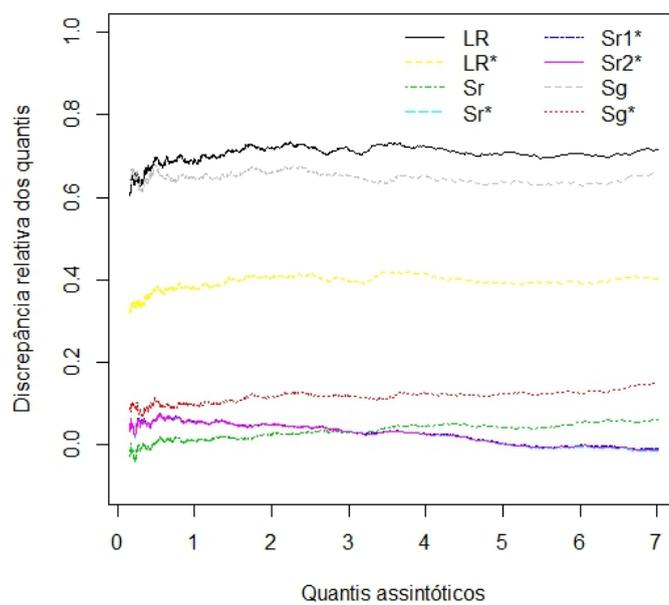


Figura 3.4: Discrepância relativa de quantis para o modelo exponencial potência com $\kappa = 0.3$, considerando $n = 30$, $p = 5$, $k = 3$



na Tabela 3.5. Notamos que, de forma geral, os testes baseados nas estatísticas S_r e S_g^* apresentaram poderes semelhantes e ligeiramente superiores aos dos demais testes. Para o modelo exponencial potência considerando o caso em que $p = 5$, pode-se observar ligeira vantagem do poder do teste baseado em S_g^* perante o teste baseado em S_r para todos os valores assumidos por δ .

Tabela 3.5: Poder dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0.3$, considerando $n = 30$, $p = 3, 5$, $k = 3$ e $\alpha = 10\%$.

δ	Modelo t_5					Modelo exponencial potência				
	<i>Teste</i>					<i>Teste</i>				
	S_r	S_r^*	$S_{r_1}^*$	$S_{r_2}^*$	S_g^*	S_r	S_r^*	$S_{r_1}^*$	$S_{r_2}^*$	S_g^*
$p = 3$										
0,1	10,8	10,3	10,3	10,3	10,1	10,1	10,6	10,6	10,6	10,8
0,5	14,3	13,6	13,7	13,6	14,8	16,4	17,0	17,0	17,0	15,3
1,0	22,6	21,4	21,4	21,4	21,2	29,1	30,3	30,4	30,4	30,1
1,5	38,7	37,4	37,5	37,4	37,9	49,7	50,1	50,1	50,5	46,6
2,0	78,8	77,8	77,8	77,8	79,1	77,2	77,5	77,5	77,5	80,3
2,5	79,3	78,2	78,2	78,3	79,8	81,4	79,7	81,1	81,1	83,4
3,0	87,8	87,3	87,3	87,3	86,8	85,9	86,1	86,1	86,1	88,5
3,5	97,5	97,3	97,3	97,3	98,8	96,2	96,1	96,1	96,1	98,0
4,0	98,0	97,9	97,9	97,9	98,6	97,4	96,2	97,4	97,4	98,2
4,5	98,7	98,6	98,6	98,6	99,7	98,0	98,0	98,0	98,0	99,0
$p = 5$										
0,1	12,5	10,8	10,8	10,8	13,4	11,6	10,5	10,6	10,6	13,5
0,5	12,7	11,4	11,4	11,4	15,6	15,3	14,9	14,9	14,9	16,9
1,0	26,9	24,4	24,4	24,4	27,3	23,7	23,3	23,4	23,3	24,5
1,5	36,5	34,5	34,5	34,5	39,8	41,4	40,3	40,4	40,4	45,5
2,0	69,6	67,9	67,9	67,9	67,1	60,2	61,6	69,0	68,9	71,8
2,5	74,6	72,5	72,5	72,5	73,8	63,2	62,6	62,6	62,6	71,9
3,0	82,2	80,2	80,2	80,2	81,7	88,3	88,2	88,4	88,3	92,5
3,5	91,2	90,6	90,6	90,6	90,5	91,6	91,0	91,5	91,5	94,8
4,0	96,5	95,9	95,9	95,9	98,6	93,4	93,4	93,4	93,4	95,1
4,5	96,8	96,4	96,4	96,4	98,7	96,3	96,4	96,4	96,4	98,0
5,0	99,2	99,1	99,1	99,1	99,9	99,1	99,1	99,1	99,1	99,9

3.5 Aplicação

No que segue, aplicamos a metodologia dos testes apresentados nas seções anteriores a um conjunto de dados reais (ver Apêndice G). Os dados analisados referem-se ao peso das lentes dos olhos de coelhos europeus na Austrália (*Oryctolagus Cuniculus*), y , em mg, e à idade do animal, x , em dias, numa amostra contendo 71 observações. Estes dados foram analisados por Wei (1998, Exemplo 6.8), que verificaram a suspeita de dois pontos aberrantes sob estimação de mínimos quadrados, assim, indicando que tal conjunto de dados suporta erros com caudas mais pesadas que as da normal. Dessa maneira, os dados foram reanalisados por Cysneiros et al. (2005), que adotaram erros seguindo distribuições simétricas, dentre elas, a distribuição t -Student com 4 graus de liberdade. A escolha dos graus de liberdade foi baseada no modelo que apresentou o menor AIC. O gráfico de resíduos contra valores ajustados (veja Cysneiros et al., 2005, Seção 3.2) mostraram que as observações 4, 15, 16 e 17 apresentaram, em valor absoluto, resíduos grandes, dando alguma evidência de heteroscedasticidade. Dessa maneira, consideramos um modelo mais geral do que o ajustado por Cysneiros et al. (2005).

O modelo heteroscedástico utilizado é dado por

$$y_\ell = \exp\left(\beta_0 - \frac{\beta_1}{x_\ell - \beta_2}\right) e^{\epsilon_\ell},$$

com $\epsilon_\ell \sim S(0, \exp\{\beta_3 + \beta_4 x_\ell\})$ e $\ell = 1, \dots, 71$. Nosso interesse principal agora é testar $H_0 : \beta_4 = 0$ contra $H_1 : \beta_4 \neq 0$. Para este teste, temos que os valores das estatísticas de teste e seu respectivos níveis descritivos, entre parênteses, são dados, respectivamente, por $LR = 8,368 (0,004)$, $LR^* = 8,348 (0,004)$, $S_r = 6.766 (0,009)$, $S_r^* = 6.678 (0,010)$, $S_{r_1}^* = 6.679 (0,010)$, $S_{r_2}^* = 6.678 (0,010)$, $S_g = 7,828 (0,005)$ e $S_g^* = 7.430 (0,006)$. Dessa maneira, ao nível nominal de 1%, temos que os testes baseados nas estatísticas score corrigida (S_r^*) e nas versões modificadas ($S_{r_1}^*$) e ($S_{r_2}^*$) conduzem a não rejeição da hipótese nula, enquanto que os demais testes conduzem à rejeição da hipótese nula.

3.6 Conclusões

Neste capítulo, apresentamos os fatores de correção tipo-Bartlett para o aperfeiçoamento dos testes escore e gradiente na classe dos modelos não-lineares simétricos heteroscedásticos, sendo o ajuste para a estatística escore nesta classe de modelos inicialmente obtido por Nascimento (2010).

A fim de avaliar o desempenho dos testes corrigidos e suas respectivas versões não corrigidas na classe dos MNLSH, realizamos um estudo de simulação de Monte Carlo sob diversos cenários. Além dos testes escore, gradiente e suas versões corrigidas, consideramos para efeito de comparação os testes da razão de verossimilhanças usual e o corrigido via correção de Bartlett, obtido para a classe de modelos em estudo por Brito (2009).

Os resultados numéricos mostraram que os testes da razão de verossimilhanças e gradiente têm comportamento similar, apresentando tendência liberal, a qual é acentuada conforme o número de parâmetros do modelo aumenta (seja parâmetros de perturbação ou de interesse). As versões corrigidas destes testes, assim como esperado, atenuam a tendência liberal de suas respectivas versões não corrigidas, aproximando as taxas de rejeição dos testes aos níveis nominais considerados, sendo o teste gradiente corrigido mais eficiente dentre os dois, isto é, apresentando taxas mais próximas dos níveis de significância considerados. O número de parâmetros do modelo tem leve influência sobre o desempenho dos testes da razão de verossimilhanças e gradiente corrigidos se comparada com a influência exercida sobre suas versões não corrigidas, isto é, quando aumentamos o número de parâmetros do modelo, a distorção destes testes é menor do que a das suas respectivas versões não corrigidas. Dentre os testes não corrigidos, o teste escore apresenta o melhor desempenho, tendo, de modo geral, caráter ligeiramente liberal, sendo tal tendência atenuada por sua versão corrigida. Além disso, os testes escore modificados propostos por Cordeiro et al. (1998) e Kakisawa (1996) apresentaram desempenho idêntico ao teste escore corrigido, salvo alguns cenários. Ainda, os testes escore usual, corrigido e as versões modificadas não sofreram influência do número de parâmetros, isto é, independente do número de parâmetros do modelo, as taxas de rejeição dos testes se apresentaram estáveis e próximas aos níveis nominais considerados. Com relação ao poder dos testes, pode-se observar que, de forma geral, os testes escore usual e gradiente corrigido apresentaram poderes semelhantes e ligeiramente superiores ao dos demais testes. Diante de todos os resultados obtidos, recomendamos para aplicações práticas o uso dos testes escore modificados e corrigido.

Poder local dos testes da razão de verossimilhanças, Wald, escore e gradiente

4.1 Introdução

Em virtude da dificuldade de se obter, quando existente, a distribuição nula exata das estatísticas de teste, é comum na literatura realizar inferências a partir de testes de hipóteses baseados em grandes amostras. Os testes assintóticos mais utilizados são os testes da razão de verossimilhanças, escore e Wald. Satisfazendo as condições de regularidade, as estatísticas desses três testes são assintoticamente equivalentes até primeira ordem, tanto sob hipótese nula como sob alternativas de Pitman (sequência de hipóteses alternativas contíguas convergindo para a hipótese nula sob taxa de convergência de $n^{-1/2}$, sendo n o tamanho da amostra).

Um novo teste assintótico, o teste gradiente, foi proposto na literatura por [Terrell \(2002\)](#). A estatística gradiente, além de apresentar estrutura bastante simples, não dependendo da matriz de informação observada, nem da esperada, compartilha das mesmas propriedades assintóticas de primeira ordem das estatísticas da razão de verossimilhanças, escore e Wald. Recentemente [Lemonte e Ferrari \(2012\)](#) obtiveram uma expansão assintótica para a distribuição da estatística gradiente sob uma sequência de hipóteses alternativas contíguas convergindo para a hipótese nula sob taxa de convergência $n^{-1/2}$,

mostrando que a estatística gradiente também é equivalente às estatísticas de teste clássicas sob hipóteses alternativas de Pitman.

Diante dos quatro testes assintóticos equivalentes existentes na literatura, um possível questionamento seria indicar qual teste mais adequado para determinado estudo. Até primeira ordem, tais testes têm poder igual sob alternativas de Pitman e até ordem $n^{-1/2}$, o mesmo tamanho. Dessa maneira, um critério de escolha dentre eles pode ser determinado através da comparação dos seus poderes até ordem $n^{-1/2}$. [Hayakawa \(1975\)](#) desenvolveu expansões assintóticas sob uma sequência de hipóteses alternativas de Pitman para as distribuições das estatísticas da razão de verossimilhanças e Wald, [Harris e Peers \(1980\)](#) obtiveram resultado análogo para a estatística score e recentemente [Lemonte e Ferrari \(2012\)](#) para a estatística gradiente. A partir dessas expansões assintóticas é feita a comparação do poder dos quatro testes até ordem $n^{-1/2}$.

Trabalhos na literatura têm explorado o estudo do poder local. [Lemonte \(2011\)](#) comparou o poder local de testes em modelos não-lineares da família exponencial, [Lemonte e Ferrari \(2012b\)](#) em modelos de dispersão, [Lemonte \(2012\)](#) em modelos lineares simétricos e [Lemonte \(2014\)](#) em modelos não-lineares de efeitos mistos. Nesta direção, nosso objetivo é realizar um estudo comparando o poder local dos testes da razão de verossimilhanças, Wald, score e gradiente para a classe dos modelos não-lineares simétricos heteroscedásticos, estendendo os resultados obtidos em [Lemonte \(2012\)](#), para isto, iremos derivar as expansões assintóticas não nulas até ordem $n^{-1/2}$ das funções de distribuição das respectivas estatísticas de teste sob uma sequência de alternativas de Pitman. No que segue, discorreremos acerca do poder local na Seção 4.2, derivando expansões assintóticas sob hipóteses alternativas de Pitman para a distribuição das estatísticas de teste em estudo. Na Seção 4.3, realizamos a comparação analítica do poder dos quatro testes. A fim de comparar o desempenho dos testes considerando amostras de tamanho pequeno e moderado, na Seção 4.4 realizamos um estudo de simulação de Monte Carlo avaliando o desempenho dos testes quanto a tamanho e poder para diversos cenários. Apresentamos conclusões acerca dos resultados obtidos na Seção 4.5.

4.2 Poder local

O modelo em estudo é definido em (3.1) por

$$Y_\ell = \mu_\ell + \sqrt{\phi_\ell^*} e_\ell; \quad e_\ell \sim S(0, 1, g), \quad \ell = 1, \dots, n,$$

com $\mu_\ell = f(x_\ell; \boldsymbol{\beta})$ e $\phi_\ell^* = m(\tau_\ell^*) > 0$ tais como definidos em (2.2) e (3.2), respectivamente. Seja $\boldsymbol{\theta} = (\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top)^\top$ o vetor de parâmetros de dimensão $(k+p) \times 1$, com $\boldsymbol{\delta} = (\delta_0, \boldsymbol{\delta}_1^\top)^\top$, sendo δ_0 escalar e $\boldsymbol{\delta}_1 = (\delta_1, \dots, \delta_{k-1})^\top$, e $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^\top$. Os vetores $\boldsymbol{\delta}$ e $\boldsymbol{\beta}$ são globalmente ortogonais no sentido de Cox e Reid (1987), isto é, a matriz de informação de Fisher para $\boldsymbol{\theta}$ tem estrutura bloco-diagonal, sendo definida na Seção 3.2.

Estamos interessados em testar $H_0 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_1^{(0)}$ contra $H_1 : \boldsymbol{\delta}_1 \neq \boldsymbol{\delta}_1^{(0)}$, sendo $\boldsymbol{\delta}_1^{(0)}$ um vetor $(k-1)$ -dimensional de constantes especificadas de tal modo que sob H_0 , $\phi_\ell^* = m(\delta_0)$, para $\ell = 1, \dots, n$, indicando que o modelo é homoscedástico.

As estatísticas da razão de verossimilhanças (LR), Wald (W), escore (S_r) e gradiente (S_g) para o teste de H_0 na classe dos MNLSH são expressas por

$$\begin{aligned} LR &= 2\{l(\hat{\boldsymbol{\delta}}_1, \hat{\delta}_0, \hat{\boldsymbol{\beta}}) - l(\boldsymbol{\delta}_1^{(0)}, \tilde{\delta}_0, \tilde{\boldsymbol{\beta}})\}, \\ W &= [\hat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}_1^{(0)}]^\top (\hat{\mathbf{R}}^\top \hat{\mathbf{V}} \hat{\mathbf{R}}) [\hat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}_1^{(0)}], \\ S_r &= \frac{1}{4} [\mathbf{W}_1 \tilde{\boldsymbol{\Lambda}} (\tilde{\mathbf{S}} \tilde{\mathbf{F}}_1 \tilde{\mathbf{u}} - \tilde{\mathbf{F}}_1 \boldsymbol{\iota})]^\top (\tilde{\mathbf{R}}^\top \tilde{\mathbf{V}} \tilde{\mathbf{R}})^{-1} [\mathbf{W}_1 \tilde{\boldsymbol{\Lambda}} (\tilde{\mathbf{S}} \tilde{\mathbf{F}}_1 \tilde{\mathbf{u}} - \tilde{\mathbf{F}}_1 \boldsymbol{\iota})] \text{ e} \\ S_g &= \frac{1}{2} [\mathbf{W}_1 \tilde{\boldsymbol{\Lambda}} (\tilde{\mathbf{S}} \tilde{\mathbf{F}}_1 \tilde{\mathbf{u}} - \tilde{\mathbf{F}}_1 \boldsymbol{\iota})]^\top (\hat{\boldsymbol{\delta}}_1 - \boldsymbol{\delta}_1^{(0)}), \end{aligned}$$

com $\mathbf{W}_1, \boldsymbol{\Lambda}, \mathbf{S}, \mathbf{F}_1, \mathbf{R}, \mathbf{V}, \mathbf{u}$ e $\boldsymbol{\iota}$ definidos no Capítulo 3. Os sinais ‘ $\hat{\cdot}$ ’ e ‘ $\tilde{\cdot}$ ’ denotam que as quantidades foram avaliadas, respectivamente, sob os estimadores de máxima verossimilhança irrestrito e restrito dos parâmetros. Para grandes amostras e sob H_0 , a distribuição das quatro estatísticas de teste LR, W, S_r e S_g é aproximadamente χ_{k-1}^2 e para uma sequência de hipóteses alternativas de Pitman convergindo para H_0 sob taxa de $n^{-1/2}$ é aproximadamente $\chi_{(k-1), \lambda}^2$, ou seja, qui-quadrado não central com $k-1$ graus de liberdade e um parâmetro de não centralidade λ . A sequência de hipóteses alternativas de Pitman é definida como: $H_1 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_1^{(0)} + \boldsymbol{\epsilon}$, com $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_{k-1})^\top$, sendo $\epsilon_r = O(n^{-1/2})$, para $r = 1, \dots, k-1$.

Seja S_i , $i = 1, 2, 3, 4$, denotando as estatísticas da razão de verossimilhanças, Wald, escore e gradiente, respectivamente. Assumindo a hipótese alternativa local $H_1 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_1^{(0)} + \boldsymbol{\epsilon}$, sendo $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_{k-1})^\top$ com $\epsilon_r = O(n^{-1/2})$ para $r = 1, \dots, k-1$, iremos derivar as expansões assintóticas para a distribuição das estatísticas S_1, S_2, S_3 e S_4 , até ordem $n^{-1/2}$, na classe dos MNLSH. Para isto, inicialmente consideremos algumas notações. Seja

$$\boldsymbol{\epsilon}^* = (\boldsymbol{\epsilon}_\delta^{*\top}, \mathbf{0}^\top), \text{ com } \boldsymbol{\epsilon}_\delta^* = \begin{bmatrix} K_{\delta_0 \delta_0}^{-1} \mathbf{K}^{-1} \boldsymbol{\delta}_0 \boldsymbol{\delta}_1 \\ -\mathbf{I}_{k-1} \end{bmatrix} \boldsymbol{\epsilon},$$

sendo \mathbf{I}_{k-1} uma matriz identidade de ordem $(k-1)$ e $\mathbf{0}$ um vetor $p \times 1$ de zeros. Ainda, os índices a, b, c, \dots variam em $\boldsymbol{\delta}$, os índices s, t, v, \dots variam em $\boldsymbol{\beta}$, $U_a = \partial l(\boldsymbol{\theta}; \mathbf{y}) / \partial \delta_a$, $U_{ab} =$

$\partial^2 l(\boldsymbol{\theta}; \mathbf{y}) / \partial \delta_a \partial \delta_b$, $U_{abc} = \partial^3 l(\boldsymbol{\theta}; \mathbf{y}) / \partial \delta_a \partial \delta_b \partial \delta_c$, $U_{sta} = \partial^3 l(\boldsymbol{\theta}; \mathbf{y}) / \partial \beta_s \partial \beta_t \partial \delta_a$, $\kappa_{ab} = E(U_{ab})$, $\kappa_{abc} = E(U_{abc})$, $\kappa_{a,bc} = E(U_a U_{bc})$, $\kappa_{a,b,c} = E(U_a U_b U_c)$ e $\kappa_{sta} = E(U_{sta})$.

A distribuição não nula das estatísticas S_i , com $i = 1, 2, 3, 4$, sob alternativas de Pitman para o teste de $H_0 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_1^{(0)}$ na classe dos modelos não-lineares simétricos heteroscedásticos pode ser expressa por

$$\mathbb{P}(S_i \leq x_\alpha) = G_{k-1,\lambda}(x_\alpha) + \sum_{k=0}^3 b_{ik} G_{k-1+2l,\lambda}(x_\alpha) + O(n^{-1}), \quad (4.1)$$

com o valor x_α obtido de $\mathbb{P}(\chi_{q-1}^2 \leq \alpha) = 1 - \alpha$ e $G_{k-1,\lambda}(x_\alpha)$ sendo a função de distribuição acumulada de uma variável aleatória qui-quadrado não central com $k - 1$ graus de liberdade e parâmetro de não centralidade λ , sendo $\lambda = \boldsymbol{\epsilon}^\top (\mathbf{K}_{\boldsymbol{\delta}_1 \boldsymbol{\delta}_1} - \mathbf{K}_{\boldsymbol{\delta}_1 \boldsymbol{\delta}_0} \mathbf{K}_{\boldsymbol{\delta}_0 \boldsymbol{\delta}_0}^{-1} \mathbf{K}_{\boldsymbol{\delta}_0 \boldsymbol{\delta}_1}) \boldsymbol{\epsilon}$ e os coeficientes b_{il} 's, com $i = 1, 2, 3, 4$ e $l = 0, 1, 2, 3$, dados por

$$\begin{aligned} b_{11} &= b_{11}^0 + \xi, \quad b_{21} = b_{21}^0 + \xi, \quad b_{31} = b_{31}^0 + \xi, \quad b_{41} = b_{41}^0 + \xi, \\ b_{12} &= b_{33} = -\frac{1}{6} \sum_{a,b,c=1}^k \kappa_{a,b,c} \epsilon_a^* \epsilon_b^* \epsilon_c^*, \\ b_{22} &= \frac{1}{2} \sum_{a,b,c=1}^k \kappa_{a,bc} \epsilon_a^* \epsilon_b^* \epsilon_c^* + \frac{1}{2} \sum_{a,b,c=1}^k \kappa_{abc} m_{ab} \epsilon_c^*, \\ b_{13} &= 0, \quad b_{23} = -2b_{43} = \frac{1}{6} \sum_{a,b,c=1}^k \kappa_{abc} \epsilon_a^* \epsilon_b^* \epsilon_c^*, \quad b_{32} = -\frac{1}{2} \sum_{a,b,c=1}^k \kappa_{a,b,c} m_{ab} \epsilon_c^* \text{ e} \\ b_{42} &= -\frac{1}{4} \sum_{a,b,c=1}^k \kappa_{abc} m_{ab} \epsilon_c^* + \frac{1}{4} \sum_{a,b,c=1}^k (\kappa_{abc} + 2\kappa_{a,bc}) \epsilon_a^* \epsilon_b^* \epsilon_c^*, \end{aligned}$$

sendo

$$\begin{aligned} b_{11}^0 &= -\frac{1}{6} \sum_{a,b,c=1}^k (\kappa_{abc} - 2\kappa_{a,b,c}) \epsilon_a^* \epsilon_b^* \epsilon_c^* - \frac{1}{2} \sum_{a,b,c=1}^k (\kappa_{abc} + 2\kappa_{a,bc}) a_{ab} \epsilon_c^* \\ &\quad - \frac{1}{2} \sum_{a=2}^k \sum_{a,b=1}^k (\kappa_{abc} + \kappa_{a,bc}) \epsilon_a \epsilon_b^* \epsilon_c^*, \\ b_{21}^0 &= -\frac{1}{2} \sum_{a,b,c=1}^k (\kappa_{abc} + 2\kappa_{a,bc}) \epsilon_a^* \epsilon_b^* \epsilon_c^* + \sum_{a,b,c=1}^k \kappa_{a,bc} m_{ab} \epsilon_c^* \\ &\quad - \frac{1}{2} \sum_{a,b,c=1}^k (\kappa_{abc} + 2\kappa_{a,bc}) \kappa^{a,b} \epsilon_c^* - \frac{1}{2} \sum_{a=2}^k \sum_{b,c=1}^k (\kappa_{abc} + \kappa_{a,bc}) \epsilon_a \epsilon_b^* \epsilon_c^*, \end{aligned}$$

$$\begin{aligned}
b_{31}^0 &= -\frac{1}{6} \sum_{a,b,c=1}^k (\kappa_{abc} - 2\kappa_{a,b,c}) \epsilon_a^* \epsilon_b^* \epsilon_c^* + \frac{1}{2} \sum_{a,b,c=1}^k \kappa_{a,b,c} m_{ab} \epsilon_c^* \\
&\quad - \frac{1}{2} \sum_{a,b,c=1}^k (\kappa_{abc} + 2\kappa_{a,b,c}) a_{ab} \epsilon_c^* - \frac{1}{2} \sum_{a=2}^k \sum_{b,c=1}^k (\kappa_{abc} + \kappa_{a,b,c}) \epsilon_a \epsilon_b^* \epsilon_c^*, \\
b_{41}^0 &= \frac{1}{4} \sum_{a,b,c=1}^k \kappa_{abc} \kappa^{a,b} \epsilon_c^* - \frac{1}{2} \sum_{a,b,c=1}^k (\kappa_{abc} + 2\kappa_{a,b,c}) \epsilon_a^* \epsilon_b^* \epsilon_c^* \\
&\quad - \frac{1}{4} \sum_{a,b,c=1}^k (4k_{a,bc} + 3k_{abc}) a_{ab} \epsilon_c^* - \frac{1}{2} \sum_{a=2}^k \sum_{b,c=1}^k (\kappa_{abc} + \kappa_{a,b,c}) \epsilon_a \epsilon_b^* \epsilon_c^* \text{ e} \\
\xi &= \frac{1}{2} \sum_{a=1}^k \sum_{s,t=1}^p \kappa_{sta} \kappa^{s,t} \epsilon_a^*,
\end{aligned}$$

e $b_{i0} = -(b_{i1} + b_{i2} + b_{i3})$, para $i = 1, 2, 3, 4$. Aqui, ϵ_r^* denota o r -ésimo elemento do vetor $\boldsymbol{\epsilon}^*$, a_{ab} e m_{ab} representam os elementos (a, b) das matrizes \mathbf{A} e \mathbf{M} definidas na Seção 3.3.2, $\mathbf{K}_\delta^{-1} = (\kappa^{r,s})$, com $a, b = 1, \dots, k$, e $\mathbf{K}_\beta^{-1} = (\kappa^{s,t})$, com $s, t = 1, \dots, p$. Além disso, todas as quantidades das expressões b_{il} 's são avaliadas sob H_0 , com exceção do $\boldsymbol{\epsilon}^*$. Ainda, podemos observar que os coeficientes b_{il} 's podem ser escritos em função da contribuição dos $\boldsymbol{\delta}$'s para o poder local (b_{il}^0) e da contribuição dos $\boldsymbol{\beta}$'s para o poder local (ξ). As expressões dos b_{il} 's podem ser utilizadas para qualquer classe de modelos que siga a mesma parametrização dos MNLSH e que os vetores $\boldsymbol{\beta}$ e $\boldsymbol{\delta}$ sejam globalmente ortogonais. Tais expressões foram obtidas inicialmente por [Lemonte \(2014\)](#).

Substituindo os κ 's referentes à classe dos MNLSH calculados no Apêndice D e considerando heteroscedasticidade com efeitos multiplicativos, isto é, $m(\tau_\ell^*) = \exp(\tau_\ell^*)$, temos que o parâmetro de não centralidade λ é dado por $\lambda = \boldsymbol{\epsilon}^\top (\mathbf{W}_1^\top \mathbf{V} \mathbf{R}) \boldsymbol{\epsilon}$ e os coeficientes b_{il} 's e ξ para o teste de $H_0 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_1^{(0)}$ por

$$\begin{aligned}
b_{12} &= -\frac{1}{3} Q_1 \text{tr}\{\mathbf{T}^{(3)}\} \quad b_{22} = -\frac{1}{2} Q_1 \text{tr}\{\mathbf{T}^{(3)}\} + \frac{1}{2} Q_1 \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{T}\}, \\
b_{23} &= \frac{1}{6} Q_1 \text{tr}\{\mathbf{T}^{(3)}\} \quad b_{32} = -Q_1 \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{T}\}, \\
b_{42} &= -\frac{1}{4} Q_1 \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{T}\} - \frac{1}{4} Q_1 \text{tr}\{\mathbf{T}^{(3)}\}, \\
b_{11}^0 &= \frac{1}{2} Q_1 \text{tr}\{\mathbf{T}^{(3)}\} + \frac{1}{2} Q_1 \text{tr}\{\mathbf{Z}_{\delta_{0d}}\}, \\
b_{21}^0 &= \frac{1}{2} Q_1 \text{tr}\{\mathbf{T}^{(3)}\} - Q_1 \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{T}\} + \frac{1}{2} Q_1 \text{tr}\{\mathbf{Z}_{\delta_d} \mathbf{T}\},
\end{aligned}$$

$$\begin{aligned}
b_{31}^0 &= \frac{1}{2}Q_1\text{tr}\{\mathbf{T}^{(3)}\} + Q_1\text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d\mathbf{T}\} + \frac{1}{2}Q_1\text{tr}\{\mathbf{Z}_{\delta_{0d}}\mathbf{T}\}, \\
b_{41}^0 &= \frac{1}{4}Q_1\text{tr}\{\mathbf{Z}_{\delta_d}\mathbf{T}\} + \frac{1}{2}Q_1\text{tr}\{\mathbf{T}^{(3)}\} + \frac{1}{4}Q_1\text{tr}\{\mathbf{Z}_{\delta_{0d}}\mathbf{T}\} \text{ e} \\
\xi &= \frac{1}{2}Q_2\text{tr}\{\Lambda\mathbf{Z}_{\beta_d}\mathbf{T}\},
\end{aligned}$$

sendo $\mathbf{t} = (t_1, \dots, t_n)^\top = \mathbf{W}\boldsymbol{\epsilon}_\delta^*$, $\mathbf{T} = \text{diag}\{t_1, \dots, t_n\}$, $\mathbf{T}^{(2)} = \mathbf{T} \odot \mathbf{T}$ e $\mathbf{T}^{(3)} = \mathbf{T}^{(2)} \odot \mathbf{T}$, sendo \odot a notação para o produto Hadamard, isto é, produto elemento a elemento (Rao, 1973, pág. 30). Detalhes sobre a obtenção dos $b_{i\ell}$'s e ξ são apresentados no Apêndice F.

4.3 Comparação entre as funções de poder

Dado que os testes baseados nas estatísticas da razão de verossimilhanças, Wald, score e gradiente têm o mesmo poder sob uma sequência de hipóteses alternativas contíguas até primeira ordem, comparemos, então, analiticamente seus poderes locais até termos de ordem $n^{-1/2}$, isto é, ignorando termos de ordem menor que $n^{-1/2}$, confrontando os poderes locais dos testes rivais. Definimos a função de poder dos testes baseados nas estatísticas S_i , $i = 1, 2, 3, 4$, para o teste de $H_0 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_1^{(0)}$ e sob uma sequência de hipóteses alternativas contíguas como

$$\pi_i = \pi_i(x_\alpha) = 1 - \mathbb{P}(S_i \leq x_\alpha) = \mathbb{P}(S_i > x_\alpha),$$

com $\mathbb{P}(S_i \leq x_\alpha)$ definido em (4.1). Dessa maneira, temos que

$$\pi_i - \pi_j = \sum_{\ell=0}^3 (b_{j\ell} - b_{i\ell})G_{k-1+2\ell, \lambda}(x_\alpha), \quad j = 1, 2, 3, 4, \quad i \neq j,$$

onde os $b_{i\ell}$'s e $b_{j\ell}$'s são apresentados na seção anterior. Sabemos ainda que

$$G_{m, \lambda}(x) - G_{m+2, \lambda}(x) = 2g_{m+2, \lambda}(x),$$

sendo $g_{m, \lambda}(x)$ a função densidade de probabilidade de uma variável aleatória qui-quadrado não central com m graus de liberdade e parâmetro de não centralidade λ . Após alguma

álgebra, apresentada com detalhes no Apêndice F, temos que

$$\begin{aligned}
\pi_1 - \pi_2 &= K_1 g_{k+1,\lambda}(x) + K_2 g_{k+3,\lambda}(x) + K_3 g_{k+5,\lambda}(x), \\
\pi_1 - \pi_3 &= -\frac{1}{2} K_2 g_{k+3,\lambda}(x) - \frac{1}{2} K_3 g_{k+5,\lambda}(x), \\
\pi_1 - \pi_4 &= K_4 g_{k+1,\lambda}(x) - \frac{1}{2} K_2 g_{k+3,\lambda}(x) - \frac{1}{2} K_3 g_{k+5,\lambda}(x), \\
\pi_2 - \pi_3 &= K_5 g_{k+1,\lambda}(x) - 3K_2 g_{k+3,\lambda}(x) - 3K_3 g_{k+5,\lambda}(x), \\
\pi_2 - \pi_4 &= -K_4 g_{k+1,\lambda}(x) - \frac{3}{2} K_2 g_{k+3,\lambda}(x) - \frac{3}{2} K_3 g_{k+5,\lambda}(x) \text{ e} \\
\pi_3 - \pi_4 &= K_4 g_{k+1,\lambda}(x) + \frac{3}{2} K_2 g_{k+3,\lambda}(x) + K_3 g_{k+5,\lambda}(x),
\end{aligned} \tag{4.2}$$

com

$$\begin{aligned}
K_1 &= Q_1 (\text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{T}\} - \text{tr}\{\mathbf{Z}_{\delta_d} \mathbf{T}\} + \text{tr}\{\mathbf{Z}_{\delta_{0d}} \mathbf{T}\}), \\
K_2 &= -Q_1 \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{T}\}; \quad K_3 = -\frac{1}{3} Q_1 \text{tr}\{\mathbf{T}^{(3)}\}, \\
K_4 &= -\frac{1}{2} Q_1 (\text{tr}\{\mathbf{Z}_{\delta_d} \mathbf{T}\} - \text{tr}\{\mathbf{Z}_{\delta_{0d}} \mathbf{T}\} - \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{T}\}) \text{ e} \\
K_5 &= \text{tr}\{\mathbf{Z}_{\delta_d} \mathbf{T}\} - \text{tr}\{\mathbf{Z}_{\delta_{0d}} \mathbf{T}\} - \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{T}\}.
\end{aligned}$$

Com base nas expressões $\pi_i - \pi_j$, podemos comparar analiticamente os poderes locais dos testes rivais. A partir da equação (4.2), podemos observar que $\pi_1 > \pi_3$ se $K_2 \leq 0$ e $K_3 < 0$ e $\pi_1 < \pi_3$ se $K_2 \geq 0$ e $K_3 > 0$. Ainda, temos que não é possível a igualdade entre π_1 e π_3 , uma vez que $K_3 \neq 0$. Satisfeitas as condições acima, este resultado se observa para as diversas distribuições pertencentes à família simétrica. De forma análoga, podemos proceder para as demais diferenças entre os poderes locais dos testes rivais, isto é, temos que

- $\pi_1 > \pi_2$ se $K_1 \geq 0$, $K_2 \geq 0$ e $K_3 > 0$ e $\pi_1 < \pi_2$ se $K_1 \leq 0$, $K_2 \leq 0$ e $K_3 < 0$,
- $\pi_1 > \pi_4$ se $K_4 \geq 0$, $K_2 \leq 0$, $K_3 < 0$ e $\pi_1 < \pi_4$ se $K_4 \leq 0$, $K_2 \geq 0$, $K_3 > 0$,
- $\pi_2 > \pi_3$ se $K_5 \geq 0$, $K_2 \leq 0$, $K_3 < 0$ e $\pi_2 < \pi_3$ se $K_5 \leq 0$, $K_2 \geq 0$, $K_3 > 0$,
- $\pi_2 > \pi_4$ se $K_4 \leq 0$, $K_2 \leq 0$, $K_3 < 0$ e $\pi_2 < \pi_4$ se $K_4 \geq 0$, $K_2 \geq 0$, $K_3 > 0$,
- $\pi_3 > \pi_4$ se $K_4 \geq 0$, $K_2 \geq 0$, $K_3 > 0$ e $\pi_3 < \pi_4$ se $K_4 \leq 0$, $K_2 \leq 0$, $K_3 < 0$.

É válido salientar que as condições impostas aos K_ℓ , $\ell = 1, 2, 3, 4, 5$, podem ser verificadas numericamente em aplicações práticas após o ajuste do MNLSH aos dados. Um

questionamento pertinente é quanto ao desempenho dos testes quando estamos lidando com amostras de tamanho pequeno ou moderado e qual deles é mais confiável. Na Seção seguinte iremos avaliar o desempenho dos testes nestas condições via simulação de Monte Carlo.

4.4 Resultados numéricos

Nesta Seção desenvolvemos um estudo de simulação de Monte Carlo a fim de avaliar o desempenho dos testes da razão de verossimilhanças, Wald, escore e gradiente em amostras de tamanho pequeno e moderado. Para isto, consideramos o modelo de regressão tal como definido na Seção 3.4, isto é,

$$y_\ell = \beta_0 + \exp\{\beta_1 x_{\ell 1}\} + \sum_{s=2}^{p-1} \beta_s x_{\ell s} + e_\ell, \quad \ell = 1, \dots, n,$$

com e_ℓ sendo variáveis aleatórias independentes com distribuição simétrica, em notação, $e_\ell \sim S(0, \exp\{\mathbf{w}_\ell^\top \boldsymbol{\delta}\}, g)$. O interesse do estudo consiste em testar $H_0 : \boldsymbol{\delta}_1 = \mathbf{0}$, contra $H_1 : \text{pelo menos uma desigualdade}$, sendo $\boldsymbol{\delta}_1$ um vetor de dimensão $k - 1$. Sob H_0 , temos que $\exp\{\mathbf{w}_\ell^\top \boldsymbol{\delta}\} = \exp\{\delta_0\}$, indicando que o modelo é homoscedástico. As variáveis resposta foram geradas das distribuições t -Student ($\nu = 5$) e exponencial potência ($\kappa = 0, 3$). Assumimos $\beta_0 = \dots = \beta_{p-1} = 1$, $\delta_1 = 0, 1$; $\delta_2 = 0, 3$; $\delta_3 = 0, 5$ e $\delta_4 = \delta_5 = \delta_6 = 1$. As covariáveis $x_{\ell 1}, \dots, x_{\ell p-1}$ e $w_{\ell 2}, \dots, w_{\ell k-1}$ foram geradas da distribuição $U(0, 1)$. O número de réplicas de Monte Carlo foi fixado em 10.000 e os níveis nominais considerados foram $\alpha = 1\%$, 5% e 10% . Para as simulações, usamos a linguagem de programação matricial `0x` (Doornik, 2006). Os gráficos de poder foram feitos usando o software estatístico R. Nas Tabelas a seguir são apresentadas, em porcentagem, as taxas de rejeição dos testes em estudo para cada cenário simulado, isto é, estimamos via simulação de Monte Carlo $\mathbb{P}(S_i \geq \chi_{\alpha; k-1}^2)$, com $i = 1, 2, 3, 4$ e $\chi_{\alpha; k-1}^2$ o percentil $(1 - \alpha)$ da distribuição χ_{k-1}^2 .

Inicialmente avaliamos o efeito do aumento do número de parâmetros β no desempenho dos testes considerando $n = 30$, $k = 3$ e $p = 3, 4, 5, 6$. Suas taxas de rejeição nula são apresentadas na Tabela 4.1. Como podemos observar, dentre os quatro testes, o de pior desempenho é o teste Wald (S_2), apresentando taxas de rejeição muito acima dos níveis nominais adotados, chegando a valores sete vezes maior do que o esperado, por exemplo, quando $\alpha = 5\%$ e $k = 5$ temos para ambos os modelos a taxa de rejeição igual a 37% .

Os testes da razão de verossimilhanças (S_1) e gradiente (S_4) apresentam desempenho bastante similar e liberal, embora não tanto como o teste Wald. Ainda, estes três testes se mostram sensíveis quanto ao número de parâmetros β , isto é, conforme aumentamos a dimensão de β , mais distorcidos se tornam os testes. Já o teste escore (S_3), apresenta taxas bastante próximas dos níveis nominais considerados, tendo comportamento ligeiramente liberal, de maneira geral, e não se mostra sensível ao aumento do número de parâmetros β , apresentando taxas próximas aos níveis nominais adotados independente da dimensão de β . A exemplo, quando $\alpha = 10\%$ e $p = 3$ temos as taxas de rejeição para o modelo t_5 iguais a 19,4% (S_1), 30,6% (S_2), 10,7% (S_3) e 18,4% (S_4), fazendo $p = 6$, obtemos as taxas de rejeição iguais a 33,5% (S_1), 52,8% (S_2), 11,9% (S_3) e 31,2% (S_4).

Tabela 4.1: Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0,3$, considerando $n = 30$, $k = 3$ e diversos valores para p .

α	Teste	Modelo t_5				Modelo exponencial potência			
		p				p			
		3	4	5	6	3	4	5	6
10%	S_1	19,4	23,3	27,8	33,5	19,2	21,3	26,2	30,8
	S_2	30,6	36,4	45,7	52,8	32,3	37,3	45,6	52,8
	S_3	10,7	11,2	11,8	11,9	9,7	10,5	10,9	11,3
	S_4	18,4	22,3	26,1	31,2	18,5	20,6	24,8	29,4
5%	S_1	12,0	14,9	18,9	24,5	11,5	13,7	17,5	21,8
	S_2	22,3	28,3	37,0	44,7	23,9	28,5	37,0	45,0
	S_3	5,1	5,5	6,0	6,3	4,9	5,4	6,0	6,0
	S_4	10,8	14,1	17,7	22,2	10,8	13,1	16,6	20,2
1%	S_1	3,7	5,2	7,5	12,0	3,3	4,6	6,7	9,9
	S_2	11,2	16,5	24,3	32,4	12,7	16,6	23,6	31,3
	S_3	3,1	1,1	1,2	1,3	1,1	1,2	1,3	1,5
	S_4	0,9	4,8	6,7	10,2	2,9	4,3	6,1	8,7

Na Tabela 4.2 avaliamos a influência do número de parâmetros δ nos modelos em estudo, apresentando as taxas de rejeição dos testes sob investigação quando $n = 30$, $p = 3$ e $k = 3, 4, 5, 6$. Notamos que os testes avaliados têm comportamento similar ao observado quando aumentamos o número de parâmetros β , isto é, os testes S_1, S_2 e S_4 apresentam taxas de rejeição bastante distorcidas, sendo o teste S_2 o de pior desempenho. Além disso, conforme aumentamos a dimensão de δ , podemos observar que os testes apresentam

taxas de rejeição ainda mais distorcidas, mostrando a sensibilidade dos testes quando aumentamos a dimensão de δ . Além disso, o teste S_3 apresenta o melhor desempenho dentre os testes em estudo, não sendo influenciado pelo número de parâmetros δ , isto é, apresentando taxas próximas aos níveis nominais adotados independente do número de parâmetros. Por exemplo, para o modelo exponencial potência quando $\alpha = 1\%$ e $k = 4$ temos as respectivas taxas de rejeição para os teste em estudo: 4,4% (S_1), 17,8% (S_2), 1,3% (S_3) e 3,9% (S_4), considerando $k = 6$, temos 6,4% (S_1), 35,2% (S_2), 1,7% (S_3) e 5,3% (S_4).

Tabela 4.2: Tamanho dos testes para o Modelo t_5 e exponencial potência com $\kappa = 0,3$, considerando $n = 30$, $p = 3$ e diversos valores para k .

α	Teste	Modelo t_5				Modelo exponencial potência			
		k				k			
		3	4	5	6	3	4	5	6
10%	S_1	19,4	22,1	24,2	28,4	19,2	20,9	22,5	26,6
	S_2	30,6	36,5	42,3	54,0	32,3	37,9	44,8	58,1
	S_3	10,7	10,8	11,0	11,3	9,7	10,5	9,9	10,6
	S_4	18,4	21,2	21,7	25,6	18,5	20,1	20,6	24,3
5%	S_1	12,0	14,2	15,2	19,3	11,5	13,0	13,6	17,4
	S_2	22,3	28,0	33,5	45,8	23,9	29,7	36,1	49,5
	S_3	5,1	5,4	5,4	5,7	4,9	5,2	5,6	6,0
	S_4	10,8	13,1	12,8	16,6	10,8	12,4	12,0	15,6
1%	S_1	3,7	4,5	5,3	7,6	3,3	4,4	4,5	6,4
	S_2	11,2	16,6	19,9	31,2	12,7	17,8	22,3	35,2
	S_3	0,9	1,0	1,0	0,9	1,1	1,3	1,2	1,7
	S_4	3,1	3,9	3,8	5,8	2,9	3,9	3,7	5,3

Nas Tabelas 4.3 e 4.4 apresentamos as taxas de rejeição dos testes em estudo para os modelos t_5 e exponencial potência, respectivamente, considerando $p = 3$, $k = 4$ e diferentes tamanhos amostrais. Nosso objetivo aqui é verificar o comportamento dos quatro testes conforme o tamanho da amostra aumenta. Como já observamos nas Tabelas anteriores, para amostras de tamanho pequeno e moderado, o teste S_2 apresenta o pior desempenho, com taxas de rejeição muito acima dos níveis nominais considerados. Os testes S_1 e S_4 são liberais, embora não tanto como o teste S_2 , e têm desempenho similar. O teste S_3 apresenta melhor desempenho, com taxas de rejeição próximas dos níveis nominais

considerados para todos os tamanhos amostrais. Como era de se esperar, conforme o tamanho da amostra aumenta, os testes S_1 , S_2 e S_4 se apresentam menos discrepantes, isto é, com taxas de rejeição menos distorcidas, mais próximas aos níveis de significância adotados. Como exemplo, para $n = 25$ e $\alpha = 5\%$, temos as taxas de rejeição dos testes iguais a 17,7% (S_1), 38,4% (S_2), 5,1% (S_3) e 15,3% (S_4) para o modelo t_5 e 16,5% (S_1), 40,5% (S_2), 5,3% (S_3) e 4,6% (S_4) para o modelo exponencial potência. Aumentando n para $n = 120$, temos 6,5% (S_1), 8,1% (S_2), 5,2% (S_3) e 6,2% (S_4) para o modelo t_5 e 5,9% (S_1), 7,9% (S_2), 4,9% (S_3) e 5,8% (S_4) para o modelo exponencial potência.

Tabela 4.3: Tamanho dos testes para o Modelo t_5 com $p = 3$, $k = 4$ e diversos valores para n .

α	Teste	n							
		25	35	45	55	75	100	120	150
$\alpha = 10\%$	S_1	27,2	19,1	16,2	15,1	12,9	12,3	12,3	11,4
	S_2	47,3	30,2	24,5	20,8	16,9	15,2	14,9	12,9
	S_3	10,5	10,7	10,0	10,9	9,7	10,1	10,4	10,1
	S_4	24,4	17,9	15,3	14,6	12,3	12,0	12,1	11,2
$\alpha = 5\%$	S_1	17,7	11,5	9,5	8,8	6,6	6,4	6,5	5,7
	S_2	38,4	21,6	16,3	13,3	9,8	8,7	8,1	7,0
	S_3	5,1	5,0	4,9	5,6	4,5	4,7	5,2	4,6
	S_4	15,3	10,2	8,6	8,2	6,2	6,2	6,2	5,4
$\alpha = 1\%$	S_1	6,8	3,5	2,5	2,4	1,4	1,5	1,6	1,2
	S_2	24,9	10,5	7,0	5,1	2,6	2,5	2,0	1,5
	S_3	0,7	0,9	0,9	0,9	0,8	1,0	1,1	0,9
	S_4	5,2	2,6	2,1	2,0	1,2	1,3	1,5	1,1

Com o objetivo de estudar o poder dos quatro testes para amostras de tamanho pequeno e moderado, simulamos as taxas de rejeição dos testes em investigação sob a hipótese alternativa $H_1 : \delta_2 = \delta_3 = \delta_4 = \delta \neq 0$, considerando os modelos t_5 e exponencial potência, com $n = 35$, $p = 3$, $k = 4$, $\alpha = 10\%$ e δ variando entre -5.0 e 5.4 . Como podemos observar nas simulações anteriores, os testes têm diferentes tamanhos quando usamos a distribuição χ^2 em amostras pequenas. Para superar este problema e garantir que todos os testes tenham o mesmo tamanho, usamos 500.000 réplicas de Monte Carlo e estimamos, sob hipótese nula, os valores críticos de cada teste para cada nível nominal considerado. Nas Figuras 4.1 e 4.2 apresentamos as curvas de poder simuladas correspondentes aos quatro testes para os modelos t_5 e exponencial potência, respectivamente. Como podemos

observar, para ambos os modelos as curvas de poder dos quatro testes se sobrepuseram, praticamente, sendo a curva de poder do teste Wald um pouco abaixo das demais, principalmente para o modelo exponencial potência. Como esperado, os poderes dos testes se aproximam de 1 conforme $|\delta|$ aumenta.

Tabela 4.4: Tamanho dos testes para o Modelo exponencial potência com $\kappa = 0,3$, $p = 3$, $k = 4$ e diversos valores para n .

α	<i>Teste</i>	<i>n</i>							
		25	35	45	55	75	100	120	150
$\alpha = 10\%$	S_1	25,5	18,5	15,5	14,3	13,1	12,0	11,3	11,1
	S_2	49,1	34,6	26,5	21,8	18,2	16,1	14,6	13,3
	S_3	9,7	9,5	10,0	10,3	10,0	9,9	9,8	9,8
	S_4	23,1	10,8	15,2	14,0	12,8	11,8	11,2	10,9
$\alpha = 5\%$	S_1	16,5	10,7	9,2	7,8	7,0	6,1	5,9	5,9
	S_2	40,5	25,3	18,3	14,2	10,8	9,1	7,9	7,9
	S_3	5,3	5,0	5,1	5,1	4,9	5,0	4,9	4,9
	S_4	14,6	9,8	8,8	7,4	6,7	5,9	5,8	5,9
$\alpha = 1\%$	S_1	6,1	3,0	2,4	2,0	1,6	1,4	1,4	1,2
	S_2	27,0	13,0	8,3	5,2	3,6	2,6	2,5	1,8
	S_3	1,3	1,1	1,2	1,0	1,2	1,0	1,0	1,0
	S_4	5,0	2,8	2,1	1,8	1,5	1,4	1,3	1,2

Figura 4.1: Curva de poder dos quatro testes para o modelo t_5 com $n = 35$, $p = 3$, $k = 4$

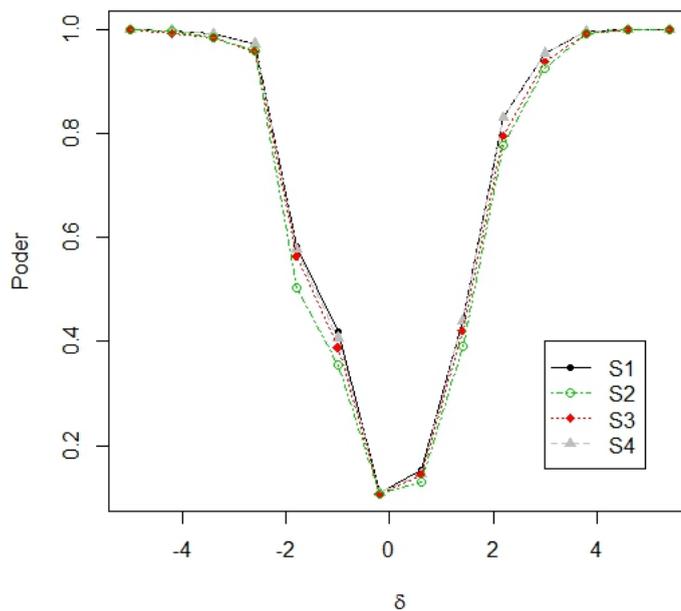
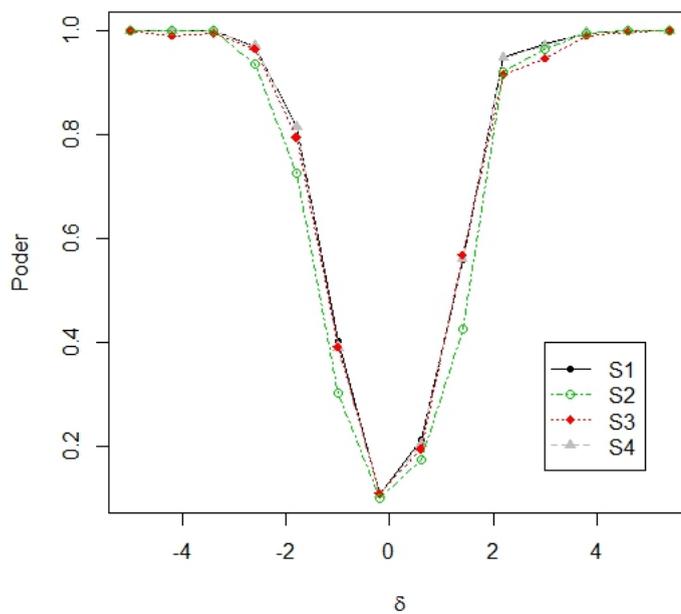


Figura 4.2: Curva de poder dos quatro testes para o modelo exponencial potência com $\kappa = 0,3$, $n = 35$, $p = 3$, $k = 4$



4.5 Conclusões

Neste capítulo, obtivemos expansões assintóticas, sob hipóteses alternativas de Pitman, das funções de distribuição das estatísticas da razão de verossimilhanças, Wald, escore e gradiente com o objetivo de comparar analiticamente o poder dos testes baseados em tais estatísticas até ordem $n^{-1/2}$. Nosso trabalho estende os resultados obtidos por [Lemonte \(2012\)](#), que compara o poder local dos quatro testes supracitados na classe dos modelos lineares simétricos. Além da comparação analítica, realizamos um estudo de simulação de Monte Carlo para avaliar o comportamento dos testes em estudo considerando amostras de tamanho pequeno e moderado. Os resultados do estudo de simulação de Monte Carlo adicionaram informações importantes. Para amostras de tamanho pequeno ou moderado, o teste escore foi o que apresentou o melhor desempenho dentre os quatro testes em estudo e o teste Wald apresentou o pior desempenho. Ainda, as simulações de poder sugeriram que, quando utilizado os valores críticos estimados corretos, os testes apresentam propriedades de poder semelhantes para pequenas amostras.

Considerações finais

As principais contribuições teóricas deste trabalho são enumeradas a seguir:

- No Capítulo 2, derivamos fatores de correção de Bartlett para os testes de heteroscedasticidade baseados nas estatísticas da razão de verossimilhanças e razão de verossimilhanças perfiladas modificadas em modelos não-lineares simétricos heteroscedásticos (MNLSH). Os resultados obtidos estendem os trabalhos de [Cordeiro \(2004\)](#), que obteve correções de Bartlett para a estatística da razão de verossimilhanças na classe dos modelos não-lineares simétricos e [Ferrari et al. \(2004\)](#), que derivaram um fator de correção de Bartlett para o teste de heteroscedasticidade baseado na estatística da razão de verossimilhanças perfiladas modificadas para o modelo normal linear heteroscedástico. Os estudos de simulação desenvolvidos apontaram que as correções de Bartlett foram efetivas, melhorando a aproximação da distribuição da estatística de teste pela distribuição χ^2 de referência, tendo o teste da razão de verossimilhanças perfiladas modificadas o melhor desempenho dentre os teste comparados, juntamente com o teste da razão de verossimilhanças Bartlett bootstrap;
- No Capítulo 3, obtivemos um fator de correção tipo-Bartlett para o teste de heteroscedasticidade baseado na estatística gradiente na classe dos MNLSH. Este é o primeiro trabalho a tratar de correções tipo-Bartlett para a estatística gradiente para modelos que têm erros seguindo distribuição simétrica. Ainda, estendemos os

resultados de Nascimento (2010), que avaliou apenas o teste escore corrigido para a média na classe dos MNLSH. Além dos testes escore e gradiente corrigidos e não corrigidos, consideramos para o estudo de simulação testes baseados em estatísticas modificadas, alternativas à estatística escore, propostas por Kakisawa (1996) e Cordeiro et al. (1998). Os resultados numéricos apontaram que os testes baseados nas estatísticas escore modificadas e corrigida e na estatística gradiente corrigida apresentaram desempenhos melhores do que os testes baseados em suas respectivas versões não corrigidas, tendo os testes baseados nas estatísticas escore modificadas e corrigida melhores desempenhos, de maneira geral;

- No Capítulo 4, estendemos os resultados obtidos por Lemonte (2012) para a classe dos MNLSH, isto é, realizamos para esta classe de modelos um estudo de poder local com o objetivo de estabelecer condições para poder comparar (localmente) o poder dos testes de heteroscedasticidade baseados nas estatísticas da razão de verossimilhanças, Wald, escore e gradiente. Em aplicações práticas, tais condições podem ser verificadas numericamente após o ajuste do MNLSH aos dados. Para avaliar o desempenho dos testes em pequenas amostras, foi realizado um estudo de simulação de Monte Carlo, cujos resultados numéricos apontaram que o teste escore tem o melhor desempenho e o teste Wald o pior desempenho dentre os quatro testes. Ainda, os resultados numéricos sugeriram que, quando utilizado os valores críticos estimados corretos, os testes apresentam propriedades de poder semelhantes para pequenas amostras.

Aspectos inferenciais do modelos não-lineares simétricos heteroscedásticos

Neste apêndice, apresentamos os cálculos das derivadas para a obtenção da função escore total do modelo não-linear simétrico heteroscedástico bem como sua matriz de informação total de Fisher. Como definido na Seção 2.2, o logaritmo da função de verossimilhança do vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top, \sigma^2)^\top$ do MNLSH definido em (2.4) é dado por

$$l(\boldsymbol{\theta}) = l(\boldsymbol{\theta}; \mathbf{y}) = -\frac{n}{2} \log \sigma^2 - \frac{1}{2} \sum_{\ell=1}^n \log m_\ell + \sum_{\ell=1}^n t(z_\ell). \quad (\text{A.1})$$

Para representar as derivadas do logaritmo da função de verossimilhança, definimos as derivadas de μ_ℓ com relação aos componentes de $\boldsymbol{\beta}$ por $(i)_\ell = \partial \mu_\ell / \partial \beta_i$, $(ij)_\ell = \partial^2 \mu_\ell / \partial \beta_i \partial \beta_j$, $(ij, l)_\ell = (\partial^2 \mu_\ell / \partial \beta_i \partial \beta_j)(\partial \mu_\ell / \partial \beta_l)$, etc, e as derivadas de m_ℓ com relação aos componentes de $\boldsymbol{\delta}$ por $\psi_{\ell a} = \partial m_\ell / \partial \delta_a$, $\psi_{\ell ab} = \partial^2 m_\ell / \partial \delta_a \partial \delta_b$, etc, onde os índices i, j, l, \dots variam em $\boldsymbol{\beta}$ e os índices a, b, c, \dots variam em $\boldsymbol{\delta}$.

Diferenciando (A.1) com relação aos componentes de $\boldsymbol{\theta}$, temos:

$$\begin{aligned}\frac{\partial l(\boldsymbol{\theta})}{\partial \delta_a} &= -\frac{1}{2} \sum_{\ell=1}^n \frac{1}{m_\ell} \psi_{\ell a} - \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{1}{m_\ell} \psi_{\ell a}, \\ \frac{\partial l(\boldsymbol{\theta})}{\partial \beta_i} &= -\sum_{\ell=1}^n t(z_\ell)^{(1)} \frac{1}{\sqrt{\sigma^2 m_\ell}} (i)_\ell \text{ e} \\ \frac{\partial l(\boldsymbol{\theta})}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} - \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{1}{2\sigma^2},\end{aligned}$$

com $a = 1, \dots, k$ e $i = 1, \dots, p$. Em notação matricial,

$$\begin{aligned}\mathbf{U}_\delta &= -\frac{1}{2} \Psi^\top (M\boldsymbol{\nu} - SM\mathbf{u}), \\ \mathbf{U}_\beta &= \tilde{X} S \Phi(\mathbf{y} - \boldsymbol{\mu}) \text{ e} \\ U_{\sigma^2} &= -\frac{n}{\sigma^2} + \frac{1}{2\sigma^2} (\mathbf{y} - \boldsymbol{\mu})^\top S \Phi(\mathbf{y} - \boldsymbol{\mu}).\end{aligned}$$

As segundas derivadas de (A.1) com respeito aos componentes de $\boldsymbol{\theta}$ são expressas por

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b} &= \frac{1}{2} \sum_{\ell=1}^n \frac{1}{m_\ell^2} \psi_{\ell a} \psi_{\ell b} - \frac{1}{2} \sum_{\ell=1}^n \frac{1}{m_\ell} \psi_{\ell ab} + \frac{1}{4} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 \frac{1}{m_\ell} \psi_{\ell a} \psi_{\ell b} \\ &+ \frac{1}{4} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{1}{m_\ell^2} \psi_{\ell a} \psi_{\ell b} + \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(1)} \frac{1}{m_\ell^2} \psi_{\ell a} \psi_{\ell b} \\ &- \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{1}{m_\ell} \psi_{\ell ab}, \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j} &= \sum_{\ell=1}^n t(z_\ell)^{(2)} \frac{1}{\sigma^2 m_\ell} (j)_\ell (i)_\ell - \frac{1}{\sqrt{\sigma^2}} \sum_{\ell=1}^n t(z_\ell)^{(1)} m_\ell^{-1/2} (ij)_\ell, \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial (\sigma^2)^2} &= \frac{n}{2\sigma^4} + \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 \frac{1}{4\sigma^4} + \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{1}{4\sigma^4} + \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{1}{2\sigma^4}, \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta_a \partial \sigma^2} &= \frac{1}{4\sigma^2} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 \frac{1}{m_\ell} \psi_{\ell a} + \frac{1}{4\sigma^2} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{1}{m_\ell} \psi_{\ell a}, \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_i \partial \delta_a} &= \sum_{\ell=1}^n \{t(z_\ell)^{(2)} z_\ell + t(z_\ell)^{(1)}\} \frac{1}{2m_\ell} \frac{1}{\sqrt{\sigma^2 m_\ell}} \psi_{\ell a} (i)_\ell \text{ e} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_i \partial \sigma^2} &= \sum_{\ell=1}^n \{t(z_\ell)^{(2)} z_\ell + t(z_\ell)^{(1)}\} \frac{1}{2\sigma^2} \frac{1}{\sqrt{\sigma^2 m_\ell}} (i)_\ell,\end{aligned}$$

com $a, b = 1, \dots, k$ e $i, j = 1, \dots, p$. Tomando as esperanças das expressões acima, temos:

$$\begin{aligned}
E \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b} \right] &= \frac{(\delta_{(0,1,0,0,2)} - 1)}{4} \sum_{\ell=1}^n \frac{1}{m_\ell^2} \psi_{\ell a} \psi_{\ell b}, \\
E \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j} \right] &= \frac{\delta_{(0,1,0,0,0)}}{\sigma^2} \sum_{\ell=1}^n \frac{1}{m_\ell} (i, j)_\ell, \\
E \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial (\sigma^2)^2} \right] &= \frac{n}{4\sigma^4} \delta_{(0,1,0,0,2)} - \frac{n}{4\sigma^2}, \\
E \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta_a \partial \sigma^2} \right] &= \left(\frac{\delta_{(0,1,0,0,2)} - 1}{4} \right) \sum_{\ell=1}^n \frac{\psi_{\ell a}}{\sigma^2 m_\ell} \text{ e} \\
E \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_i \partial \delta_a} \right] &= E \left(\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_i \partial \sigma^2} \right) = 0.
\end{aligned}$$

Assim, obtemos a matriz de informação de Fisher que é dada por

$$I = -E \left(\frac{\partial^2 l}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right) = \begin{pmatrix} \left(\frac{1 - \delta_{(0,1,0,0,2)}}{4} \right) \Psi^\top M^{(2)} \Psi & \mathbf{0} & \left(\frac{1 - \delta_{(0,1,0,0,2)}}{4} \right) \Psi^\top \Phi \boldsymbol{\nu} \\ \mathbf{0} & -\delta_{(0,1,0,0,0)} \tilde{X}^\top \Phi \tilde{X} & \mathbf{0} \\ \left(\frac{1 - \delta_{(0,1,0,0,2)}}{4} \right) \boldsymbol{\nu}^\top \Phi \Psi & \mathbf{0} & (1 - \delta_{(0,1,0,0,2)}) \frac{n}{4\sigma^4} \end{pmatrix}.$$

Cálculo do fator de correção tipo-Bartlett para a estatística LR

Neste apêndice obtemos algumas derivadas, e momentos das derivadas, do logaritmo da função de verossimilhança de $\boldsymbol{\theta}^* = (\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top, \gamma)^\top$ do modelo reparametrizado expresso por

$$l^*(\boldsymbol{\theta}) = l^*(\boldsymbol{\theta}^*; \mathbf{y}) = -\frac{n}{2} \log \gamma + \sum_{\ell=1}^n t(z_\ell) \quad (\text{B.1})$$

para a obtenção da constante c em (2.7) do fator de correção para a estatística LR . Desse modo, definimos a seguinte notação adicional: $q_{\ell a} = \frac{\partial q_\ell}{\partial \delta_a}$, $q_{\ell ab} = \frac{\partial q_{\ell a}}{\partial \delta_b}$, etc. Além disso, consideremos que os índices a, b, c, \dots variam em $\boldsymbol{\delta}$ e os índices i, j, l, \dots variam em $\boldsymbol{\beta}$.

Assim, as quatro primeiras derivadas de (B.1) com respeito aos componentes de $\boldsymbol{\delta}$ são dadas por

$$\begin{aligned} \frac{\partial l^*(\boldsymbol{\theta})}{\partial \delta_a} &= \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{1}{q_\ell} q_{\ell a}, \\ \frac{\partial^2 l^*(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b} &= \frac{1}{2} \sum_{\ell=1}^n \left\{ t(z_\ell)^{(2)} z_\ell^2 \frac{1}{2q_\ell^2} q_{\ell a} q_{\ell b} - t(z_\ell)^{(1)} z_\ell \frac{1}{2q_\ell^2} q_{\ell a} q_{\ell b} \right. \\ &\quad \left. + t(z_\ell)^{(1)} z_\ell \frac{1}{q_\ell} q_{\ell ab} \right\}, \\ \frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \delta_c} &= \frac{1}{2} \sum_{\ell=1}^n \left\{ t(z_\ell)^{(3)} z_\ell^3 \frac{1}{4q_\ell^3} q_{\ell a} q_{\ell b} q_{\ell c} - t(z_\ell)^{(2)} z_\ell^2 \frac{3}{4q_\ell^3} q_{\ell a} q_{\ell b} q_{\ell c} \right. \\ &\quad + t(z_\ell)^{(2)} z_\ell^2 \frac{3}{4q_\ell^3} [q_{\ell ac} q_{\ell b} + q_{\ell a} q_{\ell bc} + q_{\ell ab} q_{\ell c}] + t(z_\ell)^{(1)} z_\ell \frac{3}{4q_\ell^3} q_{\ell a} q_{\ell b} q_{\ell c} \\ &\quad \left. - t(z_\ell)^{(1)} z_\ell \frac{1}{2q_\ell^2} [q_{\ell ac} q_{\ell b} + q_{\ell a} q_{\ell bc} + q_{\ell ab} q_{\ell c}] + t(z_\ell)^{(1)} z_\ell \frac{1}{q_\ell} q_{\ell abc} \right\} e \end{aligned}$$

$$\begin{aligned}
\frac{\partial^4 l^*(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \delta_c \partial \delta_d} &= \frac{1}{2} \sum_{\ell=1}^n \left\{ t(z_\ell)^{(4)} z_\ell^4 \frac{1}{4q_\ell^4} q_{la} q_{lb} q_{lc} q_{ld} - t(z_\ell)^{(3)} z_\ell^3 \frac{3}{4q_\ell^4} q_{la} q_{lb} q_{lc} q_{ld} \right. \\
&- t(z_\ell)^{(3)} z_\ell^3 \frac{1}{q_\ell^3} [q_{lad} q_{lb} q_{lc} + q_{la} q_{lbd} q_{lc} + q_{la} q_{lb} q_{lcd} + q_{lac} q_{lb} q_{ld} \\
&+ q_{la} q_{lbc} q_{ld} + q_{lab} q_{lc} q_{ld}] \\
&+ t(z_\ell)^{(2)} z_\ell^2 \frac{15}{8q_\ell^4} q_{la} q_{lb} q_{lc} q_{ld} - t(z_\ell)^{(2)} z_\ell^2 \frac{9}{2q_\ell^3} q_{lad} q_{lb} q_{lc} \\
&+ t(z_\ell)^{(1)} z_\ell \frac{9}{2q_\ell^3} q_{lad} q_{lb} q_{lc} - t(z_\ell)^{(1)} z_\ell \frac{3}{q_\ell^2} q_{lacd} q_{lb} \\
&+ t(z_\ell)^{(2)} z_\ell^2 \frac{1}{2q_\ell^2} q_{labcd} q_{ld} - t(z_\ell)^{(1)} z_\ell \frac{1}{2q_\ell^2} q_{labcd} q_{ld} \\
&+ t(z_\ell)^{(1)} z_\ell \frac{1}{q_\ell} q_{labcd} + t(z_\ell)^{(2)} z_\ell^2 \frac{3}{q_\ell^2} q_{lacd} q_{lb} \\
&\left. - t(z_\ell)^{(1)} z_\ell \frac{15}{8q_\ell^4} q_{la} q_{lb} q_{lc} q_{ld} \right\}.
\end{aligned}$$

As três primeiras derivadas de (B.1) com respeito aos componentes de $\boldsymbol{\beta}$ são dadas por

$$\begin{aligned}
\frac{\partial l^*(\boldsymbol{\theta})}{\partial \beta_i} &= -\frac{1}{\gamma^{1/2}} \sum_{\ell=1}^n t(z_\ell)^{(1)} q_\ell^{1/2} (i)_\ell, \quad i = 1, \dots, p, \\
\frac{\partial^2 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j} &= \frac{1}{\gamma} \sum_{\ell=1}^n t(z_\ell)^{(2)} q_\ell (i, j)_\ell - \frac{1}{\gamma^{1/2}} \sum_{\ell=1}^n t(z_\ell)^{(1)} q_\ell^{1/2} (ij)_\ell \text{ e} \\
\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \beta_l} &= -\frac{1}{\gamma^{3/2}} \sum_{\ell=1}^n t(z_\ell)^{(3)} q_\ell^{3/2} (i, j, l)_\ell - \frac{1}{\gamma^{1/2}} \sum_{\ell=1}^n t(z_\ell)^{(1)} q_\ell^{1/2} (ijl)_\ell \\
&+ \frac{1}{\gamma} \sum_{\ell=1}^n t(z_\ell)^{(2)} q_\ell \{(ij, l)_\ell + (i, jl)_\ell + (il, j)_\ell\}.
\end{aligned}$$

As três primeiras derivadas de (B.1) com respeito a γ são dadas por

$$\begin{aligned}
\frac{\partial l^*(\boldsymbol{\theta})}{\partial \gamma} &= -\frac{n}{2\gamma} - \frac{1}{2\gamma} \sum_{\ell=1}^n t(z_\ell)^{(1)}, \\
\frac{\partial^2 l^*(\boldsymbol{\theta})}{\partial \gamma^2} &= \frac{n}{2\gamma^2} + \frac{3}{4\gamma^2} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell + \frac{1}{4\gamma^2} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 \text{ e} \\
\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \gamma^3} &= -\frac{n}{\gamma^3} - \frac{15}{12\gamma^3} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell - \frac{9}{8\gamma^3} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 - \frac{1}{8\gamma^3} \sum_{\ell=1}^n t(z_\ell)^{(3)} z_\ell^3.
\end{aligned}$$

Algumas das derivadas mistas de (B.1) são dadas por

$$\begin{aligned}
\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \beta_i} &= \frac{1}{2} \sum_{\ell=1}^n \left\{ -t(z_\ell)^{(3)} z_\ell^2 \frac{1}{2\gamma^{1/2} q_\ell^{3/2}} (i)_\ell q_{\ell a} q_{\ell b} - t(z_\ell)^{(2)} z_\ell \frac{1}{2\gamma^{1/2} q_\ell^{3/2}} (i)_\ell q_{\ell a} q_{\ell b} \right. \\
&+ t(z_\ell)^{(1)} \frac{1}{2\gamma^{1/2} q_\ell^{3/2}} (i)_\ell q_{\ell a} q_{\ell b} - t(z_\ell)^{(2)} z_\ell \frac{1}{\gamma^{1/2} q_\ell^{3/2}} (i)_\ell q_{\ell a} b \\
&\left. - t(z_\ell)^{(1)} \frac{1}{\gamma^{1/2} q_\ell^{3/2}} (i)_\ell q_{\ell a} b \right\}, \\
\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \gamma} &= \frac{1}{2} \sum_{\ell=1}^n \left\{ -t(z_\ell)^{(3)} z_\ell^3 \frac{1}{4\gamma q_\ell^2} q_{\ell a} q_{\ell b} - t(z_\ell)^{(2)} z_\ell^2 \frac{1}{4\gamma q_\ell^2} q_{\ell a} q_{\ell b} \right. \\
&+ t(z_\ell)^{(1)} z_\ell \frac{1}{4\gamma q_\ell^2} q_{\ell a} q_{\ell b} - t(z_\ell)^{(2)} z_\ell^2 \frac{1}{2\gamma q_\ell} w_{\ell a} b - t(z_\ell)^{(1)} z_\ell \frac{1}{2\gamma q_\ell} q_{\ell a} b \left. \right\}, \\
\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \delta_a} &= \sum_{\ell=1}^n \left\{ t(z_\ell)^{(3)} z_\ell \frac{1}{2\gamma} q_{\ell a} (i, j)_\ell + t(z_\ell)^{(2)} \frac{1}{\gamma} q_{\ell a} (i, j)_\ell \right. \\
&\left. - t(z_\ell)^{(2)} z_\ell \frac{1}{2\gamma^{1/2} q_\ell} q_{\ell a} q_\ell^{1/2} (ij)_\ell - t(z_\ell)^{(1)} \frac{1}{\gamma^{1/2} q_\ell} q_{\ell a} (ij)_\ell \right\}, \\
\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \gamma \partial \gamma \partial \delta_a} &= \frac{1}{8\gamma^2} \sum_{\ell=1}^n \left\{ 5t(z_\ell)^{(2)} z_\ell^2 + 3t(z_\ell)^{(1)} z_\ell + t(z_\ell)^{(3)} z_\ell^3 \right\} \frac{q_{\ell a}}{q_\ell}, \\
\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_l \partial \gamma} &= \sum_{\ell=1}^n \left\{ -t(z_\ell)^{(2)} \frac{1}{\gamma^2} q_\ell (i, l)_\ell - t(z_\ell)^{(3)} z_\ell \frac{1}{2\gamma^2} q_\ell (i, l)_\ell \right. \\
&+ t(z_\ell)^{(1)} \frac{1}{2\gamma^{3/2}} q_\ell^{1/2} (il)_\ell + t(z_\ell)^{(2)} z_\ell \frac{1}{2\gamma^{3/2}} q_\ell^{1/2} (il)_\ell \left. \right\}, \\
\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \delta_a \partial \gamma} &= \sum_{\ell=1}^n \left\{ t(z_\ell)^{(2)} z_\ell \frac{1}{4\gamma^{3/2} q_\ell^{1/2}} q_{\ell a} (i)_\ell + t(z_\ell)^{(3)} z_\ell^2 \frac{1}{4\gamma^{3/2} q_\ell^{1/2}} q_{\ell a} (i)_\ell \right. \\
&+ t(z_\ell)^{(2)} z_\ell \frac{1}{4\gamma^{3/2} q_\ell^{1/2}} q_{\ell a} (i)_\ell + t(z_\ell)^{(1)} \frac{1}{4\gamma^{3/2} q_\ell^{1/2}} q_{\ell a} (i)_\ell \\
&\left. + t(z_\ell)^{(2)} z_\ell \frac{1}{4\gamma^{3/2} q_\ell^{1/2}} q_{\ell a} (i)_\ell \right\}, \\
\frac{\partial^4 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \delta_a \partial \delta_b} &= \frac{1}{2} \sum_{\ell=1}^n \left\{ t(z_\ell)^{(4)} z_\ell^2 \frac{1}{2\gamma q_\ell} q_{\ell a} q_{\ell b} (i, j)_\ell + t(z_\ell)^{(3)} z_\ell \frac{1}{\gamma q_\ell} q_{\ell a} q_{\ell b} (i, j)_\ell \right. \\
&+ t(z_\ell)^{(3)} z_\ell \frac{1}{2\gamma q_\ell} q_{\ell a} q_{\ell b} (i, j)_\ell + t(z_\ell)^{(2)} \frac{1}{2\gamma q_\ell} q_{\ell a} q_{\ell b} (i, j)_\ell \\
&- t(z_\ell)^{(2)} \frac{1}{2\gamma q_\ell} q_{\ell a} q_{\ell b} (i, j)_\ell + t(z_\ell)^{(3)} z_\ell \frac{1}{\gamma q_\ell} q_{\ell a} b (i, j)_\ell \\
&\left. + t(z_\ell)^{(2)} \frac{1}{\gamma q_\ell} q_{\ell a} b (i, j)_\ell + t(z_\ell)^{(2)} \frac{1}{\gamma q_\ell} q_{\ell a} b (i, j)_\ell \right\} e
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^4 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \gamma \partial \gamma} &= \sum_{\ell=1}^n \left\{ t(z_\ell)^{(2)} \frac{2}{\gamma^3} q_\ell(i, l)_\ell + t(z_\ell)^{(3)} z_\ell \frac{1}{2\gamma^3} q_\ell(i, l)_\ell - t(z_\ell)^{(3)} z_\ell \frac{1}{\gamma^3} q_\ell(i, l)_\ell \right. \\
&+ t(z_\ell)^{(4)} z_\ell^2 \frac{1}{4\gamma^3} q_\ell(i, l)_\ell + t(z_\ell)^{(3)} z_\ell \frac{1}{4\gamma^3} q_\ell(i, l)_\ell + t(z_\ell)^{(1)} \frac{3}{4\gamma^{5/2}} q_\ell^{1/2}(il)_\ell \\
&- t(z_\ell)^{(2)} z_\ell \frac{1}{4\gamma^{5/2}} q_\ell^{1/2}(il)_\ell - t(z_\ell)^{(2)} z_\ell \frac{3}{4\gamma^{5/2}} q_\ell^{1/2}(il)_\ell - t(z_\ell)^{(3)} z_\ell^2 \frac{1}{4\gamma^{5/2}} q_\ell^{1/2}(il)_\ell \\
&\left. - t(z_\ell)^{(2)} z_\ell \frac{1}{4\gamma^{5/2}} q_\ell^{1/2}(il)_\ell \right\}.
\end{aligned}$$

A esperança das derivadas acima são expressas por

$$\begin{aligned}
\kappa_{ab} &= E \left[\frac{\partial^2 l^*(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b} \right] = \frac{1}{2} \sum_{\ell=1}^n \left\{ \delta_{(0,1,0,0,2)} \frac{1}{2q_\ell^2} q_{la} q_{lb} - \delta_{(1,0,0,0,1)} \frac{1}{2q_\ell^2} q_{la} q_{lb} - \delta_{(1,0,0,0,1)} \frac{1}{q_\ell} q_{lab} \right\}, \\
\kappa_{abc} &= E \left[\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \delta_c} \right] = \sum_{\ell=1}^n \left\{ \delta_{(0,0,1,0,3)} \frac{1}{8q_\ell^3} q_{la} q_{lb} q_{lc} - \delta_{(0,1,0,0,2)} \frac{3}{8q_\ell^3} q_{la} q_{lb} q_{lc} \right. \\
&+ \delta_{(0,1,0,0,2)} \frac{1}{4q_\ell^2} [q_{lac} q_{lb} + q_{la} q_{lbc} + q_{lab} q_{lc}] + \delta_{(1,0,0,0,1)} \frac{1}{2q_\ell} q_{labc} + \delta_{(1,0,0,0,1)} \frac{3}{8q_\ell^3} q_{la} q_{lb} q_{lc} \\
&\left. - \delta_{(1,0,0,0,1)} \frac{1}{4q_\ell^2} [q_{lac} q_{lb} + q_{la} q_{lbc} + q_{lab} q_{lc}] \right\}, \\
\kappa_{abcd} &= E \left[\frac{\partial^4 l^*(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \delta_c \partial \delta_d} \right] = \sum_{\ell=1}^n \left\{ \delta_{(0,0,0,1,4)} \frac{1}{16q_\ell^4} q_{la} q_{lb} q_{lc} q_{ld} - \frac{3}{8} \delta_{(0,0,1,0,3)} \left[\frac{1}{q_\ell^4} q_{la} q_{lb} q_{lc} q_{ld} \right. \right. \\
&+ \left. \frac{6}{q_\ell^3} q_{lad} q_{lb} q_{lc} \right] + \delta_{(0,1,0,0,2)} \left[\frac{16}{16q_\ell^4} q_{la} q_{lb} q_{lc} q_{ld} - \frac{9}{4q_\ell^3} q_{lad} q_{lb} q_{lc} + \frac{1}{4q_\ell^2} q_{labcd} q_{ld} \right. \\
&+ \left. \frac{3}{2q_\ell^2} q_{lacd} q_{lb} \right] + \delta_{(1,0,0,0,1)} \left[\frac{9}{4q_\ell^3} q_{lad} q_{lb} q_{lc} - \frac{3}{2q_\ell^2} q_{lacd} q_{lb} - \frac{1}{4q_\ell^2} q_{labcd} q_{lc} \right. \\
&\left. + \frac{1}{2q_\ell} q_{labcd} - \frac{15}{16q_\ell^4} q_{la} q_{lb} q_{lc} q_{ld} \right] \right\}, \\
\kappa_{ij} &= E \left[\frac{\partial^2 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j} \right] = \frac{1}{\gamma} \delta_{(0,1,0,0,0)} \sum_{\ell=1}^n q_\ell(i, j)_\ell, \\
\kappa_{ijl} &= E \left[\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \beta_l} \right] = \frac{1}{\gamma} \delta_{(0,1,0,0,0)} \sum_{\ell=1}^n q_\ell \{ (ij, l)_\ell + (i, jl)_\ell + (il, j)_\ell \}, \\
\kappa_{\gamma\gamma} &= E \left[\frac{\partial^2 l^*(\boldsymbol{\theta})}{\partial \gamma^2} \right] = \frac{n}{4\gamma^2} \{ 2 + 3\delta_{(1,0,0,0,1)} + \delta_{(0,1,0,0,2)} \}, \\
\kappa_{\gamma\gamma\gamma} &= E \left[\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \gamma^3} \right] = -\frac{n}{8\gamma^3} \{ 8 + 15\delta_{(1,0,0,0,1)} + 9\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)} \},
\end{aligned}$$

$$\begin{aligned}
\kappa_{ab\gamma} &= E \left[\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \gamma} \right] = \frac{1}{4\gamma} \sum_{\ell=1}^n \left\{ -\delta_{(0,0,1,0,3)} \frac{1}{2q_\ell^2} q_{\ell a} q_{\ell b} - \delta_{(0,1,0,0,2)} \frac{1}{2q_\ell^2} q_{\ell a} q_{\ell b} \right. \\
&\quad \left. + \delta_{(1,0,0,0,1)} \frac{1}{2q_\ell^2} q_{\ell a} q_{\ell b} - \delta_{(0,1,0,0,2)} \frac{1}{q_\ell} q_{\ell ab} - \delta_{(1,0,0,0,1)} \frac{1}{q_\ell} q_{\ell ab} \right\}, \\
\kappa_{ija} &= E \left[\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \delta_a} \right] = \frac{1}{2\gamma} \sum_{\ell=1}^n \left\{ \delta_{(0,0,1,0,1)} q_{\ell a}(i, j)_\ell + 2\delta_{(0,1,0,0,0)} q_{\ell a}(i, j)_\ell \right\}, \\
\kappa_{\gamma\gamma a} &= E \left[\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \gamma \partial \gamma \partial \delta_a} \right] = \frac{1}{8\gamma^2} \left\{ 5\delta_{(0,1,0,0,2)} + 3\delta_{(1,0,0,0,1)} + \delta_{(0,0,1,0,3)} \right\} \sum_{\ell=1}^n \frac{q_{\ell a}}{q_\ell}, \\
\kappa_{ij\gamma} &= E \left[\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \gamma} \right] = -\frac{1}{2\gamma^2} \left\{ 2\delta_{(0,1,0,0,0)} + \delta_{(0,0,1,0,1)} \right\} \sum_{\ell=1}^n q_\ell(i, j)_\ell, \\
\kappa_{ijab} &= E \left[\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \delta_a \partial \delta_b} \right] = \frac{1}{4\gamma} \sum_{\ell=1}^n \left\{ \delta_{(0,0,0,1,2)} \frac{1}{q_\ell} q_{\ell a} q_{\ell b}(i, j)_\ell + 3\delta_{(0,0,1,0,1)} \frac{1}{q_\ell} q_{\ell a} q_{\ell b}(i, j)_\ell \right. \\
&\quad \left. + 2\delta_{(0,0,1,0,1)} \frac{1}{q_\ell} q_{\ell ab}(i, j)_\ell + 4\delta_{(0,1,0,0,0)} \frac{1}{q_\ell} q_{\ell ab}(i, j)_\ell \right\}, \\
\kappa_{ij\gamma\gamma} &= E \left[\frac{\partial^3 l^*(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \gamma \partial \gamma} \right] = \frac{1}{4\gamma^3} \left\{ 8\delta_{(0,1,0,0,0)} - \delta_{(0,0,1,0,1)} + \delta_{(0,0,0,1,2)} \right\} \sum_{\ell=1}^n q_\ell(i, j)_\ell \text{ e} \\
\kappa_{iab} &= \kappa_{ia\gamma} = \kappa_{\gamma\gamma r} = 0.
\end{aligned}$$

Ainda, como consequência da ortogonalidade global entre os parâmetros de interesse e perturbação, tem-se que $\kappa_{ai} = \kappa_{\gamma a} = \kappa_{i\gamma} = 0$.

Algumas derivadas das esperanças acima são expressas por

$$\begin{aligned}
(\kappa_{ab})_c &= \frac{\partial \kappa_{ab}}{\partial \delta_c} = -\frac{1}{4} \delta_{(0,1,0,0,2)} \sum_{\ell=1}^n \left\{ \frac{2}{q_\ell} q_{\ell a} q_{\ell b} q_{\ell c} - \frac{1}{q_\ell^2} [q_{\ell ac} q_{\ell b} + q_{\ell a} q_{\ell bc}] \right\} \\
&\quad + \frac{1}{4} \delta_{(1,0,0,0,1)} \sum_{\ell=1}^n \left\{ \frac{2}{q_\ell^3} q_{\ell a} q_{\ell b} q_{\ell c} - \frac{1}{q_\ell^2} [q_{\ell ac} q_{\ell b} + q_{\ell a} q_{\ell bc} - 2q_{\ell ab} q_{\ell c}] + \frac{1}{q_\ell q_{\ell abc}} \right\}, \\
(\kappa_{ij})_l &= \frac{\partial \kappa_{ij}}{\partial \beta_l} = \frac{1}{\gamma} \sum_{\ell=1}^n q_\ell \{ (il, j)_\ell + (i, jl)_\ell \}, \\
(\kappa_{ab\gamma})_\gamma &= \frac{\partial \kappa_{ab\gamma}}{\partial \gamma} = \frac{1}{8\gamma^2} \delta_{(0,0,1,0,3)} \sum_{\ell=1}^n \frac{1}{q_\ell^2} q_{\ell a} q_{\ell b} + \frac{1}{8\gamma^2} \delta_{(0,1,0,0,2)} \sum_{\ell=1}^n \left\{ \frac{1}{q_\ell^2} q_{\ell a} q_{\ell b} + \frac{2}{q_\ell} q_{\ell ab} \right\} \\
&\quad - \frac{1}{8\gamma^2} \delta_{(1,0,0,0,1)} \sum_{\ell=1}^n \left\{ \frac{1}{q_\ell^2} q_{\ell a} q_{\ell b} + \frac{2}{q_\ell} q_{\ell ab} \right\}, \\
(\kappa_{ij})_a &= \frac{\partial \kappa_{ij}}{\partial \delta_a} = \frac{1}{\gamma} \delta_{(0,1,0,0,0)} \sum_{\ell=1}^n q_{\ell a}(i, j)_\ell,
\end{aligned}$$

$$\begin{aligned}
(\kappa_{ija})_b &= \frac{\partial \kappa_{ija}}{\partial \delta_b} = \frac{1}{2\gamma} \{ \delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)} \} \sum_{\ell=1}^n q_{\ell ab}(i, j)_\ell, \\
(\kappa_{\gamma\gamma a})_b &= \frac{\partial \kappa_{ija}}{\partial \delta_b} = \frac{1}{8\gamma^2} \{ 5\delta_{(0,1,0,0,2)} + 3\delta_{(1,0,0,0,1)} + \delta_{(0,0,1,0,3)} \} \sum_{\ell=1}^n \left\{ \frac{q_{\ell ab}}{q_\ell} - \frac{q_{\ell a} q_{\ell b}}{q_\ell^2} \right\} e \\
(\kappa_{ij})_\gamma &= \frac{\partial \kappa_{ij}}{\partial \delta_\gamma} = -\frac{1}{\gamma^2} \delta_{(0,1,0,0,0)} \sum_{\ell=1}^n q_\ell(i, j)_\ell.
\end{aligned}$$

Como definido na Subseção 2.3.1, o fator de correção de Bartlett para a estatística LR é definido por $1 + c/k$, com

$$c = \epsilon_{p,k} - \epsilon_p.$$

Das expansões de Lawley (Lawley, 1956), temos que $E(LR) = k + \epsilon_{p,k} - \epsilon_p + O(n^{-2})$, sendo

$$\epsilon_{p,k} = \sum_{\boldsymbol{\theta}^*} (l_{rstu} - l_{rstuvw}), \quad (\text{B.2})$$

$$\begin{aligned}
l_{rstu} &= \kappa^{rs} \kappa^{tu} \left\{ \frac{\kappa_{rstu}}{4} - (\kappa_{rst})_u + (\kappa_{rt})_{su} \right\}, \\
l_{rstuvw} &= \kappa^{rs} \kappa^{tu} \kappa^{vw} \left\{ \kappa_{rtv} \left(\frac{\kappa_{suw}}{6} - (\kappa_{sw})_u \right) \right. \\
&\quad \left. + \kappa_{rtu} \left(\frac{\kappa_{svw}}{4} - (\kappa_{sw})_v + (\kappa_{rt})_v (\kappa_{sw})_u + (\kappa_{rt})_u (\kappa_{sw})_v \right) \right\},
\end{aligned}$$

em que $-\kappa^{rs}$ é o elemento (r, s) da inversa da matriz de informação de Fisher e a soma em (B.2) varia sob todos os parâmetros de $\boldsymbol{\theta}^*$. A expressão ϵ_p é obtida de (B.2) considerando a soma sob todos os parâmetros de perturbação. Dessa forma, podemos escrever c como

$$c = \epsilon_k(\boldsymbol{\delta}) + \epsilon_{p,k}(\boldsymbol{\beta}, \boldsymbol{\delta}) + \epsilon_{p,k}(\boldsymbol{\delta}, \gamma) + \epsilon_{p,q}(\boldsymbol{\beta}, \boldsymbol{\delta}, \gamma),$$

assim como em (2.7), com $\epsilon_k(\boldsymbol{\delta})$ indicando que a soma em (B.2) é realizada sob todos os componentes de $\boldsymbol{\delta}$, analogamente para os demais casos, considerando os respectivos parâmetros. Considerando o caso de heteroscedasticidade com efeitos multiplicativos, isto é, $m_\ell = \exp\{\boldsymbol{\omega}_\ell^\top \boldsymbol{\delta}\}$, e substituindo os κ 's obtidos anteriormente nas expressões dos ϵ 's, após uma extensiva álgebra obtemos os componentes da constante c definida em (2.7) expressos por

$$\begin{aligned}
\epsilon_k(\boldsymbol{\delta}) &= \frac{M_4}{4} \text{tr}(\mathbf{H}_d^2) + \frac{M_3^2}{6} \boldsymbol{\iota}^\top \mathbf{H}^{(3)} \boldsymbol{\iota} + \frac{M_3^2}{4} \boldsymbol{\iota}^\top \mathbf{H} \mathbf{H}_d \mathbf{H} \boldsymbol{\iota}, \\
\epsilon_{p,k}(\boldsymbol{\beta}, \boldsymbol{\delta}) &= -\frac{M_7}{4\delta_{(0,1,0,0,0)}} \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} - \frac{M_8}{4\delta_{(0,1,0,0,0)}} \boldsymbol{\iota} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} \\
&+ \frac{\delta_{(0,0,1,0,1)}}{2\delta_{(0,1,0,0,0)}} \boldsymbol{\iota} \mathbf{Q} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} + \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} \\
&- \frac{M_7}{4\delta_{(0,1,0,0,0)}} \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} - \frac{M_8}{4\delta_{(0,1,0,0,0)}} \boldsymbol{\iota} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota} \\
&+ \left(\frac{M_{10}^2}{2(\delta_{(0,1,0,0,0)})^2} + \frac{M_{10}}{\delta_{(0,1,0,0,0)}} \right) \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{Z}_{\beta} \odot \mathbf{H} \odot \mathbf{Z}_{\beta} \mathbf{Q} \boldsymbol{\iota} \\
&+ \frac{M_{10}^2}{4(\delta_{(0,1,0,0,0)})^2} \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{Z}_{\beta d} \mathbf{H} \mathbf{Z}_{\beta d} \mathbf{Q} \boldsymbol{\iota} - \frac{M_3 M_{10}}{2\delta_{(0,1,0,0,0)}} \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{Z}_{\beta d} \mathbf{H} \mathbf{H}_d \boldsymbol{\iota}, \\
\epsilon_{p,k}(\boldsymbol{\delta}, \gamma) &= -\frac{M_9 M_{11}}{2} \text{tr}(\mathbf{H}_d) + M_2 M_{11} \text{tr}(\mathbf{H}_d) - \frac{M_1^2 M_{11}}{2} \boldsymbol{\iota}^\top \mathbf{H}^{(2)} \boldsymbol{\iota} \\
&- \frac{M_1^2 M_{11}}{4} [\text{tr}(\mathbf{H}_d)]^2 + \left(\frac{M_1 M_6}{4} + \frac{M_1 (M_6 - 4M_5)}{4} \right) M_{11}^2 \text{tr}(\mathbf{H}_d) \text{ e} \\
\epsilon_{p,k}(\boldsymbol{\beta}, \boldsymbol{\delta}, \gamma) &= \frac{M_1 M_{10} M_{11}}{4\delta_{(0,1,0,0,0)}} [(\boldsymbol{\iota}^\top \mathbf{W} \mathbf{Z}_{\beta d} \boldsymbol{\iota}) \odot (\boldsymbol{\iota}^\top \mathbf{H}_d \boldsymbol{\iota}) + \boldsymbol{\iota}^\top \mathbf{Q} \mathbf{H}_d \mathbf{Z}_{\beta d} \boldsymbol{\iota}].
\end{aligned}$$

Cálculo do fator de correção c_m para a estatística LR_m

Neste apêndice obtemos com detalhes a constante c_m do fator de correção para a estatística LR_m para a classe dos MNLSH definido em (2.8). No que segue adotaremos as seguintes notações: os índices a, b, c, \dots variam em $\boldsymbol{\delta}$, os índices i, j, l, \dots variam em $\boldsymbol{\beta}$ e os índices r, s, t, \dots variam em $\boldsymbol{\theta}^* = (\boldsymbol{\delta}^\top, \boldsymbol{\beta}^\top, \gamma)^\top$. Para os valores esperados das derivadas do logaritmo da função de verossimilhança: $\kappa_{rs} = E(\partial^2 l^* / \partial \theta_r^* \partial \theta_s^*)$, $\kappa_{rst} = E(\partial^3 l^* / \partial \theta_r^* \partial \theta_s^* \partial \theta_t^*)$, etc. Denotamos as derivadas dos valores esperados das derivadas do logaritmo da função de verossimilhança por $(\kappa_{rs})_t = \partial \kappa_{rs} / \partial \theta_t^*$, $(\kappa_{rs})_{tu} = \partial^2 \kappa_{rs} / \partial \theta_t^* \partial \theta_u^*$, etc. A matriz de informação de Fisher tem como elementos $-\kappa_{rs}$, com $-\kappa^{rs}$ os respectivos elementos da sua inversa. Ainda, definimos $\tau^{rs} = \kappa^{ra} \kappa^{sb} \sigma_{ab}$, sendo σ_{ab} a inversa da matriz κ^{ab} e $\nu^{rs} = \kappa^{rs} - \tau^{rs}$.

A equação para o cálculo de c_m obtida em Ferrari et al. (2004, Eq. 5) é expressa por

$$\begin{aligned}
 c_m &= \frac{1}{4} \kappa^{ab} \kappa^{cd} \kappa_{abcd} - \kappa^{ab} \kappa^{cd} (\kappa_{acd})_b + \kappa^{ab} \kappa^{cd} (\kappa_{ac})_{db} - \kappa^{ij} \kappa^{ab} (\kappa_{iab})_j \\
 &- \kappa^{\gamma\gamma} \kappa^{ab} (\kappa_{ab\gamma})_\gamma - \left(\frac{1}{4} \kappa^{ab} \kappa^{cd} \kappa^{ef} + \frac{1}{2} \kappa^{ab} \kappa^{cf} \kappa^{de} - \frac{1}{3} \kappa^{ab} \kappa^{cf} \kappa^{de} \right) \kappa_{acd} \kappa_{bef} \\
 &+ (\kappa^{ab} \kappa^{cd} \kappa^{ef} + \kappa^{ab} \kappa^{cf} \kappa^{de}) \kappa_{acd} (\kappa_{be})_f - (\kappa^{ab} \kappa^{cd} \kappa^{ef} + \kappa^{ab} \kappa^{cf} \kappa^{de}) (\kappa_{ac})_d (\kappa^{be})_f \\
 &- \left(\frac{1}{4} \kappa^{ij} \kappa^{ab} \kappa^{cd} + \frac{1}{2} \kappa^{ij} \kappa^{ad} \kappa^{bc} \right) \kappa_{iab} \kappa_{jcd} + (\kappa^{ij} \kappa^{ab} \kappa^{kl}) \kappa_{iab} (\kappa_{jk})_l \\
 &- \left(\frac{1}{4} \kappa^{\gamma\gamma} \kappa^{ab} \kappa^{cd} + \frac{1}{2} \kappa^{\gamma\gamma} \kappa^{ad} \kappa^{bc} \right) \kappa_{ab\gamma} \kappa_{cd\gamma} + (\kappa^{\gamma\gamma} \kappa^{ab} \kappa^{\gamma\gamma}) \kappa_{ab\gamma} (\kappa_{\gamma\gamma})_\gamma. \tag{C.1}
 \end{aligned}$$

Considerando o caso de heteroscedasticidade com efeitos multiplicativos, isto é, o

caso em que $m_\ell = \exp\{\boldsymbol{\omega}_\ell^\top \boldsymbol{\delta}\}$, temos $q_\ell = \frac{(\prod_{s=1}^n m_s)^{1/n}}{m_\ell} = \exp\{-(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})^\top \boldsymbol{\delta}\}$, com $\bar{\boldsymbol{\omega}} = (\bar{\omega}_1, \dots, \bar{\omega}_k)^\top$. Do Apêndice B, sabemos que $q_{\ell a} = \frac{\partial q_\ell}{\partial \delta_a}$, $q_{\ell ab} = \frac{\partial q_{\ell a}}{\partial \delta_b}$, *etc*, portanto, temos $q_{\ell a} = -(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a q_\ell$, $q_{\ell ab} = (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{ab} q_\ell$, $q_{\ell abc} = -(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{abc} q_\ell$ e $q_{\ell abcd} = (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{abcd} q_\ell$, com $\bar{\omega}_a = \frac{1}{n} \sum_{\ell=1}^n \omega_{\ell a}$ para $a = 1, \dots, k$. Ainda, seja $(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a = \omega_{\ell a} - \bar{\omega}_a$, $(\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{ab} = (\omega_{\ell a} - \bar{\omega}_a)(\omega_{\ell b} - \bar{\omega}_b)$, $\ell = 1, \dots, n$, e assim sucessivamente. Dessa maneira, tomando a esperança de algumas derivadas do logaritmo da função de verossimilhança obtidas no Apêndice B temos os κ 's necessários ao cálculo de (C.1) expressos por

$$\begin{aligned}\kappa_{ab\gamma} &= \frac{\Delta_1}{\gamma} \sum_{\ell=1}^n (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{ab}, \\ \kappa_{acd} &= \Delta_3 \sum_{\ell=1}^n (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{acd}, \\ \kappa_{abcd} &= \Delta_4 \sum_{\ell=1}^n (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{abcd}, \\ (\kappa_{\gamma\gamma})^{-1} &= \kappa^{\gamma\gamma} = \Delta_6 \gamma^2.\end{aligned}$$

Diferenciando algumas das esperanças acima, temos:

$$\begin{aligned}(\kappa_{\gamma\gamma})_\gamma &= \frac{\Delta_5}{\gamma^3}, \\ (\kappa_{ab\gamma})_\gamma &= \frac{\Delta_2}{\gamma^2} \sum_{\ell=1}^n (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_{ab}, \\ (\kappa_{ac})_d &= (\kappa_{ac})_{bd} = (\kappa_{acd})_b = \kappa_{iab} = (\kappa_{iab})_j = 0,\end{aligned}$$

Substituindo os κ 's na equação (C.1), temos que

$$\begin{aligned}c_m &= \frac{1}{4} \Delta_4 \sum_{\ell=1}^n (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a \kappa^{ab} (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_b (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_c \kappa^{cd} (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_d \\ &- \frac{1}{4} \Delta_6 \Delta_1^2 \sum_{\ell=1}^n (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a \kappa^{ab} (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_b (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_c \kappa^{cd} (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_d \\ &- \frac{1}{4} \Delta_3^2 \sum_{\ell=1}^n \sum_{s=1}^n (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a \kappa^{ab} (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_a (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_c \kappa^{cd} (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_d (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_e \kappa^{ef} (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_f \\ &- \frac{1}{6} \Delta_3^2 \sum_{\ell=1}^n \sum_{s=1}^n (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a \kappa^{ab} (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_b (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_c \kappa^{ef} (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_f (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_d \kappa^{de} (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_e \\ &- \Delta_2 \Delta_6 \sum_{s=1}^n (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_a \kappa^{ab} (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_b + \Delta_1 \Delta_5 \Delta_6^2 \sum_{\ell=1}^n (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_a \kappa^{ab} (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_b \\ &- \frac{1}{2} \Delta_1^2 \Delta_6 \sum_{\ell=1}^n \sum_{s=1}^n (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_a \kappa^{ad} (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_d (\boldsymbol{\omega}_\ell - \bar{\boldsymbol{\omega}})_b \kappa^{bc} (\boldsymbol{\omega}_s - \bar{\boldsymbol{\omega}})_c.\end{aligned}$$

Seja $H = \{h_{\ell s}\} = -(\mathbf{W} - \bar{\mathbf{W}})[(\mathbf{W} - \bar{\mathbf{W}})^\top V(\mathbf{W} - \bar{\mathbf{W}})]^{-1}(\mathbf{W} - \bar{\mathbf{W}})^\top$, com $(\mathbf{W} - \bar{\mathbf{W}}) = (\mathbf{w}_1 - \bar{\mathbf{w}}, \dots, \mathbf{w}_n - \bar{\mathbf{w}})^\top$ e V matriz diagonal de ordem n com entradas $v_\ell = (1 - \delta_{(0,1,0,0,2)})/4$, $\ell, s = 1, \dots, n$. Assim, temos que $h_{\ell s} = -(\mathbf{w}_\ell - \bar{\mathbf{w}})_a \kappa^{ab} (\mathbf{w}_s - \bar{\mathbf{w}})$. Dessa forma, c_m pode ser reescrito como

$$\begin{aligned} c_m &= \frac{1}{4} \Delta_4 \sum_{\ell=1}^n h_{\ell\ell}^2 - \frac{1}{4} \Delta_1^2 \Delta_6 \left(\sum_{\ell=1}^n h_{\ell\ell} \right)^2 + \frac{1}{4} \Delta_2^3 \sum_{\ell=1}^n \sum_{s=1}^n h_{\ell\ell} h_{\ell s} h_{s s} \\ &+ \frac{1}{6} \Delta_3^2 \sum_{\ell=1}^n \sum_{s=1}^n h_{\ell s}^3 + \Delta_2 \Delta_6 \sum_{s=1}^n h_{s s} - \Delta_1 \Delta_5 \Delta_6^2 v \sum_{\ell=1}^n h_{\ell\ell} \\ &- \frac{1}{2} \Delta_1^2 \Delta_6 \sum_{\ell=1}^n \sum_{s=1}^n h_{\ell s}^2, \end{aligned}$$

em notação matricial, temos

$$\begin{aligned} c_m &= \frac{1}{4} \Delta_4 \text{tr}(\mathbf{H}_d^2) - \frac{1}{4} \Delta_1^2 \Delta_6 [\text{tr}(\mathbf{H}_d)]^2 + \frac{1}{4} \Delta_3^2 \boldsymbol{\iota}^\top \mathbf{H}_d \mathbf{H} \mathbf{H}_d \boldsymbol{\iota} + \frac{1}{6} \Delta_3^2 \boldsymbol{\iota}^\top \mathbf{H}^{(3)} \boldsymbol{\iota} \\ &+ \Delta_2 \Delta_6 \text{tr}(\mathbf{H}_d) - \Delta_1 \Delta_5 \Delta_6^2 \text{tr}(\mathbf{H}_d) - \frac{1}{2} \Delta_1^2 \Delta_6 \boldsymbol{\iota}^\top \mathbf{H}^{(2)} \boldsymbol{\iota}, \end{aligned}$$

sendo $H_d = \text{diag}\{h_{11}, \dots, h_{nn}\}$, $H^{(2)} = (h_{\ell s}^2)$, $H^{(3)} = (h_{\ell s}^3)$ e os Δ 's são escalares definidos na Subseção 2.3.2.

APÊNDICE D

Derivadas do logaritmo da função de verossimilhança de $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\delta})^\top$

Neste apêndice obtemos algumas derivadas, e momentos das derivadas, do logaritmo da função de verossimilhança de $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\delta})^\top$ do modelo definido no Capítulo 3 expresso por

$$l(\boldsymbol{\theta}) = l(\boldsymbol{\theta}; \mathbf{y}) = -\frac{1}{2} \sum_{\ell=1}^n \log(\phi_\ell^*) + \sum_{\ell=1}^n t(z_\ell) \quad (\text{D.1})$$

para a obtenção das quantidades A_1, A_2 e A_3 do fator de correção tipo-Bartlett para a estatística S_r e quantidades A_1^g, A_2^g e A_3^g do fator de correção tipo-Bartlett para a estatística S_g definidas nas Subseções 3.3.1 e 3.3.2, respectivamente. Para isto, consideremos algumas notações: os índices a, b, c, \dots variam em $\boldsymbol{\delta}$, os índices i, j, l, \dots variam em $\boldsymbol{\beta}$, as derivadas de μ_ℓ com relação aos componentes de $\boldsymbol{\beta}$ são denotados por $(i)_\ell = \partial\mu_\ell/\partial\beta_i$, $(ij)_\ell = \partial^2\mu_\ell/\partial\beta_i\partial\beta_j$, $(ij, l)_\ell = (\partial^2\mu_\ell/\partial\beta_i\partial\beta_j)(\partial\mu_\ell/\partial\beta_l)$, etc, as derivadas de τ_ℓ^* com respeito aos componentes de $\boldsymbol{\delta}$ são denotadas por $(a)_\ell = \partial\tau_\ell^*/\partial\delta_a$, $(ab)_\ell = \partial^2\tau_\ell^*/\partial\delta_a\partial\delta_b$, $(ab, c)_\ell = (\partial^2\tau_\ell^*/\partial\delta_a\partial\delta_b)(\partial\tau_\ell^*/\partial\delta_c)$, etc, $h_\ell'' = \partial^2\phi_\ell^*/\partial\tau_\ell^{*2}$, $h_\ell''' = \partial^3\phi_\ell^*/\partial\tau_\ell^{*3}$ e $h_\ell'''' = \partial^4\phi_\ell^*/\partial\tau_\ell^{*4}$.

As quatro primeiras derivadas de $l(\boldsymbol{\theta})$ com respeito aos componentes de $\boldsymbol{\beta}$ são dadas

por

$$\begin{aligned}
\frac{\partial l(\boldsymbol{\theta})}{\partial \beta_i} &= -\sum_{\ell=1}^n \frac{t(z_\ell)^{(1)}}{\sqrt{\phi_\ell^*}}(i)_\ell, \\
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j} &= \sum_{\ell=1}^n t(z_\ell)^{(2)} \frac{1}{\phi_\ell^*}(i, j)_\ell - \sum_{\ell=1}^n t(z_\ell)^{(1)} \frac{1}{\sqrt{\phi_\ell^*}}(ij)_\ell, \\
\frac{\partial^3 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \beta_l} &= -\sum_{\ell=1}^n t(z_\ell)^{(3)} \frac{1}{\phi_\ell^{3/2}}(i, j, l)_\ell + \sum_{\ell=1}^n t(z_\ell)^{(2)} \frac{1}{\phi_\ell^*} \{(ij, l)_\ell + (i, jl)_\ell + (il, j)_\ell\} \\
&\quad - \sum_{\ell=1}^n t(z_\ell)^{(1)} \frac{1}{\sqrt{\phi_\ell^*}}(ijl)_\ell \text{ e} \\
\frac{\partial^4 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \beta_l \partial \beta_m} &= \sum_{\ell=1}^n t(z_\ell)^{(4)} \frac{1}{\phi_\ell^{*2}}(i, j, l, m)_\ell - \sum_{\ell=1}^n t(z_\ell)^{(3)} \frac{1}{\phi_\ell^{*3/2}} \{(im, j, l)_\ell + (i, jm, l)_\ell \\
&\quad + (i, j, lm)_\ell\} - \sum_{\ell=1}^n t(z_\ell)^{(3)} \frac{1}{\phi_\ell^{*3/2}} \{(il, j, m)_\ell + (i, jl, m)_\ell + (ij, l, m)_\ell\} \\
&\quad + \sum_{\ell=1}^n t(z_\ell)^{(2)} \frac{1}{\phi_\ell^*} \{(ilm, j)_\ell + (il, jm)_\ell + (im, jl)_\ell + (i, jlm)_\ell + (ijm, l)_\ell \\
&\quad + (ij, lm)_\ell\} + \sum_{\ell=1}^n t(z_\ell)^{(2)} \frac{1}{\phi_\ell^*}(ijl, m)_\ell - \sum_{\ell=1}^n t(z_\ell)^{(1)} \frac{1}{\sqrt{\phi_\ell^*}}(ijlm)_\ell.
\end{aligned}$$

As quatro primeiras derivadas de $l(\boldsymbol{\theta})$ com respeito aos componentes de $\boldsymbol{\delta}$ são dadas

por

$$\begin{aligned}
\frac{\partial l(\boldsymbol{\theta})}{\partial \delta_a} &= -\frac{1}{2} \sum_{\ell=1}^n \frac{h'_\ell}{\phi_\ell^*}(a)_\ell - \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{h'_\ell}{\phi_\ell^*}(a)_\ell, \\
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b} &= -\frac{1}{2} \sum_{\ell=1}^n \frac{h''_\ell}{\phi_\ell^*}(a, b)_\ell + \frac{1}{2} \sum_{\ell=1}^n \frac{h_\ell'^2}{\phi_\ell^{*2}}(a, b)_\ell + \frac{1}{4} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 \frac{h_\ell'^2}{\phi_\ell^{*2}}(a, b)_\ell \\
&\quad + \frac{1}{4} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{h_\ell'^2}{\phi_\ell^{*2}}(a, b)_\ell - \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{h_\ell''}{\phi_\ell^*}(a, b)_\ell + \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{h_\ell'^2}{\phi_\ell^{*2}}(a, b)_\ell, \\
\frac{\partial^3 l(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \delta_c} &= -\frac{1}{2} \sum_{\ell=1}^n \frac{h_\ell'''}{\phi_\ell^*}(a, b, c)_\ell + \frac{3}{2} \sum_{\ell=1}^n \frac{h_\ell' h_\ell''}{\phi_\ell^{*2}}(a, b, c)_\ell - \sum_{\ell=1}^n \frac{h_\ell'^3}{\phi_\ell^{*3}}(a, b, c)_\ell \\
&\quad - \frac{1}{8} \sum_{\ell=1}^n t(z_\ell)^{(3)} z_\ell^3 \frac{h_\ell'^3}{\phi_\ell^{*3}}(a, b, c)_\ell - \frac{9}{8} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 \frac{h_\ell'^3}{\phi_\ell^{*3}} \\
&\quad + \frac{3}{4} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 \frac{h_\ell' h_\ell''}{\phi_\ell^{*2}}(a, b, c)_\ell - \frac{15}{8} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{h_\ell'^3}{\phi_\ell^{*3}}(a, b, c)_\ell \\
&\quad + \frac{9}{4} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{h_\ell' h_\ell''}{\phi_\ell^{*2}}(a, b, c)_\ell - \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(1)} z_\ell \frac{h_\ell'''}{\phi_\ell^*}(a, b, c)_\ell \text{ e}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^4 l(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \delta_c \partial \delta_d} &= -\frac{1}{2} \sum_{\ell=1}^n \frac{h_\ell''''}{\phi_\ell^*} (a, b, c, d)_\ell + 2 \sum_{\ell=1}^n \frac{h'_\ell h_\ell''''}{\phi_\ell^{*2}} (a, b, c, d)_\ell + \frac{3}{2} \sum_{\ell=1}^n \frac{h_\ell''^2}{\phi_\ell^{*2}} (a, b, c, d)_\ell \\
&- 6 \sum_{\ell=1}^n \frac{h_\ell'^2 h_\ell''}{\phi_\ell^{*3}} (a, b, c, d)_\ell + 3 \sum_{\ell=1}^n \frac{h_\ell'^4}{\phi_\ell^{*4}} (a, b, c, d)_\ell \\
&+ \frac{9}{8} \sum_{\ell=1}^n t(z_\ell)^{(3)} z_\ell^3 \frac{h_\ell'^4}{\phi_\ell^{*4}} (a, b, c, d)_\ell - \frac{3}{4} \sum_{\ell=1}^n t(z_\ell)^{(3)} z_\ell^3 \frac{h_\ell'^2 h_\ell''}{\phi_\ell^{*3}} (a, b, c, d)_\ell \\
&+ \frac{87}{16} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 \frac{h_\ell'^4}{\phi_\ell^{*4}} (a, b, c, d)_\ell - \frac{27}{4} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 \frac{h_\ell'^2 h_\ell''}{\phi_\ell^{*3}} (a, b, c, d)_\ell \\
&+ \frac{3}{4} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 \frac{h_\ell''^2}{\phi_\ell^{*2}} (a, b, c, d)_\ell + \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell^2 \frac{h'_\ell h_\ell''}{\phi_\ell^{*2}} (a, b, c, d)_\ell \\
&+ \frac{1}{16} \sum_{\ell=1}^n t(z_\ell)^{(4)} z_\ell^4 \frac{h_\ell'^4}{\phi_\ell^{*4}} (a, b, c, d)_\ell.
\end{aligned}$$

Algumas das derivadas mistas de (D.1) são dadas por

$$\begin{aligned}
\frac{\partial^3 l(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \beta_i} &= -\frac{1}{4} \sum_{\ell=1}^n t(z_\ell)^{(3)} z_\ell^2 \frac{h_\ell'^2}{\phi_\ell^{*3/2}} (a, b, i)_\ell \\
&- \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell \frac{h_\ell'^2}{\phi_\ell^{*5/2}} (a, b, i)_\ell - \frac{1}{4} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell \frac{h_\ell'^2}{\phi_\ell^{*5/2}} (a, b, i)_\ell \\
&- \frac{1}{4} \sum_{\ell=1}^n t(z_\ell)^{(1)} \frac{h_\ell'^2}{\phi_\ell^{*5/2}} (a, b, i)_\ell + \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell \frac{h_\ell''}{\phi_\ell^{*5/2}} (a, b, i)_\ell \\
&+ \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(1)} \frac{h_\ell''}{\phi_\ell^{*5/2}} (a, b, i)_\ell - \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell \frac{h_\ell'^2}{\phi_\ell^{*5/2}} (a, b, i)_\ell \\
&- \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(1)} \frac{h_\ell'^2}{\phi_\ell^{*5/2}} (a, b, i)_\ell, \\
\frac{\partial^3 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \delta_a} &= -\frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(3)} z_\ell \frac{h'_\ell}{\phi_\ell^{*2}} (i, j, a)_\ell - \sum_{\ell=1}^n t(z_\ell)^{(2)} \frac{h'_\ell}{\phi_\ell^{*2}} (i, j, a)_\ell \\
&+ \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell \frac{h'_\ell}{\phi_\ell^{*3/2}} (ij, a)_\ell + \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(1)} \frac{h'_\ell}{\phi_\ell^{*3/2}} (ij, a)_\ell \text{ e} \\
\frac{\partial^4 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \delta_a \partial \delta_b} &= \frac{1}{4} \sum_{\ell=1}^n t(z_\ell)^{(4)} z_\ell^2 \frac{h_\ell'^2}{\phi_\ell^{*3}} (i, j, a, b)_\ell \\
&+ \frac{7}{4} \sum_{\ell=1}^n t(z_\ell)^{(3)} z_\ell \frac{h_\ell'^2}{\phi_\ell^{*3}} (i, j, a, b)_\ell - \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(3)} z_\ell \frac{h_\ell''}{\phi_\ell^{*2}} (i, j, a, b)_\ell
\end{aligned}$$

$$\begin{aligned}
& - \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(2)} \frac{h_\ell''}{\phi_\ell^{*2}}(i, j, a, b)_\ell + 2 \sum_{\ell=1}^n t(z_\ell)^{(2)} \frac{h_\ell'^2}{\phi_\ell^{*3}}(i, j, a, b)_\ell \\
& - \frac{1}{4} \sum_{\ell=1}^n t(z_\ell)^{(3)} z_\ell^2 \frac{h_\ell'^2}{\phi_\ell^{*5/2}}(ij, a, b)_\ell - \frac{5}{4} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell \frac{h_\ell'^2}{\phi_\ell^{*5/2}}(ij, a, b)_\ell \\
& + \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(2)} z_\ell \frac{h_\ell''}{\phi_\ell^{*3/2}}(ij, a, b)_\ell + \frac{1}{2} \sum_{\ell=1}^n t(z_\ell)^{(1)} \frac{h_\ell''}{\phi_\ell^{*3/2}}(ij, a, b)_\ell \\
& - \frac{3}{4} \sum_{\ell=1}^n t(z_\ell)^{(1)} \frac{h_\ell'^2}{\phi_\ell^{*5/2}}(ij, a, b)_\ell.
\end{aligned}$$

Tomando a esperança das derivadas anteriores, temos:

$$\begin{aligned}
\kappa_{ij} &= E \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j} \right] = \delta_{(0,1,0,0,0)} \sum_{\ell=1}^n \frac{1}{\phi_\ell^*} (i, j)_\ell, \\
\kappa_{ijl} &= E \left[\frac{\partial^3 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \beta_l} \right] = \delta_{(0,1,0,0,0)} \sum_{\ell=1}^n \frac{1}{\phi_\ell^*} \{ (il, j)_\ell + (i, jl)_\ell + (ij, l)_\ell \}, \\
\kappa_{ijklm} &= E \left[\frac{\partial^4 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \beta_l \partial \beta_m} \right] = \delta_{(0,0,0,1,0)} \sum_{\ell=1}^n \frac{1}{\phi_\ell^{*2}} (i, j, l, m)_\ell \{ (il, j)_\ell + (i, lm)_\ell + (ij, l)_\ell \}, \\
\kappa_{ab} &= E \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b} \right] = \frac{\delta_{(0,1,0,0,2)} - 1}{4} \sum_{\ell=1}^n \frac{h_\ell'^2}{\phi_\ell^{*2}}(a, b)_\ell, \\
\kappa_{abc} &= E \left[\frac{\partial^3 l(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \delta_c} \right] = \frac{3(\delta_{(0,1,0,0,2)} - 1)}{4} \sum_{\ell=1}^n \frac{h_\ell' h_\ell''}{\phi_\ell^{*2}}(a, b, c)_\ell, \\
& + \frac{(7 - \delta_{(0,0,1,0,3)} - 9\delta_{(0,1,0,0,2)})}{8} \sum_{\ell=1}^n \frac{h_\ell'^3}{\phi_\ell^{*3}}(a, b, c)_\ell,
\end{aligned}$$

$$\begin{aligned}
\kappa_{abcd} &= E \left[\frac{\partial^4 l(\boldsymbol{\theta})}{\partial \delta_a \partial \delta_b \partial \delta_c \partial \delta_d} \right] = - \sum_{\ell=1}^n \frac{h'_\ell h''_\ell}{\phi_\ell^{*2}}(a, b, c, d)_\ell - \frac{3}{4} \sum_{\ell=1}^n \frac{h''^2_\ell}{\phi_\ell^{*2}}(a, b, c, d)_\ell, \\
&+ \frac{21}{4} \sum_{\ell=1}^n \frac{h'^2_\ell h''_\ell}{\phi_\ell^{*3}}(a, b, c, d)_\ell - \frac{57}{16} \sum_{\ell=1}^n \frac{h'^4_\ell}{\phi_\ell^{*4}}(a, b, c, d)_\ell, \\
&+ \frac{1}{16} \delta_{(0,0,0,4,4)} \sum_{\ell=1}^n \frac{h'^4_\ell}{\phi_\ell^{*4}}(a, b, c, d)_\ell + \frac{9}{8} \delta_{(0,0,1,0,3)} \sum_{\ell=1}^n \frac{h'^4_\ell}{\phi_\ell^{*4}}(a, b, c, d)_\ell, \\
&- \frac{3}{4} \delta_{(0,0,1,0,3)} \sum_{\ell=1}^n \frac{h'^2_\ell h''_\ell}{\phi_\ell^{*3}}(a, b, c, d)_\ell + \frac{87}{16} \delta_{(0,1,0,0,2)} \sum_{\ell=1}^n \frac{h'^4_\ell}{\phi_\ell^{*4}}(a, b, c, d)_\ell, \\
&- \frac{27}{4} \delta_{(0,1,0,0,2)} \sum_{\ell=1}^n \frac{h'^2_\ell h''_\ell}{\phi_\ell^{*3}}(a, b, c, d)_\ell + \frac{3}{4} \delta_{(0,1,0,0,2)} \sum_{\ell=1}^n \frac{h''^2_\ell}{\phi_\ell^{*2}}(a, b, c, d)_\ell, \\
&+ \delta_{(0,1,0,0,2)} \sum_{\ell=1}^n \frac{h'_\ell h''_\ell}{\phi_\ell^{*2}}(a, b, c, d)_\ell, \\
\kappa_{ija} &= E \left[\frac{\partial^3 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \delta_a} \right] = -\frac{1}{2} \delta_{(0,0,1,0,1)} \sum_{\ell=1}^n \frac{h'_\ell}{\phi_\ell^{*2}}(i, j, a)_\ell - \delta_{(0,1,0,0,0)} \sum_{\ell=1}^n \frac{h'_\ell}{\phi_\ell^{*2}}(i, j, a)_\ell \text{ e} \\
\kappa_{ijab} &= E \left[\frac{\partial^4 l(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j \partial \delta_a \partial \delta_b} \right] = \frac{1}{4} \delta_{(0,0,0,1,2)} \sum_{\ell=1}^n \frac{h'^2_\ell}{\phi_\ell^{*3}}(i, j, a, b)_\ell + \frac{7}{4} \delta_{(0,0,1,0,1)} \sum_{\ell=1}^n \frac{h'^2_\ell}{\phi_\ell^{*3}}(i, j, a, b)_\ell \\
&- \frac{1}{2} \delta_{(0,0,1,0,1)} \sum_{\ell=1}^n \frac{h''_\ell}{\phi_\ell^{*2}}(i, j, a, b)_\ell - \frac{1}{2} \delta_{(0,1,0,0,0)} \sum_{\ell=1}^n \frac{h''_\ell}{\phi_\ell^{*2}}(i, j, a, b)_\ell \\
&+ 2 \delta_{(0,1,0,0,0)} \sum_{\ell=1}^n \frac{h'^2_\ell}{\phi_\ell^{*3}}(i, j, a, b)_\ell \text{ e} \\
\kappa_{ai} &= \kappa_{abi} = 0.
\end{aligned}$$

Derivando algumas das esperanças acima, temos

$$\begin{aligned}
(\kappa_{ij})_l &= \frac{\partial \kappa_{ij}}{\partial \beta_l} = \delta_{(0,1,0,0,0)} \sum_{\ell=1} \frac{1}{\phi_\ell} \{(ij, l)_\ell + (i, jl)_\ell\}, \\
(\kappa_{ab})_c &= \frac{\partial \kappa_{ab}}{\partial \delta_c} = \left(\frac{\delta_{(0,1,0,0,2)} - 1}{2} \right) \sum_{\ell=1} \left\{ \frac{h'_\ell h''_\ell}{\phi_\ell^{*2}} - \frac{h_\ell^3}{\phi_\ell^{*3}} \right\} (b, c)_\ell, \\
(\kappa_{ab})_{cd} &= \frac{\partial^2 \kappa_{ab}}{\partial \delta_c \partial \delta_d} = \left(\frac{\delta_{(0,1,0,0,2)} - 1}{2} \right) \sum_{\ell=1} \left\{ \frac{h''_\ell^2}{\phi_\ell^{*2}} + \frac{h'_\ell h''_\ell}{\phi_\ell^{*2}} - \frac{5h_\ell^2 h''_\ell}{\phi_\ell^{*3}} + \frac{3h_\ell^4}{\phi_\ell^{*4}} \right\} (a, b, c, d)_\ell, \\
(\kappa_{abc})_d &= \frac{\partial \kappa_{abc}}{\partial \delta_d} = \frac{3(\delta_{(0,1,0,0,2)} - 1)}{4} \sum_{\ell=1} \left\{ \frac{h''_\ell^2}{\phi_\ell^{*2}} + \frac{h'_\ell h''_\ell}{\phi_\ell^{*2}} - \frac{2h_\ell^2 h''_\ell}{\phi_\ell^{*3}} \right\} (a, b, c, d)_\ell, \\
&+ \frac{(7 - \delta_{(0,0,1,0,3)} - 9\delta_{(0,1,0,0,2)})}{8} \sum_{\ell=1} \left\{ \frac{3h_\ell^2 h''_\ell}{\phi_\ell^{*3}} - \frac{3h_\ell^4}{\phi_\ell^{*4}} \right\} (a, b, c, d)_\ell \quad e \\
(\kappa_{abi})_j &= (\kappa_{ab})_i = 0.
\end{aligned}$$

APÊNDICE E

Cálculo dos termos A_1^g , A_2^g e A_3^g do fator de correção tipo-Bartlett para a estatística S_g

Neste apêndice discutimos o cálculo dos termos A_1^g , A_2^g e A_3^g que compõem o fator de correção tipo-Bartlett para a estatística gradiente (S_g), definido na Subseção 3.3.2 por

$$\{1 - (c_g + b_g S_g + a_g S_g^2)\}, \quad (\text{E.1})$$

sendo

$$a_g = \frac{A_3^g}{12(k-1)((k-1)+2)((k-1)+4)}, \quad b_g = \frac{A_2^g - 2A_3^g}{12(k-1)((k-1)+21)} \text{ e}$$

$$c_g = \frac{A_1^g - A_2^g + A_3^g}{12(k-1)},$$

para a classe dos MNLSH. No que segue, estaremos considerando $m_\ell(\tau_\ell^*) = \exp(\tau_\ell^*)$, isto é, heteroscedasticidade com efeitos multiplicativos. A obtenção de A_1^g , A_2^g e A_3^g é feita substituindo em suas expressões, dadas na Subseção 3.3.2, os momentos do logaritmo da função de verossimilhança calculados no Apêndice D. Aqui, iremos obter com detalhes o termo A_3^g .

Como definido na Subseção 3.3.2, temos que

$$A_3^g = \frac{1}{4} \sum' \kappa_{jrs} \kappa_{klu} (3m^{jr} m^{sk} m^{lu} + 2m^{jk} m^{rl} m^{su}),$$

onde \sum' denota que o somatório foi realizado sob os $(p+k)$ parâmetros, isto é, os índices j, r, s, k, l, u variam sob todos os parâmetros do modelo. Substituindo os momentos calculados no Apêndice D em A_3^g e rearranjando os termos, temos que:

$$A_3^g = \frac{3}{4}Q_1^2 \sum_{i,\ell} \left(\sum' (a)_i m^{ab}(b)_i \right) \left(\sum' (c)_i m^{cd}(d)_\ell \right) \left(\sum' (e)_\ell m^{ef}(b)_\ell \right) \\ + \frac{1}{2}Q_1^2 \sum_{i,\ell} \left(\sum' (a)_i m^{ad}(d)_\ell \right) \left(\sum' (b)_i m^{be}(e)_\ell \right) \left(\sum' (c)_i m^{cf}(f)_\ell \right),$$

em que $Q_1 = \frac{1}{8}\{1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\}$. Em notação matricial, temos

$$A_3^g = \frac{3}{4}Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0}) (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \boldsymbol{\nu} + \frac{1}{2}Q_1^2 \boldsymbol{\nu}^\top (\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})^{(3)} \boldsymbol{\nu}.$$

Para os termos A_1^g e A_2^g , podemos reescrevê-los como

$$A_1^g = A_{11}^g + A_{12}^g + A_{13}^g + A_{14}^g + A_{15}^g + A_{16}^g \quad \text{e} \quad A_2^g = A_{21}^g + A_{22}^g + A_{23}^g + A_{24}^g,$$

em que

$$A_{11}^g = 3 \sum' \kappa_{trs} \kappa_{kvu} [m^{tr} a^{vu} (m^{sk} + 2a^{sk}) + a^{tr} m^{sk} a^{vu} + 2m^{tk} a^{rv} a^{su}], \\ A_{12}^g = -12 \sum' (\kappa_{tr})_s (\kappa_{kv})_u (\kappa^{st} \kappa^{rk} \kappa^{vu} + a^{st} a^{rk} a^{vu} + \kappa^{sk} \kappa^{vt} \kappa^{ru} + a^{sk} a^{vt} a^{ru}), \\ A_{13}^g = -6 \sum' \kappa_{trs} (\kappa_{kv})_u [(a^{su} - k^{su})(k^{tk} k^{vr} - a^{tk} a^{vr}) + m^{rt} (a^{sk} a^{vu} + k^{sk} k^{vu})], \\ A_{14}^g = -12 \sum' \kappa_{trs} (\kappa_{kv})_u [a^{rs} (\kappa^{tk} \kappa^{vu} - a^{tk} a^{vu}) + a^{rk} a^{vs} m^{tu}], \\ A_{15}^g = 6 \sum' \kappa_{trsu} m^{tr} a^{su} - 6 \sum' \kappa_{trs}^u [m^{tr} (a^{su} - \kappa^{su}) + 2m^{tu} a^{rs}], \\ A_{16}^g = 12 \sum' \kappa_{rs}^{tu} (\kappa^{tr} k^{su} - a^{tr} a^{su}), \\ A_{21}^g = -3 \sum' \kappa_{trs} \kappa_{kvu} [m^{tr} m^{sk} m^{vu} + m^{tr} m^{sk} m^{vu} + 2m^{tk} m^{rv} m^{su}], \\ A_{22}^g = -\frac{3}{4} \sum' \kappa_{trs} \kappa_{kvu} [(3m^{tr} m^{sk} m^{vu} + 2m^{tk} m^{rv} m^{su})], \\ A_{23}^g = 6 \sum' \kappa_{trs} (\kappa_{kv})_u [m^{su} (\kappa^{tk} \kappa^{vr} - a^{tk} a^{vr}) + m^{tr} (\kappa^{sk} \kappa^{vu} - a^{sk} a^{vu})] \text{ e} \\ A_{24}^g = 6 \sum' (\kappa_{trs})_u m^{tr} m^{su} - 3 \sum' \kappa_{trsu} m^{tr} m^{su},$$

e substituir os momentos calculados no Apêndice D em cada um dos A_{1i}^g 's e A_{2j}^g 's, com $i = 1, 2, 3, 4, 5, 6$ e $j = 1, 2, 3, 4$, e assim obter as expressões matriciais de A_1^g e A_2^g , apresentadas na Subseção 3.3.2.

Poder local: obtenção dos b_{il} 's e ξ e comparação de funções de poder

Neste apêndice iremos obter com detalhes os coeficientes b_{il} 's e ξ que determinam as expansões assintóticas das funções de distribuição das estatísticas da razão de verossimilhanças (S_1), Wald (S_2), escore (S_3) e gradiente (S_4) sob uma sequência de alternativas de Pitman. Ainda, iremos apresentar com detalhes a comparação das funções de poder dos testes baseados nas estatísticas supracitadas até ordem $n^{-1/2}$ sob uma sequência de alternativas de Pitman. Para isto, introduziremos as seguintes notações: os índices a, b, c, \dots variam em $\boldsymbol{\delta}$, os índices s, t, v, \dots variam em $\boldsymbol{\beta}$, $U_a = \partial l(\boldsymbol{\theta}; \mathbf{y}) / \partial \delta_a$, $U_{ab} = \partial^2 l(\boldsymbol{\theta}; \mathbf{y}) / \partial \delta_a \partial \delta_b$, $U_{abc} = \partial^3 l(\boldsymbol{\theta}; \mathbf{y}) / \partial \delta_a \partial \delta_b \partial \delta_c$, $U_{sta} = \partial^3 l(\boldsymbol{\theta}; \mathbf{y}) / \partial \beta_s \partial \beta_t \partial \delta_a$, $\kappa_{ab} = E(U_{ab})$, $\kappa_{abc} = E(U_{abc})$, $\kappa_{a,bc} = E(U_a U_{bc})$, $\kappa_{a,b,c} = E(U_a U_b U_c)$, $\kappa_{sta} = E(U_{sta})$, $\kappa_{a,b,cd} = E(U_a U_b U_{cd}) - \kappa_{a,b} \kappa_{cd}$, $\kappa_{a,b,c,d} = E(U_a U_b U_c U_d) - \kappa_{a,b} \kappa_{c,d} - \kappa_{a,c} \kappa_{b,d} - \kappa_{a,d} \kappa_{b,c}$.

Os κ 's acima satisfazem algumas relações, as chamadas identidades de Bartlett, que facilitam seus cálculos. Algumas destas relações são expressas por: $\kappa_{a,b} = -\kappa_{ab}$, $\kappa_{a,bc} = (\kappa_{bc})_a - \kappa_{abc}$ e $\kappa_{a,b,c} = 2\kappa_{abc} - \{(\kappa_{ab})_c + (\kappa_{ac})_b + (\kappa_{bc})_a\}$.

Além de alguns dos κ 's calculados no Apêndice D, outros κ 's necessários ao cálculo

dos coeficientes b_{il} 's e ξ são expressos abaixo:

$$\begin{aligned}
\kappa_{s,t,v} &= \delta_{(0,1,0,0,0)} \sum_{\ell=1}^n \frac{1}{\phi_\ell^*} \{(sv, t)_\ell + (s, tv)_\ell + (st, v)_\ell \\
&\quad - (s, vt)_\ell - (st, v)_\ell - (t, vs)_\ell\}, \\
\kappa_{a,b,c} &= 2\kappa_{abc} = \frac{3(\delta_{(0,1,0,0,2)} - 1)}{2} \sum_{\ell=1}^n \frac{h'_\ell h''_\ell}{\phi_\ell^{*2}} (a, b, c)_\ell, \\
&\quad + \frac{(7 - \delta_{(0,0,1,0,3)} - 9\delta_{(0,1,0,0,2)})}{4} \sum_{\ell=1}^n \frac{h_\ell^3}{\phi_\ell^{*3}} (a, b, c)_\ell, \\
\kappa_{s,ta} &= -\kappa_{sta} = \frac{1}{2} \delta_{(0,0,1,0,1)} \sum_{\ell=1}^n \frac{h'_\ell}{\phi_\ell^{*2}} (s, t, a)_\ell + \delta_{(0,1,0,0,0)} \sum_{\ell=1}^n \frac{h'_\ell}{\phi_\ell^{*2}} (s, t, a)_\ell, \\
\kappa_{s,ta} &= -\delta_{(0,1,0,0,0)} \sum_{\ell=1}^n \phi_\ell^* (s, ta)_\ell, \\
\kappa_{a,bc} &= -\frac{3(\delta_{(0,1,0,0,2)} - 1)}{4} \sum_{\ell=1}^n \frac{h'_\ell h''_\ell}{\phi_\ell^{*2}} (a, b, c)_\ell, \\
&\quad - \frac{(7 - \delta_{(0,0,1,0,3)} - 9\delta_{(0,1,0,0,2)})}{8} \sum_{\ell=1}^n \frac{h_\ell^3}{\phi_\ell^{*3}} (a, b, c)_\ell \text{ e} \\
\kappa_{a,bs} &= \kappa_{a,b,s} = 0.
\end{aligned}$$

Para obter os coeficientes b_{il} 's e ξ na classe dos MNLSH para o teste de $H_0 : \boldsymbol{\delta}_1 = \boldsymbol{\delta}_1^{(0)}$, devemos substituir os κ 's obtidos neste Apêndice e no Apêndice D nas expressões dos b_{il} 's e ξ apresentadas na Seção 4.2. Aqui, obteremos em detalhes os coeficientes b_{41}^0 e ξ . Os coeficientes são obtidos de forma análoga.

Da Seção 4.2, temos que o coeficiente b_{41}^0 é dado por

$$\begin{aligned}
b_{41}^0 &= \frac{1}{4} \sum_{a,b,c=1}^k \kappa_{abc} \kappa^{a,b} \epsilon_c^* - \frac{1}{2} \sum_{a,b,c=1}^k (\kappa_{abc} + 2\kappa_{a,bc}) \epsilon_a^* \epsilon_b^* \epsilon_c^* \\
&\quad - \frac{1}{4} \sum_{a,b,c=1}^k (4\kappa_{a,bc} + 3\kappa_{abc}) a_{ab} \epsilon_t^* - \frac{1}{2} \sum_{r=2}^k \sum_{b,c=1}^k (\kappa_{abc} + \kappa_{a,bc}) \epsilon_a \epsilon_b^* \epsilon_c^*.
\end{aligned}$$

Substituindo os κ 's em b_{41}^0 , temos que

$$\begin{aligned}
b_{41}^0 &= \frac{1}{4} \sum_{a,b,c=1}^k \sum_{\ell=1}^n Q_1(a, b, c)_\ell \kappa^{a,b} \epsilon_c^* + \frac{1}{2} \sum_{a,b,c=1}^k \sum_{\ell=1}^n Q_1(a, b, c)_\ell \epsilon_a^* \epsilon_b^* \epsilon_c^* \\
&\quad + \frac{1}{4} \sum_{a,b,c=1}^k \sum_{\ell=1}^n Q_1(a, b, c)_\ell a_{ab} \epsilon_c^*,
\end{aligned}$$

em que $Q_1 = \frac{1}{8}\{1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\}$. Reorganizando os termos, temos que

$$b_{41}^0 = \frac{1}{4}Q_1 \sum_{\ell=1}^n \left(\sum_{a,b=1}^k (a)_\ell \kappa^{a,b}(b)_\ell \right) \left(\sum_{c=1}^k (c)_\ell \epsilon_c^* \right) + \frac{1}{2}Q_1 \sum_{\ell=1}^n \left(\sum_{a=1}^k (a)_\ell \epsilon_a^* \right) \left(\sum_{b=1}^k (b)_\ell \epsilon_b^* \right) \\ + \left(\sum_{c=1}^k (c)_\ell \epsilon_c^* \right) + \frac{1}{4}Q_1 \sum_{\ell=1}^n \left(\sum_{a,b=1}^k (a)_\ell a_{ab}^*(b)_\ell \right) \left(\sum_{c=1}^k (c)_\ell \epsilon_c^* \right).$$

Em notação matricial, temos

$$b_{41}^0 = \frac{1}{4}Q_1 \text{tr}\{\mathbf{Z}_{\delta d} \mathbf{T}\} + \frac{1}{2}Q_1 \text{tr}\{\mathbf{T}^{(3)}\} + \frac{1}{4}Q_1 \text{tr}\{\mathbf{Z}_{\delta d} \mathbf{T}\}.$$

Da Seção 4.2, temos que ξ é dado por

$$\xi = \frac{1}{2} \sum_{a=1}^k \sum_{s,t=1}^p \kappa_{sta} \kappa^{s,t} \epsilon_a^*,$$

com $Q_2 = -\frac{1}{2}\{\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}\}$. Substituindo os valores dos κ 's e rearranjando os termos, temos

$$\xi = \frac{1}{2}Q_2 \sum_{\ell=1}^n \frac{1}{\phi_\ell^*} \left(\sum_{s,t=1}^p (s)_\ell a_{st}(t)_\ell \right) \left(\sum_{a=1}^k (a)_\ell \epsilon_a^* \right).$$

Em notação matricial,

$$\xi = \frac{1}{2}Q_2 \text{tr}\{\mathbf{\Lambda} \mathbf{Z}_{\beta d} \mathbf{T}\}.$$

Para a comparação dos poderes dos testes da razão de verossimilhanças, Wald, score e gradiente até ordem $n^{-1/2}$, sob uma sequência de hipóteses contíguas convergindo para a hipótese nula à taxa de $n^{-1/2}$, temos da Seção 4.3 que

$$\pi_i - \pi_j = \sum_{\ell=0}^3 (b_{j\ell} - b_{i\ell}) G_{k-1+2\ell,\lambda}(x_\alpha), \quad i, j = 1, 2, 3, 4, \quad i \neq j,$$

sendo os $b_{i\ell}$'s e $b_{j\ell}$'s os coeficientes apresentados na Seção 4.2 e π_i a função de poder dos testes supracitados. Dessa maneira, temos

$$\begin{aligned} \pi_1 - \pi_3 &= \sum_{k=0}^3 (b_{3k} - b_{1k}) G_{k-1+2\ell,\lambda}(x) \\ &= (b_{30} - b_{10}) G_{k-1,\lambda}(x) + (b_{31} - b_{11}) G_{k+1,\lambda}(x) \\ &\quad + (b_{32} - b_{12}) G_{k+3,\lambda}(x) + (b_{33} - b_{13}) G_{k+5,\lambda}(x). \end{aligned} \tag{F.1}$$

Substituindo os coeficientes $b_{i\ell}$'s e $b_{j\ell}$'s em (F.1), temos

$$\begin{aligned} \pi_1 - \pi_3 &= Q_1 \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{T}\} (G_{k+1,\lambda}(x) - G_{k+3,\lambda}(x)) \\ &\quad + \frac{1}{3} Q_1 \text{tr}\{\mathbf{T}^{(3)}\} (G_{k+3,\lambda}(x) - G_{k+5,\lambda}(x)). \end{aligned} \tag{F.2}$$

Sabemos ainda que

$$G_{m,\lambda}(x) - G_{m+2,\lambda}(x) = 2g_{m+2,\lambda}(x),$$

sendo $g_{m,\lambda}(x)$ a função densidade de probabilidade de uma variável aleatória qui-quadrado não central com m graus de liberdade e parâmetro de não centralidade λ . Dessa maneira, podemos reescrever (F.2) como

$$\pi_1 - \pi_3 = \frac{1}{2}Q_1 \text{tr}\{(\mathbf{Z}_\delta - \mathbf{Z}_{\delta_0})_d \mathbf{T}\} g_{k+3,\lambda}(x) + \frac{1}{6}Q_1 \text{tr}\{\mathbf{T}^{(3)}\} g_{k+5,\lambda}(x).$$

De maneira análoga procedemos para obter as demais expressões $\pi_i - \pi_j$.

APÊNDICE G

Coelhos europeus na Austrália

Tabela G.1: Peso das lentes dos olhos de coelhos europeus (y), em miligramas, e a idade (x), em dias, numa amostra contendo 71 observações. (Wei, 1998, Exemplo 6.8)

x	y	x	y	x	y	x	y
15	21,66	75	94,60	218	174,03	347	188,38
15	22,75	82	92,50	219	173,54	354	189,70
15	22,30	85	105,00	224	178,86	357	195,31
18	31,25	91	101,70	225	177,68	375	202,63
28	44,79	91	102,90	227	173,73	394	224,82
29	40,55	97	110,00	232	159,98	513	203,30
37	50,25	98	104,30	232	161,29	535	209,70
37	46,88	125	134,90	237	187,07	554	233,90
44	52,03	142	130,68	246	176,13	591	234,70
50	63,47	142	140,58	258	183,40	648	244,30
50	61,13	147	155,30	276	186,26	660	231,00
60	81,00	147	152,20	285	189,66	705	242,40
61	73,09	150	144,50	300	186,09	723	230,77
64	79,09	159	142,15	301	186,70	756	242,57
65	79,51	165	139,81	305	186,80	768	232,12
65	65,31	183	153,22	312	195,10	860	246,70
72	71,90	192	145,72	317	216,41		
75	86,10	195	161,10	338	203,23		

- Barndorff-Nielsen, O.E. (1983). On a formula to the distribution of the maximum likelihood estimator. *Biometrika*, **70**, 343–365.
- Barndorff-Nielsen, O.E. (1994). Adjusted versions of profile likelihood and directed likelihood, and extended likelihood. *Journal of the Royal Statistical Society B*, **56**, 125–140.
- Bartlett, M.S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London A*, **160**, 268–282.
- Bayer, F.M.; Cribari-Neto, F. (2013). Bartlett corrections in beta regression models. *Journal of Statistical Planning and Inference*, **143**, 531–547.
- Brito, C.C.R. (2009). Correção de Bartlett nos modelos não-lineares simétricos heteroscedásticos. Dissertação de Mestrado, Programa de pós-graduação em Estatística, Universidade Federal de Pernambuco – *UFPE*.
- Cordeiro, G.M.; Vasconcellos, K.L.P.; Santos, M.L.F. (1998). On the second-order bias of parameter estimates in nonlinear regression models with Student t errors. *Journal of Statistical Computation and Simulation*, **60**, 363–378.
- Cordeiro, G.M. (1999). Introdução à Teoria Assintótica. *IMPA, XXII*. Colóquio Brasileiro de Matemática, Rio de Janeiro.
- Cordeiro, G.M. (2003). Three Corrected Score Tests for Generalized Linear Models with Dispersion Covariates. *Statistica Neerlandica*, **57**, 391–409.

- Cordeiro, G.M. (2004). Corrected likelihood ratio tests in symmetric nonlinear regression models. *Journal of Statistical Computation and Simulation*, **74**, 609–620.
- Cordeiro, G.M.; Cribari-Neto, F. (2014). An introduction to Bartlett correction and bias reduction. *Springer*, New York.
- Cordeiro, G.M.; Ferrari, S.L.P.; Uribe-Opazo, M.A.; Vasconcellos, K.L.P. (2000). Corrected maximum-likelihood estimation in a class of symmetric nonlinear regression models. *Statistics and Probability Letters*, **46**, 317–328.
- Cordeiro, G.M.; Ferrari, S.L.P. (1991). A modified score test statistic having chi-squared distribution to order $1/n$. *Biometrika*, **78**, 573–582.
- Cordeiro, G.M.; Ferrari, S.L.P.; Cysneiros, A.H.M.A. (1998). A formula to improve score test statistics *Journal of Statistical Computation and Simulation*, **62**, 123–136.
- Cox, D.R.; Hinkley, D.V. (1974). Theoretical Statistics. *Chapman and Hall*, London.
- Cox, D.R.; Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*, **49**, 1–39.
- Cribari-Neto, F.; Queiroz, M.P.F. (2014). On testing inference in beta regression. *Journal of Statistical Computation and Simulation*, **84**, 186–203.
- Cysneiros, F.J.A.; Cordeiro, G.M.; Cysneiros, A.H.M.A. (2010). Corrected maximum likelihood estimators in heteroscedastic symmetric nonlinear models. *Journal of Statistical Computational and Simulation*, **80**, 451–461.
- Cysneiros, F.J.A.; Paula, G.A.; Galea, M. (2005). Modelos Simétricos Aplicados. 9º Escola de Modelos de Regressão, São Paulo.
- Cysneiros, F.J.A.; Paula, G.A.; Galea, M. (2007). Heteroscedastic symmetrical linear models. *Statistics and Probability Letters*, **77**, 1084–1090.
- Cysneiros, A.H.M.A.; Rodrigues, K.S.P.; Cordeiro, G.M.; Ferrari, S.L.P. (2010b). Three-Bartlett-type corrections for Score statistics in Symmetric nonlinear regression models. *Statistical Papers*, **51**, 273–284.

- DiCiccio, T.J.; Stern, S.E. (1994). Frequentist and Bayesian Bartlett correction of test statistics based on adjusted profile likelihoods. *Journal of the Royal Statistical Society. Series B (Methodological)*, **56**, 397–408.
- Doornik, J.A. (2006). Ox: an object-oriented matrix programming language, 4th edn. Timberlake Consultants Ltd, London.
- Efron, B. (1979). Bootstrap Methods: another look at the jackknife. *Annals of Statistics* **7(1)**, 1–26.
- Efron, B.; Tibshirani R.J. (1993). An Introduction to the Bootstrap. *Chapman & Hall*, New York.
- Fang, K.T.; Anderson, T.W. (1990). Statistical Inference in Elliptical Contoured and Related Distributions. *Allerton Press*, New York.
- Fang, K.L.; Kotz, S.; Ng, K.W. (1990). Symmetric Multivariate and Related Distributions, *Chapman and Hall*, London.
- Ferrari, S.L.P.; Cysneiros, A.H.M.A.; Cribari-Neto, F. (2004). An improved test for heteroskedasticity using adjusted modified profile likelihood inference. *Journal of Statistical Planning and Inference*, **124**, 423–437.
- Ferrari, S.L.P.; Lucambio, F.; Cribari-Neto, F. (2005). Improved profile likelihood inference. *Journal of Statistical Planning and Inference*, **134**, 373–391.
- Ferrari, S.L.P.; Uribe-Opazo, M.A. (2001). Corrected likelihood ratio test in class of symmetric linear regression models. *Brazilian Journal of Probability and Statistics*, **15**, 49–67.
- Hayakawa, T. (1975). The Likelihood ratio criterion for composite hypothesis under local alternative. *Biometrika*, **62**, 451–460.
- Harris, P.; Peers, H.W. (1980). The local power of the efficient score test statistic. *Biometrika*, **67**, 525–529.
- Kakisawa, Y. (1996). Higher order monotone Bartlett-type adjustment for some multivariate test statistics. *Biometrika*, **83**, 923–927.

- Lang, K.L.; Little, R.J.A.; Taylor, J.M.G. (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, **84**, 881–896.
- Lawley, D.N. (1956). A general method for approximating to the distribution of the likelihood ratio criteria. *Biometrika*, **71**, 233–244.
- Lemonte, A.J. (2011). Local power of some asymptotic tests in exponential nonlinear models. *Journal of Statistical Planning and Inference*, **141**, 1981–1989.
- Lemonte, A.J. (2012). Local power properties of some asymptotic tests in symmetric linear regression models. *Journal of Statistical Planning and Inference*, **142**, 1178–1188.
- Lemonte, A.J. (2014). On local power properties of the LR, Wald, score and gradient tests in nonlinear mixed-effects models. *Annals of the Institute of Statistical Mathematics*.
- Lemonte, A.J.; Ferrari, S.L.P. (2012). The Local power of gradient test. *Annals of The Institute of Statistical Mathematics*, **64**, 373–381.
- Lemonte, A.J.; Ferrari, S.L.P. (2012b). Local power and size properties of the LR, Wald, score and gradient tests in dispersion models. *Statistical Methodology*, **9**, 537–554.
- Lucas, A. (1997). Robustness of the Student- t based on M -estimator. *Communications in Statistics, Theory and Methods*, **26**, 1165–1182.
- McCullagh, P.; Tibishirani, R. (1990). A simple method for the adjustment of profile likelihood. *Journal of the Royal Statistical Society B*, **52**, 325–344.
- Nascimento, K.P. (2010). Correção tipo-Bartlett em modelos não lineares simétricos heteroscedásticos. Dissertação de Mestrado, Programa de pós-graduação em Biometria e Estatística Aplicada, Universidade Federal Rural de Pernambuco – *UFRPE*.
- Nocedal, J.; Wright, S.J. (1999). Numerical Optimization. *Springer Verlag*, New York.
- Pace, L.; Savan, A. (1997). Principles of Statistical Inference from a Neo-Fisherian Perspective. *World Scientific*, Singapore.
- Rao, C.R. (1973). Linear Statistical Inference and its Applications - Second Edition. *Wiley*, New York.
- Rocke, D.M. (1989). Bootstrap Bartlett adjustment in seemingly unrelated regression. *Journal of the American Statistical Association*, **84**, 598–601.

- Sen, P.K.; Singer, J.M. (1993). Large Sample Methods in Statistics: An Introduction with Applications. *Chapman and Hall/CRC*, Boca Raton.
- Severini, T.A. (2000). Likelihood Methods in Statistics. *Oxford University Press*, Oxford.
- Silva, M.F. (2005). Estimação e teste de hipóteses baseados em verossimilhança perfilada. Tese de doutorado, Programa de pós-graduação em Estatística, Universidade de São Paulo – *USP*.
- Simonoff, J.S.; Tsai, C.H. (1994). Use of modified profile likelihood for improved tests of constancy of variance in regression. *Applied Statistics*, **43**, 357–370.
- Stern, S.E. (1997). A second-order adjustment to the profile likelihood in the case of a multidimensional parameter of interest. *Journal of the Royal Statistical Society B*, **59**, 653–665.
- Terrell, G.R. (2002) The gradient statistic. *Computing Science and Statistics*, **34**, 206–215.
- Vargas, T.M. (2013). Estatística gradiente: teoria assintótica de alta ordem e correção tipo-Bartlett. Tese de doutorado, Programa de pós-graduação em Estatística, Universidade de São Paulo – *USP*.
- Vargas, T.M.; Ferrari, S.L.P.; Lemonte, A.J. (2013). Gradient statistic: Higher-order asymptotics and Bartlett-type correction. *Electronic Journal of Statistics*, **7**, 43–61.
- Vargas, T.M.; Ferrari, S.L.P.; Lemonte, A.J. (2014). Improved likelihood inference in generalized linear models. *Computational Statistics and 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 and Data Analysis*, **59**, 161–170.
- Uribe-Opazo, M.A.; Ferrari, S.L.P.; Cordeiro, G.M. (2008). Improved score test in symmetric linear regression model. *Communications in Statistics Theory Methods*, **37**, 261–276.
- Wei, B.L. (1998). Exponential Family Nonlinear Models. *Springer-Verlag*, Singapore.