

VEROSSIMILHANÇA PERFILADA NOS
MODELOS NÃO LINEARES SIMÉTRICOS
HETEROSCEDÁSTICOS

Mariana Correia de Araújo

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

Área de concentração: Estatística Matemática

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

Recife, fevereiro de 2012

Catálogo na fonte
Bibliotecária Jane Souto Maior, CRB4-571

Araújo, Mariana Correia de
Verossimilhança perfilada nos modelos não lineares
simétricos heteroscedásticos / Mariana Correia de Araújo
- Recife: O Autor, 2012.
xi, 64 folhas: il., fig., tab.

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

Inclui bibliografia e apêndice.

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

519.9

CDD (23. ed.)

MEI2012 – 017

Universidade Federal de Pernambuco
Pós-Graduação em Estatística

17 de fevereiro de 2012

Nós recomendamos que a dissertação de mestrado de autoria de

Mariana Correia de Araújo

intitulada

“Verossimilhança perfilada nos modelos não lineares simétricos heteroscedásticos”

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

Coordenador da Pós-Graduação em Estatística

Banca Examinadora:

Audrey Helen Mariz de Aquino Cysneiros orientadora

Betsabé Grimalda Blas Achic

Miguel Angel Uribe Opazo (UNIOESTE)

Este documento será anexado à versão final da dissertação.

*Dedico este trabalho aos
meus amados pais.*

Agradecimentos

Agradeço primeiramente a Deus, por mais esta vitória e por me ter feito perseverar nas horas mais difíceis.

É com muita alegria que agradeço aos meus pais, por todo carinho, apoio e confiança. Tenho a certeza que hoje, eles mais do que ninguém, independente de estar presente ou apenas nas lembranças, vibram com esta conquista.

Agradeço ao meu irmão, pela horas de conversa, pela companhia e por compreender minha falta em algumas atividades.

Meu muitíssimo obrigada à professora Audrey Helen Mariz de Aquino Cysneiros, por sua orientação, amizade, paciência, ensinamentos, sugestões e principalmente pela confiança em mim depositada, fatores primordiais para a conclusão deste mestrado.

Ao meu amigo e companheiro de estudo William Marciano, por tudo que compartilhamos ao longo destes dois anos, por sua paciência de monge, pelos ensinamentos e por sua amizade.

A Fernando, Iván, Manoel, Rodrigo e Flávius, pelas horas de estudo dispensadas, pela companhia, pelas conversas, pelo convívio, pelas palavras de conforto e pelo incentivo.

Aos demais colegas do mestrado, pelo convívio e horas de estudo dispensadas.

Aos meus queridos amigos Larissa, Romison, Rafaella, Natália, Mayra e Silvana, por entender minha falta nestes últimos dois anos, pelo incentivo, pela amizade, pelas palavras de conforto nos momentos difíceis, pelo carinho e por vibrarem comigo a cada conquista.

Aos professores da pós-graduação em Estatística da UFPE, pelo conhecimento repassado.

A Valéria Bittencourt, pela atenção e excelência em seu trabalho.

Aos professores Betsabé Grimalda Blas Achic e Miguel Angel Uribe Opazo, pelas sugestões e correções.

À CAPES, pelo apoio financeiro.

*"Linda menina abra as asas
Voe o mais alto que puder
Acredite, você pode ir ainda mais alto
Feche os olhos, imagine-se
Nunca olhe para baixo
Dê a mão ao céu
Alcance o universo, almeje o infinito
Transpasse as barreiras da realidade
E siga seu coração."*

Daniel Lessa

Resumo

Nesta dissertação abordamos o desempenho do teste da razão de verossimilhanças usual e suas versões modificadas em pequenas amostras na classe dos modelos não lineares simétricos heteroscedásticos (MNLSH), mais especificamente, nos modelos t -Student com 4 graus de liberdade e Exponencial potência com parâmetro de forma $k = 0, 3$. Além do teste usual, são considerados os testes baseados na estatística da razão de verossimilhanças corrigida via Bartlett (1937), LR^* , na estatística da razão de verossimilhanças perfiladas modificadas via Cox e Reid (1987), LR_m , e sua respectiva versão corrigida via DiCiccio e Stern (1994), LR_m^* . Desse modo, os objetivos principais deste trabalho são obter um fator de correção de Bartlett para a estatística LR_m na classe dos MNLSH e realizar um estudo de simulação para avaliar o desempenhos dos testes de hipóteses baseados na estatística da razão de verossimilhanças usual, LR , e nas estatísticas LR^* , LR_m e LR_m^* . Neste estudo de simulação avaliamos o comportamento dos quatro testes em questão com relação ao tamanho, poder e discrepância relativa de quantis em amostras de tamanhos finitos e pode-se observar que, de modo geral, o teste baseado na estatística LR_m^* apresentou o melhor desempenho.

Palavras chave: Verossimilhança perfilada modificada; Teste da razão de verossimilhanças; Correção de Bartlett.

Abstract

This thesis addresses the performance of the usual likelihood ratio test and its modified versions for small samples in a class of heteroskedastic symmetric nonlinear models (MNLSH), more specifically, on the models Student-t with 4 degrees of freedom and power exponential with shape parameter $k = 0, 3$. We consider the tests based on the likelihood ratio statistic, LR , its Bartlett-corrected version (Bartlett, 1937), LR^* , the modified profile likelihood ratio statistic, LR_m , and its Bartlett-corrected version (DiCiccio Stern, 1994), LR_m^* . On this way, our main goal is derive a Bartlett adjustment to the modified profile likelihood ratio statistic for testing heteroskedasticity in a class of MNLSH and perform a simulation study to evaluate the performance of the tests based on the statistics LR , LR^* , LR_m and LR_m^* . In this simulation study, we evaluated the behavior of this four tests in question with respect to size, power and relative quantile discrepancy in samples of finite size and was observed that the test based on the LR_m^* outperformed all the others tests considered.

Keywords: Modified profile likelihood; likelihood ratio test; Bartlett correction.

Sumário

1	Introdução	1
2	Modelos não lineares simétricos heteroscedásticos	6
2.1	Introdução	6
2.2	Definição	7
2.3	Estimação dos parâmetros e testes de hipóteses nos MNLSH	9
3	Ajustes para o teste da razão de verossimilhanças	16
3.1	Introdução	16
3.2	Correção de Bartlett nos MNLSH	17
4	Resultados Numéricos	28
5	Considerações finais	42
A	Cálculo dos cumulantes	45
A.1	Derivadas do logaritmo da função de verossimilhança	45
A.2	Cálculo dos Cumulantes	49
A.3	Derivadas dos Cumulantes	50

B	Valores dos $\delta's$	52
B.1	Distribuição Normal	52
B.2	Distribuição de Cauchy	53
B.3	Distribuição t-Student	54
B.4	Distribuição t-Student generalizada	55
B.5	Distribuição Logística I	56
B.6	Distribuição Logística II	57
B.7	Exponencial potência	58
	Referências Bibliográficas	60

Lista de Figuras

- 4.1 *Gráfico das discrepâncias relativas de quantis - Modelo t-Student*
com $\nu = 4$, $n = 35$, $p = 5$ e $q = 3$ 37
- 4.2 *Gráfico das discrepâncias relativas de quantis - Modelo Expo-*
nencial potência com $k = 0,3$; $n = 35$, $p = 5$ e $q = 3$ 37

Lista de Tabelas

2.1	Valores de ξ para algumas distribuições simétricas	8
2.2	Expressões para $g(z)$, $\frac{g'(z)}{g(z)}$ e $s = -\frac{2g'(z)}{g(z)}$ para algumas distribuições simétricas.	11
4.1	Tamanho dos testes - Modelo t -Student com $\nu = 4, 10$ e 30 e Exponencial potência com $k = 0, 3$; $p = 3$, $q = 3$ e diversos valores para n	38
4.2	Tamanho dos testes para os Modelo t -Student com $\nu = 4$ e Exponencial potência com $k = 0, 3$; ambos com $n = 35$, $q = 2$ e diversos valores para p	39
4.3	Tamanho dos testes para os Modelo t -Student com $\nu = 4$ e Exponencial potência com $k = 0, 3$; ambos com $n = 35$, $p = 3$ e diversos valores para q	40
4.4	Poder dos testes - Modelo t -Student com $\nu = 4$, $p = 5$, $q = 3$, $n = 35$ e $\alpha = 10\%$	41
4.5	Poder dos testes - Modelo Exponencial potência com $k = 0, 3$; $p = 5$, $q = 3$, $n = 35$ e $\alpha = 10\%$	41

CAPÍTULO 1

Introdução

A suposição de normalidade para os erros de modelos de regressão com resposta contínua, embora seja muito atrativa, pode ser inadequada em algumas situações, sobretudo se os dados a serem modelados apresentarem observações extremas, também chamadas de observações aberrantes.

Sabe-se, na literatura, que a modelagem estatística assumindo distribuição normal para os erros pode ser expressivamente influenciada pela presença de observações aberrantes. Como alternativa à suposição de normalidade, novos modelos menos sensíveis à presença de observações extremas têm sido propostos na literatura. A família de distribuições simétricas fornece uma extensão útil da distribuição normal para a modelagem estatística de dados contendo observações aberrantes, uma vez que contempla distribuições com caudas mais pesadas que as da distribuição normal, e tem sido frequentemente usada na literatura como uma extensão robusta do modelo normal clássico. Grande parte desses resultados podem ser encontrados em Fang *et al.* (1990) e Fang e Anderson (1990).

Ao modelar dados simétricos, frequentemente é assumida a suposição de variância constante (homoscedasticidade). Entretanto, em diversas situações práticas esta condição não é satisfeita. Sob heteroscedasticidade, Cysneiros

et al. (2010) propuseram a classe dos modelos não lineares simétricos heteroscedásticos (MNLSH) e obtiveram uma correção de viés para os estimadores de máxima verossimilhança desta classe de modelos. Brito (2009) obteve uma correção de Bartlett para a estatística da razão de verossimilhanças para os MNLSH e Nascimento (2010) obteve uma correção tipo-Bartlett para a estatística *score* para os MNLSH. Nesta direção, estamos interessados em fazer inferências sob parte dos parâmetros do MNLSH. Tais parâmetros são chamados de parâmetros de interesse, enquanto os demais são denominados de parâmetros de perturbação.

Para modelos com parâmetros de perturbação é prática comum realizar inferências baseando-se em uma pseudo-verossimilhança, que é uma função da verossimilhança genuína envolvendo apenas os parâmetros de interesse. Tal pseudo-verossimilhança, denominada verossimilhança perfilada, é obtida substituindo o vetor de parâmetros de perturbação por uma estimativa consistente na verossimilhança original. Por não se tratar de uma verossimilhança genuína, embora goze de algumas propriedades da mesma, a verossimilhança perfilada pode apresentar, por exemplo, vícios na função *score* e na informação. Outro problema comum refere-se à presença de um número considerável de parâmetros de perturbação, que deteriora a qualidade das aproximações envolvidas nas inferências que se baseiam em resultados assintóticos. A fim de contornar problemas oriundos da não genuinidade de tal função, vários ajustes para a função de verossimilhança perfilada, incluindo os propostos por Barndorff-Nielsen (1983, 1994), Cox e Reid (1987, 1992), McCullagh e Tibishirani (1990) e Stern (1997), foram desenvolvidos, resultando uma verossimilhança perfilada modificada. Algumas dessas versões são descritas em Severini (2000, Capítulo 9) e Pace e Salvan (1997, Capítulo 11). Uma vez que os parâmetros do modelo em estudo são globalmente ortogonais, consideraremos neste trabalho a versão proposta por Cox e Reid (1987).

A grande dificuldade em se obter, quando existentes, as distribuições nulas exatas das estatísticas de teste tem tornado frequente o uso de estatísticas

cujas distribuições são baseadas em aproximações para grandes amostras. Testes baseados nessas estatísticas são fundamentados em valores críticos obtidos de uma distribuição limite conhecida, sendo assim, são denominados assintóticos de primeira ordem. Os testes de hipóteses envolvendo grandes amostras mais comuns são os testes da razão de verossimilhanças, escore e Wald. As estatísticas destes três testes são equivalentes em grandes amostras e, sob condições gerais de regularidade, convergem sob hipótese nula (H_0) para a distribuição χ^2_q , onde q é o número de restrições impostas por H_0 . No mais, para pequenas amostras ou até mesmo amostras de tamanho moderado, a aproximação da distribuição da estatística de teste por χ^2 pode não ser satisfatória, apresentando taxas de rejeição da hipótese nula bastantes distorcidas. Concentrando-nos na estatística da razão de verossimilhanças, uma maneira de melhorar a aproximação da distribuição da estatística de teste pela distribuição χ^2 de referência é aplicar um fator de correção à esta estatística.

Bartlett (1937) propôs um fator de correção a ser incorporado à estatística da razão de verossimilhanças de modo que sua versão corrigida (LR^*) apresenta melhor aproximação da distribuição χ^2 de referência do que a estatística usual (LR). A partir da verossimilhança perfilada modificada deriva-se a estatística da razão de verossimilhanças perfiladas modificadas (LR_m), que sob H_0 e certas condições de regularidade, tem distribuição χ^2_q , onde q é o número de restrições impostas pela hipótese nula. Mesmo oriunda de uma verossimilhança ajustada, a estatística LR_m apresenta a mesma problemática que a estatística usual para amostras de tamanho pequeno e até moderado. Visando melhorar a aproximação da distribuição da estatística LR_m por χ^2 , DiCiccio e Stern (1994) propuseram um fator de correção para ser incorporado a esta estatística de modo que a distribuição de sua versão corrigida (LR_m^*) apresenta melhor aproximação por χ^2 .

Na literatura, diversos trabalhos apresentam correções de Bartlett para a estatística da razão de verossimilhanças dos mais diversos modelos estatísticos e em problemas específicos. Cordeiro (1983, 1987) apresentam correções

de Bartlett para os modelos lineares generalizados (MLGs) sendo o fator de escala conhecido e desconhecido, respectivamente. Cribari-Neto e Ferrari (1995) obtiveram o fator de correção de Bartlett para os modelos Lineares normais heteroscedásticos, Cribari-Neto e Zarkos (1995) para os modelos de regressão multivariada. Ferrari e Uribe-Opazo (2001) derivaram a correção de Bartlett para os modelos lineares simétricos, sendo estes resultados generalizados por Cordeiro (2004), que obteve uma correção de Bartlett para a estatística da razão de verossimilhanças para os modelos não lineares simétricos e posteriormente estendidos por Brito (2009), que obteve uma correção de Bartlett para a estatística em questão nos MNLSH. Ferrari e Cribari-Neto (2002) obtiveram o fator de correção para a estatística da razão de verossimilhanças perfiladas modificadas para o modelo de regressão normal linear heteroscedástico considerando o parâmetro que define a componente da heteroscedasticidade sendo escalar. Estes resultados foram generalizados por Ferrari *et al.* (2004), que consideraram o caso em que o parâmetro de interesse é multidimensional. Ainda para a estatística da razão de verossimilhanças perfiladas modificadas, Ferrari *et al.* (2005) obtiveram um fator de correção de Bartlett para a classe dos MLGs e Cysneiros e Ferrari (2006) para os modelos de regressão não lineares da família exponencial.

Nesta direção, esta dissertação tem dois objetivos principais. O primeiro é obter um fator de correção de Bartlett para a estatística da razão de verossimilhanças perfiladas modificadas para o teste de heteroscedasticidade nos MNLSH, baseando-nos na verossimilhança perfilada modificada proposta por Cox e Reid (1987). O segundo objetivo desta dissertação é realizar um estudo de simulação a fim de comparar o desempenho dos testes de heteroscedasticidade nos MNLSH baseados nas estatística da razão de verossimilhanças, sua versão corrigida via Bartlett (1937), na estatística da razão de verossimilhanças perfiladas modificadas e sua versão corrigida via DiCiccio e Stern (1994). É importante observar que não há na literatura nenhum estudo de simulação para avaliar o desempenho do teste de heteroscedasticidade baseado na estatística LR^* para os MNLSH, apenas foi proposto um fator de correção

de Bartlett para a estatística em questão, vide Brito (2009).

Esta dissertação está dividida em cinco capítulos e dois apêndices. No segundo Capítulo, apresentamos a classe de distribuições simétricas juntamente com suas características, propriedades e também os modelos não lineares simétricos heteroscedásticos, explanando seus aspectos inferenciais.

No terceiro Capítulo, discutimos sobre os ajustes propostos por Bartlett (1937) e por DiCiccio e Stern (1994) para as estatísticas da razão de verossimilhanças usual e razão de verossimilhanças perfiladas modificadas baseada na verossimilhança perfilada modificada proposta por Cox e Reid (1987), respectivamente. Ainda no Capítulo 3, fornecemos uma expressão em forma fechada para fator de correção para a estatística LR_m para a classe dos MNLSH.

No quarto Capítulo, desenvolvemos um estudo de simulação de Monte Carlo a fim de comparar o desempenho dos testes da razão de verossimilhanças baseados nas estatísticas LR , LR^* , LR_m e LR_m^* . Por fim, no quinto Capítulo apresentamos algumas considerações finais.

As avaliações numéricas e simulações computacionais descritas ao longo deste trabalho foram realizadas utilizando a linguagem matricial de programação `0x` (Doornik, 2006) em sua versão 6.2 para o sistema operacional Windows. As maximizações não lineares necessárias para o cálculo de estimativas de máxima verossimilhança foram realizadas a partir do algoritmo quasi-Newton BFGS (Nocedal e Wright, 1999, Capítulo 8) e o algoritmo SQP (Nocedal e Wright, 1999, Capítulo 18, similar ao Algoritmo 18.7), disponíveis em funções pré-definidas da linguagem `0x`. Os gráficos apresentados neste texto foram produzidos utilizando o ambiente de programação R em sua versão 2.14.0 para o sistema operacional Windows.

Modelos não lineares simétricos heteroscedásticos

2.1 Introdução

Como alternativa robusta ao modelo normal, a classe de distribuições simétricas tem recebido crescente atenção na literatura estatística. Por contemplar distribuições com caudas mais pesadas que as da distribuição normal, essa família de distribuições sofre menor influência da presença de observações extremas do que a distribuição normal.

Na literatura, diversos trabalhos contemplam a classe de distribuições simétricas. Lang *et al.* (1989) propõem o modelo baseado na suposição de erros t -Student, enquanto que Little (1988) e Yamaguchi (1990) utilizam a distribuição normal contaminada. Arellano-Valle (1994) desenvolve um estudo baseado no modelo de regressão t -Student, sendo estes resultados estendidos para modelos de regressão lineares simétricos por Ferrari e Uribe-Opazo (2001). Cordeiro *et al.* (2000) obtiveram a correção de viés dos estimadores de máxima verossimilhança na classe de modelos não lineares simétricos. Cordeiro (2004) obteve uma correção de Bartlett para os modelos não lineares simétricos. Cysneiros *et al.* (2005) propuseram a classe dos modelos simétricos heteroscedásticos, sendo esta classe posteriormente

estendida por Cysneiros *et al.* (2010), que propuzeram a classe dos modelos não lineares simétricos heteroscedástico e apresentaram uma correção de viés para os estimadores de máxima verossimilhança nessa classe de modelos, generalizando os resultados desenvolvidos por Cordeiro (2004). Brito (2009) obteve uma correção de Bartlett para os modelos não lineares simétricos heteroscedásticos, estendendo os resultados obtidos por Cordeiro (2004). Cysneiros *et al.* (2010) derivou um fator de correção tipo-Bartlett para a estatística *escore* nos modelos não lineares simétricos, sendo este resultado posteriormente estendido por Nascimento (2010), que considerou os modelos não lineares simétricos heteroscedásticos.

2.2 Definição

Sejam $Y_1, \dots, Y_n$ variáveis aleatórias independentes. Dizemos que cada variável aleatória Y_i tem distribuição simétrica com suporte em $\mathbb{R}$, parâmetro de locação $\mu_i \in \mathbb{R}$, e escala $\phi_i > 0$, se sua função de densidade é da forma

$$\pi(y; \mu_i, \phi_i) = \frac{1}{\sqrt{\phi_i}} g(u_i), y \in \mathbb{R}, \quad (2.1)$$

em que $g(u) : \mathbb{R} \rightarrow [0, \infty)$, comumente denominada função geradora de densidade, é tal que $\int_0^\infty g(u) du < \infty$ com $u_i = \phi_i^{-1}(y - \mu_i)^2$. Denotaremos por $Y_i \sim S(\mu_i, \phi_i, g)$, $i = 1, \dots, n$, e denominaremos de variável aleatória simétrica.

Sendo Y_i uma variável aleatória pertencente à família de distribuições dada em (2.1), temos que sua função característica $\psi(t) = E(e^{itY_i})$ é dada por $\psi(t) = e^{it\mu_i} \varphi(t^2\phi_i)$, em que $t \in \mathbb{R}$ para alguma função φ , com $\varphi(x) \in \mathbb{R}$ para $x > 0$. Quando os dois primeiros momentos existem são válidas as seguintes relações:

$$E(Y_i) = \mu_i \quad e \quad Var(Y_i) = \xi\phi_i,$$

sendo $\xi = -2\varphi'(0)$ uma constante positiva que não depende de μ_i e ϕ_i , com $\varphi'(0) = \frac{d\varphi(u)}{du}|_{u=0}$. Alguns valores de ξ podem ser encontrados na Tabela 2.1. Para maiores detalhes, ver Fang *et al.* (1990).

A família simétrica (2.1) mantém a estrutura da distribuição normal, entretanto, elimina a forma específica da densidade de tal distribuição, o que a torna mais abrangente, uma vez que todas as distribuições simétricas contínuas de caudas mais leves e mais pesadas que as da distribuição normal são contempladas. Pertencem à classe das distribuições simétricas, por exemplo, as distribuições normal, normal contaminada, Cauchy, t -Student, t -Student generalizada, logística tipos I e II , logística generalizada, Kotz, Kotz generalizada e exponencial potência. Podemos encontrar aplicações da classe simétrica nas áreas de engenharia, biologia, medicina, economia, entre outras.

Tabela 2.1: Valores de ξ para algumas distribuições simétricas

Distribuições	ξ
Normal	1
t -Student	$\frac{\nu}{\nu-2}, \nu > 2$
t -Student generalizada	$\frac{w}{r-2}, w > 0, r > 2$
Logística tipo I	0,79569
Logística tipo II	$\pi^2/3$
Exponencial potência	$2^{(1+k)} \frac{\Gamma\{3(k+1)/2\}}{\Gamma\{(k+1/2)\}}$

À classe de distribuições dada em (2.1), introduziremos duas estruturas de regressão. Assumiremos que a resposta média seja $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^\top$, com

$$\mu_i = f(x_i; \boldsymbol{\beta}), \quad i = 1, \dots, n, \quad (2.2)$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ é um vetor de dimensão $p \times 1$ ($p < n$) de parâmetros desconhecidos a serem estimados, $\mathbf{x}_i = (x_{i1}, \dots, x_{im})^\top$ representa os m valores das variáveis explicativas associadas à i -ésima observação e $f(\cdot; \cdot)$ é uma função possivelmente não linear no segundo argumento, contínua e duplamente diferenciável com respeito aos componentes de $\boldsymbol{\beta}$, cuja matriz de

derivadas $\tilde{X} = \partial\boldsymbol{\mu}/\partial\boldsymbol{\beta}$ tem posto completo p para todo $\boldsymbol{\beta}$. Os elementos da matriz $\tilde{X}$ são, em geral, funções do vetor de parâmetros $\boldsymbol{\beta}$ desconhecidos. Assumimos ainda uma componente sistemática para o vetor de parâmetros de dispersão $\boldsymbol{\phi} = (\phi_1, \dots, \phi_n)^\top$ dada por

$$\phi_i = h(\tau_i), \quad i = 1, \dots, n, \quad (2.3)$$

onde $h(\cdot)$ é uma função monótona conhecida, contínua e diferenciável do preditor linear de dispersão, definido por $\tau_i = \mathbf{z}_i^\top \boldsymbol{\gamma}$, sendo $\mathbf{z}_i = (z_{i1}, \dots, z_{iq})^\top$ um vetor de dimensão $q \times 1$ de variáveis explicativas que podem ter componentes comuns com $\mathbf{x}_i$ e $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_q)^\top$ um vetor $q \times 1$ de parâmetros desconhecidos a serem estimados. A função $h(\cdot)$ é tipicamente denominada de função de ligação de dispersão e uma escolha comum na literatura para $h(\cdot)$ é $h(\tau) = \exp(\tau)$.

Dessa forma, temos que o modelo não linear simétrico heteroscedástico (MNLSH) é dado por

$$Y_i = \mu_i + \sqrt{\phi_i} \epsilon_i; \quad \epsilon_i \sim S(0, 1, g), \quad (2.4)$$

$i = 1, \dots, n$, onde μ_i e ϕ_i variam sobre as observações com as estruturas não lineares dadas por (2.2) e (2.3), respectivamente.

2.3 Estimação dos parâmetros e testes de hipóteses nos MNLSH

Seja $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ um vetor $(p + q) \times 1$ de parâmetros do modelo definido por (2.4), sendo $\boldsymbol{\gamma}$ e $\boldsymbol{\beta}$ vetores de parâmetros de interesse e perturbação, com dimensões $q \times 1$ e $p \times 1$, respectivamente. A inferência neste trabalho está centrada no vetor de parâmetros $\boldsymbol{\gamma}$. O logaritmo da função de verossimilhança do vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$, dado o vetor de observações $(y_1, \dots, y_n)^\top$, do modelo definido por (2.4) é expresso por

$$l(\boldsymbol{\theta}) = \sum_{i=1}^n \log \left(\frac{g(u_i)}{\sqrt{\phi_i}} \right),$$

sendo $u_i = \frac{(y_i - \mu_i)^2}{\phi_i}$ e μ_i e ϕ_i como definidos em (2.2) e (2.3), respectivamente. Reescrevendo $l(\boldsymbol{\theta})$, temos que

$$l(\boldsymbol{\theta}) = -\frac{1}{2} \sum_{i=1}^n \log(\phi_i) + \sum_{i=1}^n t(z_i), \quad (2.5)$$

em que $t(z_i) = \log g(z_i^2)$, sendo $z_i = \sqrt{u_i} = \frac{(y_i - \mu_i)}{\sqrt{\phi_i}}$. Assumimos que $l(\boldsymbol{\theta})$ seja regular com respeito às derivadas dos componentes de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ até quarta ordem (Cox e Hinkley, 1974, Cap 9).

A fim de representar as derivadas do logaritmo da função de verossimilhança, será introduzida a notação $U_j = \partial l(\boldsymbol{\theta}) / \partial \beta_j$, $U_{ja} = \partial^2 l(\boldsymbol{\theta}) / \partial \beta_j \partial \gamma_a$, $U_{jab} = \partial^3 l(\boldsymbol{\theta}) / \partial \beta_j \partial \gamma_a \partial \gamma_b$, *etc.*, onde os índices $j, l, m, \dots$ estão reservados para denotar os componentes do vetor $\boldsymbol{\beta}$ e os índices $a, b, c, \dots$ para denotar os componentes do vetor $\boldsymbol{\gamma}$. Tais derivadas podem ser encontradas no Apêndice A. Denotamos ainda as derivadas de μ_i com relação aos componentes de $\boldsymbol{\beta}$ por $(j)_i = \partial \mu_i / \partial \beta_j$, $(jl)_i = \partial^2 \mu_i / \partial \beta_j \partial \beta_l$, $(jl, m)_i = (\partial^2 \mu_i / \partial \beta_j \partial \beta_l)(\partial \mu_i / \partial \beta_m)$, *etc.*, e as derivadas de τ_i com relação aos componentes de $\boldsymbol{\gamma}$ por $(a)_i = \partial \tau_i / \partial \gamma_a$, $(ab)_i = \partial^2 \tau_i / \partial \gamma_a \partial \gamma_b$, $(ab, c)_i = (\partial^2 \tau_i / \partial \gamma_a \partial \gamma_b)(\partial \tau_i / \partial \gamma_c)$, *etc.*

As primeiras derivadas do logaritmo da função de verossimilhança dado em (2.5) com relação aos j e a -ésimos componentes de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$, respectivamente, são dadas por:

$$\begin{aligned} U_j &= \frac{\partial l(\boldsymbol{\theta})}{\partial \beta_j} = - \sum_{i=1}^n t_{(z_i)}^{(1)} \frac{1}{\sqrt{\phi_i}} (j)_i, \quad j = 1, \dots, p \quad \text{e} \\ U_a &= \frac{\partial l(\boldsymbol{\theta})}{\partial \gamma_a} = - \frac{1}{2} \sum_{i=1}^n \frac{h'_i}{\phi_i} (a)_i - \frac{1}{2} \sum_{i=1}^n t_{(z_i)}^{(1)} z_i \frac{h'_i}{\phi_i} (a)_i, \quad a = 1, \dots, q, \end{aligned}$$

sendo $t_{z_i}^{(1)} = \partial t(z_i) / \partial z_i$, $i = 1, \dots, n$.

A função escore total de $\boldsymbol{\theta}$, $\mathbf{U}(\boldsymbol{\theta})$, tem a forma $\mathbf{U}(\boldsymbol{\theta}) = (\mathbf{U}_{\boldsymbol{\beta}}^\top, \mathbf{U}_{\boldsymbol{\gamma}}^\top)^\top$, sendo seus componentes $\mathbf{U}_{\boldsymbol{\beta}}^\top = (U_1, \dots, U_p)^\top$ e $\mathbf{U}_{\boldsymbol{\gamma}}^\top = (U_1, \dots, U_q)$. Em notação matricial, temos:

$$\begin{aligned} \mathbf{U}_\beta &= \tilde{X} S \Lambda(\mathbf{y} - \boldsymbol{\mu}) \quad \text{e} \\ \mathbf{U}_\gamma &= -\frac{1}{2} \tilde{P}^\top \Lambda(S F_1 \mathbf{u} - F_1 \mathbf{1}), \end{aligned} \quad (2.6)$$

em que $\tilde{X} = \partial \boldsymbol{\mu} / \partial \boldsymbol{\beta}$, $\tilde{P} = \partial \boldsymbol{\tau} / \partial \boldsymbol{\gamma}$, $\Lambda = \text{diag}\{\frac{1}{\phi_1}, \dots, \frac{1}{\phi_n}\}$, $S = \text{diag}\{s_1, \dots, s_n\}$, com $s_i = \frac{-2g'(u_i)}{g(u_i)}$, com u_i definido na Seção 2.2, $F_1 = \text{diag}\{h'_1, \dots, h'_n\}$, com $h'_i = \partial \phi_i / \partial \tau_i$, $\mathbf{u} = (u_1, \dots, u_n)^\top$ e $\mathbf{1}$ é um vetor $n \times 1$ de 1's. Daqui em diante, o símbolo “ $\prime$ ” sobre h_i denota a derivada com relação a τ_i . Na Tabela 2.2 são apresentadas características de algumas distribuições pertencentes à família simétrica.

Tabela 2.2: Expressões para $g(z)$, $\frac{g'(z)}{g(z)}$ e $s = -\frac{2g'(z)}{g(z)}$ para algumas distribuições simétricas.

Distribuições	$g(z)$	$\frac{g'(z)}{g(z)}$	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. t -Student generalizada	$\frac{w^{\frac{r}{2}} [w+z^2]^{-\frac{r+1}{2}}}{B(1/2, r/2)}$	$-\frac{r+1}{2(w+z^2)}$	$\frac{r+1}{(w+z^2)}$
5. Logística tipo I	$c \frac{\exp\{-z^2\}}{(1+\exp\{-z^2\})^2}$	$-\frac{1-\exp\{-z^2\}}{1+\exp\{-z^2\}}$	$\frac{2(1-\exp\{-z^2\})}{1+\exp\{-z^2\}}$
6. Logística tipo II	$\frac{\exp\{-z\}}{(1+\exp\{-z\})^2}$	$-\frac{1}{2} \frac{(\exp\{ z\} - 1)}{(z \exp\{ z\} + 1)}$	$\frac{(\exp\{ z\} - 1)}{(z \exp\{ z\} + 1)}$
7. Exponencial potência	$c(k) \exp\left\{-\frac{1}{2}z^{2/(1+k)}\right\}$	$-\frac{z^{-k/(1+k)}}{2(1+k)}$	$\frac{z^{-k/(1+k)}}{(1+k)}$

As segundas derivadas do logaritmo da função de verossimilhança dado em (2.5) são determinadas pelas expressões dadas abaixo:

$$\begin{aligned}
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial l} &= \sum_{i=1}^n t_{(z_i)}^{(2)} \frac{1}{\phi_i} (j, l)_i - \sum_{i=1}^n t_{(z_i)}^{(1)} \frac{1}{\sqrt{\phi_i}} (jl)_i, \\
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial \gamma_a} &= \frac{1}{2} \sum_{i=1}^n t_{(z_i)}^{(2)} z_i \frac{h'_i}{\phi_i^{3/2}} (a)_i (j)_i + \frac{1}{2} \sum_{i=1}^n t_{(z_i)}^{(1)} \frac{h'_i}{\phi_i^{3/2}} (a)_i (j)_i \quad \text{e} \\
\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma_a \partial \gamma_b} &= -\frac{1}{2} \sum_{i=1}^n \frac{h''_i \phi_i - h_i'^2}{\phi_i^2} (a, b)_i + \frac{1}{4} \sum_{i=1}^n t_{(z_i)}^{(2)} z_i^2 \frac{h_i'^2}{\phi_i^2} (a, b)_i \\
&\quad + \frac{1}{4} \sum_{i=1}^n t_{(z_i)}^{(1)} z_i \frac{3h_i'^2 - 2h''_i \phi_i}{\phi_i^2} (a, b)_i.
\end{aligned}$$

Assim, a matriz de informação total de Fisher para $\boldsymbol{\theta}$ é dada por

$$\mathbf{K}(\boldsymbol{\theta}) = - \begin{bmatrix} \mathbf{K}_{\beta\beta} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{\gamma\gamma} \end{bmatrix}.$$

Uma vez que $\mathbf{K}_{\beta\gamma} = \mathbf{0}$, temos que a matriz de informação total $\mathbf{K}(\boldsymbol{\theta})$ tem estrutura bloco-diagonal, vide Cysneiros *et al.* (2010). Desse modo, temos que β e γ são parâmetros globalmente ortogonais (Cox & Reid, 1987). As submatrizes $\mathbf{K}_{\beta\beta}$ e $\mathbf{K}_{\gamma\gamma}$ são definidas, respectivamente, por

$$\begin{aligned}
\mathbf{K}_{\beta\beta} &= E \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_r \partial \beta_s} \right] = \delta_{(0,1,0,0,0)} \tilde{X}^\top \Lambda \tilde{X} \quad \text{e} \\
\mathbf{K}_{\gamma\gamma} &= E \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma_R \partial \gamma_S} \right] = -\tilde{P}^\top V \tilde{P},
\end{aligned}$$

onde V é uma matriz diagonal $n \times n$ cujos elementos são dados por $v_i = (1 - \delta_{(0,1,0,0,2)}) h_i'^2 / 4\phi_i^2$. Assim sendo, temos que

$$\mathbf{K}(\boldsymbol{\theta}) = \begin{bmatrix} -\delta_{(0,1,0,0,0)} \tilde{X}^\top \Lambda \tilde{X} & \mathbf{0} \\ \mathbf{0} & \tilde{P}^\top V \tilde{P} \end{bmatrix},$$

onde a notação $\delta_{(a,b,c,d,e)} = E\{t^{(1)a} t^{(2)b} t^{(3)c} t^{(4)d} z_i^e\}$ para $a, b, c, d, e = 1, 2, 3$ e 4 , $t_{(z_i)}^{(k)} = \partial^k t / \partial z^k$, $k = 1, \dots, 4$. O apêndice B contém os δ 's correspondentes a algumas distribuições pertencentes à classe dos MNLSH.

Quando o modelo a ser estimado envolve parâmetros de perturbação é prática comum realizar inferências baseando-se numa função de verossimilhança perfilada. Tal função é obtida substituindo-se os parâmetros de perturbação por suas respectivas estimativas de máxima verossimilhança para valores fixados do parâmetro de interesse. Dessa forma, a função de verossimilhança perfilada depende apenas dos parâmetros de interesse. O logaritmo da função de verossimilhança perfilada do modelo definido em (2.4) para o parâmetro de interesse γ é dado por

$$l_p(\gamma) = l(\gamma, \hat{\beta}_\gamma), \quad (2.7)$$

em que $\hat{\beta}_\gamma$ é a estimativa de máxima verossimilhança de β dado γ .

A expressão (2.7) sugere um procedimento de maximização em duas etapas, onde a primeira consiste em encontrar o valor único $\hat{\beta}_\gamma$ que maximiza $l(\beta, \gamma)$ com respeito a β para γ fixado e a segunda tem por finalidade determinar o valor de γ que maximiza $l_p(\gamma)$. Sob condições gerais de regularidade, $\hat{\beta}_\gamma$ é solução da equação $U_\beta = \mathbf{0}$, derivada de (2.6). No mais, o estimador de máxima verossimilhança de β restrito a γ não pode ser expresso em forma fechada, sendo necessário o uso de técnicas de maximização restrita para a sua obtenção. Para maiores detalhes destas técnicas, ver Nocedal e Wright (1999, Capítulo 18).

Para a obtenção do estimador de máxima verossimilhança de γ , maximiza-se o logaritmo da função de verossimilhança dado em (2.7) sujeito à restrição $U_\beta = \mathbf{0}$.

Estamos interessados em testar a hipótese nula $H_0 : \gamma = \gamma^{(0)}$ contra a hipótese alternativa $H_1 : \gamma \neq \gamma^{(0)}$, onde $\gamma^{(0)}$ é um vetor de dimensão $q \times 1$ de constantes especificadas. A estatística da razão de verossimilhanças (LR) para o teste acima pode ser expressa por:

$$LR = 2\{l_p(\hat{\gamma}) - l_p(\gamma^{(0)})\},$$

sendo $\hat{\gamma}$ o estimador de máxima verossimilhança de γ . Vale salientar que para o teste de H_0 em questão, a estatística da razão de verossimilhanças baseada

no logaritmo da função de verossimilhança perfilada l_p é igual a estatística da razão de verossimilhanças usual.

Em problemas regulares e sob hipótese nula, a estatística da razão de verossimilhanças tem distribuição assintótica χ_q^2 , em que q é o número de restrições impostas por H_0 .

A função de verossimilhança perfilada $l_p(\gamma)$, por não se tratar de uma verossimilhança genuína, pode apresentar viés na função escore e na informação. A fim de atenuar tais vieses e explorando as consequências da ortogonalidade entre os parâmetros de interesse e perturbação, Cox e Reid (1987) definiram uma versão modificada da verossimilhança perfilada adicionando a ela um termo de ajuste anteriormente à estimação. O logaritmo da função de verossimilhança perfilada modificada proposta por Cox e Reid (1987) para o modelo definido por (2.4) é expresso por:

$$l_{CR}(\gamma) = l(\gamma, \hat{\beta}_\gamma) - \frac{1}{2} \log |j_{\beta\beta}(\gamma, \hat{\beta}_\gamma)|, \quad (2.8)$$

em que $j_{\beta\beta}(\gamma, \hat{\beta}_\gamma)$ é a matriz de informação observada do vetor β quando γ é fixado, sendo expressa por

$$j_{\beta\beta} = -\frac{\partial^2 l(\theta)}{\partial \beta_j \partial \beta_l} = -\sum_{i=1}^n t_{(z_i)}^{(2)} \frac{1}{\phi_i} (j, l)_i + \sum_{i=1}^n t_{(z_i)}^{(1)} \frac{1}{\sqrt{\phi_i}} (jl)_i.$$

Podemos observar que l_{CR} , definido em (2.8), é uma versão penalizada do logaritmo da função de verossimilhança perfilada em que a penalização considera a informação sobre o parâmetro de perturbação. É válido salientar que a utilização de l_{CR} requer ortogonalidade sobre os parâmetros de interesse e perturbação.

A estatística da razão de verossimilhanças baseada na função de verossimilhança perfilada modificada para o teste de $H_0 \times H_1$ é dada por

$$LR_m = 2\{l_{CR}(\hat{\gamma}) - l_{CR}(\gamma^{(0)})\},$$

onde $\hat{\gamma}$ é o estimador de máxima verossimilhança de γ . Sob certas condições de regularidade e sob H_0 , a estatística da razão de verossimilhanças perfiladas

modificadas (LR_m) tem distribuição assintótica χ_q^2 , com erro de aproximação de ordem $O(n^{-1})$.

Ajustes para o teste da razão de verossimilhanças

3.1 Introdução

Em problemas de testes de hipóteses envolvendo grandes amostras é prática comum o uso dos testes da razão de verossimilhanças, escore e Wald. Neste trabalho, nos concentraremos no teste da razão de verossimilhanças. Para grandes amostras e, sob condições gerais de regularidade, a estatística de teste da razão de verossimilhanças (LR) converge sob hipótese nula (H_0) para a distribuição χ_q^2 , onde q é o número de restrições impostas por H_0 e o erro dessa aproximação é de ordem $O(n^{-1})$. Entretanto, para pequenas amostras ou até mesmo para amostras de tamanho moderado, a aproximação da distribuição da estatística LR por χ^2 pode não ser satisfatória, conduzindo a taxas de rejeição bastantes distorcidas e um caminho para melhorar esta aproximação é usar a correção de Bartlett.

Bartlett (1937) propôs a primeira idéia para aprimorar os testes estatísticos. Considerando o teste da razão de verossimilhanças, o autor introduziu um fator de correção à estatística LR de modo que a distribuição da estatística corrigida (LR^*) tem melhor aproximação da distribuição χ^2 de refe-

rência, uma vez que, sob hipótese nula, o valor esperado $E(LR)$ equivale a $q + c + O(n^{-2})$, em que c é uma constante de ordem $O(n^{-1})$ que pode ser estimada sob H_0 , q é o número de restrições impostas por H_0 e n é o tamanho amostral, enquanto que o valor esperado $E(LR^*)$ equivale a $q + O(n^{-1})$.

Em problemas regulares, Lawley (1956) derivou uma fórmula geral para a obtenção de c envolvendo momentos das derivadas de até quarta ordem da função de log-verossimilhança e mostrou que os momentos da estatística LR^* , segundo H_0 , concordam com seus respectivos da distribuição χ^2 de referência exceto por termos de ordem $O(n^{-2})$. O fator de correção de Bartlett não depende do valor da estatística da razão de verossimilhanças, mas pode depender de parâmetros desconhecidos, e assim sendo, estes devem ser substituídos por suas respectivas estimativas de máxima verossimilhança sob hipótese nula.

DiCiccio e Stern (1994) obtiveram a correção de Bartlett para a estatística da razão de verossimilhanças oriunda de uma verossimilhança perfilada modificada, assim, derivando a estatística da razão de verossimilhanças perfiladas modificadas corrigida, cuja distribuição assintótica sob H_0 é χ_q^2 com erro de aproximação de ordem $O(n^{-2})$. O objetivo desta correção é reduzir a distorção de tamanho do teste em amostras de tamanho típico por meio de uma aceleração da taxa de convergência do tamanho verdadeiro para o tamanho nominal (assintótico).

3.2 Correção de Bartlett nos MNLSH

Sejam $Y_1, \dots, Y_n$ variáveis aleatórias independentes. O modelo não linear simétrico heteroscedástico é dado por

$$Y_i = \mu_i + \sqrt{\phi_i} \epsilon_i; \quad \epsilon_i \sim S(0, 1, g),$$

$i = 1, \dots, n$, onde μ_i e ϕ_i variam sobre as observações com as estruturas não lineares dadas por

$$\mu_i = f(x_i; \beta) \text{ e } \phi_i = h(\tau_i); \text{ com } i = 1, \dots, n,$$

respectivamente.

Considere o vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ de um modelo pertencente à classe dos modelos não lineares simétricos heteroscedásticos, sendo $\boldsymbol{\gamma}$ e $\boldsymbol{\beta}$ vetores de parâmetros de interesse e perturbação, e de ordens q e p , respectivamente. Seja $l_p(\boldsymbol{\gamma})$ o respectivo logaritmo da função de verossimilhança perfilada de $\boldsymbol{\gamma}$. A fim de testar as hipóteses $H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}^{(0)}$ versus $H_1 : \boldsymbol{\gamma} \neq \boldsymbol{\gamma}^{(0)}$, sendo $\boldsymbol{\gamma}^{(0)}$ um vetor de constantes conhecidas, a estatística da razão de verossimilhanças usual é definida por

$$LR = 2\{l_p(\hat{\boldsymbol{\gamma}}) - l_p(\boldsymbol{\gamma}^{(0)})\}.$$

Sob condições gerais de regularidade e sob H_0 , LR tem distribuição assintótica χ_q^2 , onde q é o número de restrições impostas por H_0 e o erro desta aproximação é de ordem $O(n^{-1})$.

A fim de aprimorar a qualidade da aproximação da distribuição da estatística LR pela distribuição χ^2 de referência, Bartlett (1937) propôs multiplicar a estatística LR por um fator de correção, comumente chamando de fator de correção de Bartlett, dado por $(1 + c/q)^{-1}$, assim, derivando a estatística da razão de verossimilhanças corrigida LR^* , expressa por

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

em que c é uma constante de ordem $O(n^{-1})$ que pode ser estimada sob H_0 . Lawley (1956) obteve, sob condições de regularidade, uma expressão geral para c envolvendo momentos das derivadas de até quarta ordem da função de log-verossimilhança. Esta fórmula geral é dada por

$$c = \epsilon_{p+q} - \epsilon_p,$$

em que o termo ϵ_{p+q} é de ordem $O(n^{-1})$, podendo ser decomposto como

$$\epsilon_{p+q} = \epsilon_p + \epsilon_q + \epsilon_{p,q},$$

em que ϵ_p depende apenas da componente sistemática para o parâmetro de locação, ϵ_q depende apenas da componente sistemática para o parâmetro de

escala e $\epsilon_{p,q}$ depende de ambos os componentes. Em função de cumulantes e derivadas de cumulantes, temos que ϵ_{p+q} é escrito na forma

$$\epsilon_{p+q} = \sum' (l_{rstu} - l_{rstuvw}),$$

sendo que $\sum$ é o somatório sobre todos os componentes do vetor $\boldsymbol{\theta}$, isto é, os índices r, s, t, u, v, w variam sobre os $p + q$ parâmetros, com l_{rstu} e l_{rstuvw} expressos, respectivamente, por

$$\begin{aligned} l_{rstu} &= \lambda^{rs} \lambda^{tu} \left(\frac{\lambda_{rstu}}{4} - \lambda_{rst}^{(u)} + \lambda_{rt}^{(su)} \right), \text{ e} \\ l_{rstuvw} &= \lambda^{rs} \lambda^{tu} \lambda^{vw} \left\{ \lambda_{rtv} \left(\frac{\lambda_{suw}}{6} - \lambda_{sw}^{(u)} \right) \right. \\ &\quad \left. + \lambda_{rtu} \left(\frac{\lambda_{svw}}{4} - \lambda_{sw}^{(v)} \right) + \lambda_{rt}^{(v)} \lambda_{sw}^{(u)} + \lambda_{rt}^{(u)} \lambda_{sw}^{(v)} \right\}, \end{aligned} \quad (3.1)$$

sendo $\lambda_{rs} = E(\partial^2 l(\boldsymbol{\theta}) / \partial \theta_r \partial \theta_s)$, $\lambda_{rst} = (E(\partial^2 l(\boldsymbol{\theta}) / \partial \theta_r \partial \theta_s \partial \theta_t))$ e *etc.*, $\lambda_{rs}^{(t)} = \partial \lambda_{rs} / \partial \theta_t$, $\lambda_{rs}^{(tu)} = \partial^2 \lambda_{rs} / \partial \theta_t \partial \theta_u$ e $-\lambda^{rs}$ representa o elemento (r, s) da inversa da matriz de informação de Fisher $\mathbf{K}(\boldsymbol{\theta})$ de $\boldsymbol{\theta}$. O termo ϵ_p pode ser deduzido a partir da equação (3.1) estendendo o somatório $\sum$ apenas sobre os p parâmetros de perturbação. Em problemas regulares e sob H_0 , a estatística LR^* tem distribuição assintótica χ_q^2 , onde o erro desta aproximação é de ordem $O(n^{-2})$. Uma revisão de correção de Bartlett para a estatística da razão de verossimilhanças pode ser obtida em Cribari-Neto e Cordeiro (1996).

Considerando o vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ de um modelo pertencente à classe dos MNLSH, Brito (2009) obteve um fator de correção de Bartlett para a estatística da razão de verossimilhanças do teste de $H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}^{(0)}$ versus $H_1 : \boldsymbol{\gamma} \neq \boldsymbol{\gamma}^{(0)}$ considerando o caso de heteroscedasticidade multiplicativa, isto é, $\phi_i = \exp\{\mathbf{z}_i^\top \boldsymbol{\gamma}\}$, onde $\boldsymbol{\gamma} = \mathbf{0}$ corresponde ao modelo homoscedástico. Desse modo, a constante c é dada por

$$c = \epsilon_q + \epsilon_{p,q}, \quad (3.2)$$

com

$$\begin{aligned}
\epsilon_q &= \text{tr} \left((Q_3 + Q_4 + Q_5 + Q_6) Z_{\gamma d}^{(2)} \right) + (N_1 + N_3 + N_9) \mathbf{1}^\top \Lambda_2 Z_{\gamma}^{(3)} \Lambda_2 \mathbf{1} \\
&+ (N_2 - 2N_9) \mathbf{1}^\top \Lambda_1 Z_{\gamma}^{(3)} \Lambda_2 \mathbf{1} + (N_4 + N_9) \mathbf{1}^\top \Lambda_1 Z_{\gamma}^{(3)} \Lambda_1 \mathbf{1} \\
&+ (N_5 + N_9) \mathbf{1}^\top \Lambda_2 Z_{\gamma d}^{(2)} Z_{\gamma} \Lambda_2 \mathbf{1} + N_7 \mathbf{1}^\top \Lambda_2 Z_{\gamma d}^{(2)} Z_{\gamma} \Lambda_1 \mathbf{1} \\
&+ (N_6 - 2N_9) \mathbf{1}^\top \Lambda_1 Z_{\gamma d}^{(2)} Z_{\gamma} \Lambda_2 \mathbf{1} + (N_8 + N_9) \mathbf{1}^\top \Lambda_1 Z_{\gamma d}^{(2)} Z_{\gamma} \Lambda_1 \mathbf{1}
\end{aligned}$$

e

$$\begin{aligned}
\epsilon_{p,q} &= -\frac{1}{\delta_{(0,1,0,0,0)}} \text{tr} \left((Q_1 + Q_2) Z_{\beta d} Z_{\gamma d} \right) - N_{18} \mathbf{1}^\top \Lambda_3 Z_{\beta d} Z_{\gamma} Z_{\gamma d} \Lambda_2 \mathbf{1} \\
&- N_{19} \mathbf{1}^\top \Lambda_3 Z_{\beta d} Z_{\gamma} Z_{\gamma d} \Lambda_1 \mathbf{1} - N_{20} \mathbf{1}^\top \Lambda_2 Z_{\gamma d} Z_{\gamma} Z_{\beta d} \Lambda_3 \mathbf{1} \\
&- N_{21} \mathbf{1}^\top \Lambda_1 Z_{\gamma d} Z_{\gamma} Z_{\beta d} \Lambda_3 \mathbf{1} + (2N_{22} + N_{24}) \mathbf{1}^\top \Lambda_3 Z_{\gamma} Z_{\beta}^{(2)} \Lambda_3 \mathbf{1} \\
&+ N_{23} \mathbf{1}^\top \Lambda_3 Z_{\beta d} Z_{\gamma} Z_{\beta d} \Lambda_3 \mathbf{1},
\end{aligned}$$

sendo $Z_\beta = \tilde{X}(\tilde{X}^\top \Lambda \tilde{X})^{-1} \tilde{X}^\top$, $Z_{\beta d} = \text{diag}\{z_{\beta_{11}}, \dots, z_{\beta_{nn}}\}$, $Z_\gamma = \tilde{P}(\tilde{P}^\top V \tilde{P})^{-1} \tilde{P}^\top$ e $Z_{\gamma d} = \text{diag}\{z_{\gamma_{11}}, \dots, z_{\gamma_{nn}}\}$ matrizes de dimensão $n \times n$ positivas semi-definidas de posto completo. Denotamos $Z_\gamma^{(3)} = Z_\gamma^{(2)} \odot Z_\gamma$, $Z_\gamma^{(2)} = Z_\gamma \odot Z_\gamma$, $Z_{\gamma d}^{(2)} = Z_{\gamma d} \odot Z_{\gamma d}$, $Z_\beta^{(2)} = Z_\beta \odot Z_\beta$, em que $\odot$ denota o produto de Hadamard (Rao, 1973, p. 30), isto é, produto elemento a elemento. Ainda, Q_1 , Q_2 , Q_3 , Q_4 , Q_5 e Q_6 são matrizes diagonais de dimensão $n \times n$ cujos elementos são dados, respectivamente, por

$$\begin{aligned}
q_{1i} &= \frac{\delta_{(0,0,0,1,2)}}{16} \Lambda_{7i} + \frac{7\delta_{(0,0,1,0,1)}}{16} \Lambda_{7i} - \frac{\delta_{(0,0,1,0,1)}}{8} \Lambda_{8i} + \frac{\delta_{(0,1,0,0,0)}}{2} \Lambda_{7i} \\
&- \frac{\delta_{(0,1,0,0,0)}}{4} \Lambda_{8i}, \\
q_{2i} &= \frac{\delta_{(0,0,0,1,2)}}{16} \Lambda_{7i} - \frac{9\delta_{(0,0,1,0,1)}}{16} \Lambda_{7i} + \frac{3\delta_{(0,0,1,0,1)}}{8} \Lambda_{8i} + \frac{3\delta_{(0,1,0,0,0)}}{4} \Lambda_{8i} \\
&- \frac{3\delta_{(0,1,0,0,0)}}{2} \Lambda_{8i}, \\
q_{3i} &= -\frac{1}{4} \Lambda_{13i} - \frac{3}{16} \Lambda_{12i} + \frac{21}{16} \Lambda_{14i} - \frac{57}{64} \Lambda_{9i} + \frac{\delta_{(0,0,0,1,1)}}{64} \Lambda_{9i} + \frac{9\delta_{(0,0,1,0,3)}}{32} \Lambda_{9i} \\
&- \frac{3\delta_{(0,0,1,0,1)}}{16} \Lambda_{14i} + \frac{87\delta_{(0,1,0,0,2)}}{64} \Lambda_{9i} - \frac{27\delta_{(0,1,0,0,2)}}{16} \Lambda_{14i} + \frac{\delta_{(0,1,0,0,2)}}{4} \Lambda_{13i} \\
&+ \frac{3\delta_{(0,1,0,0,2)}}{16} \Lambda_{12i},
\end{aligned}$$

$$\begin{aligned}
q_{4i} &= -\frac{3\delta_{(0,1,0,0,2)}}{4}\Lambda_{5i} - \frac{3\delta_{(0,1,0,0,2)}}{4}\Lambda_{13i} + \frac{3\delta_{(0,1,0,0,2)}}{2}\Lambda_{14i} + \frac{3}{4}\Lambda_{12i} \\
&+ \frac{3}{4}\Lambda_{13i} - \frac{3}{2}\Lambda_{14i}, \\
q_{5i} &= -\frac{1}{8}(7 - \delta_{(0,0,1,0,3)} - 9\delta_{(0,1,0,0,2)}) \{3\Lambda_{14i} - 3\Lambda_{9i}\} \text{ e} \\
q_{6i} &= \frac{\delta_{(0,1,0,0,2)} - 1}{2} \{\Lambda_{12i} + \Lambda_{13i} - 5\Lambda_{14i} + 3\Lambda_{9i}\},
\end{aligned}$$

em que $\delta_{(a,b,c,d,e)} = E\{t^{(1)a}t^{(2)b}t^{(3)c}t^{(4)d}z_i^e\}$ para $a, b, c, d, e = 1, 2, 3$ e 4 , $t_{(z_i)}^{(k)} = \partial^k t / \partial z^k$, $k = 1, \dots, 4$. Os valores dos δ 's correspondentes a algumas distribuições são apresentados no Apêndice B. Além disso, $\Lambda_1 = \text{diag}\{\frac{h_1'^3}{\phi_1^3}, \dots, \frac{h_n'^3}{\phi_n^3}\}$, $\Lambda_2 = \text{diag}\{\frac{h_1'h_1''}{\phi_1^2}, \dots, \frac{h_n'h_n''}{\phi_n^2}\}$, $\Lambda_3 = \text{diag}\{\frac{h_1'}{\phi_1^2}, \dots, \frac{h_n'}{\phi_n^2}\}$, $\Lambda_5 = \text{diag}\{\frac{h_1'^2}{\phi_1^2}, \dots, \frac{h_n'^2}{\phi_n^2}\}$, $\Lambda_7 = \text{diag}\{\frac{h_1'^2}{\phi_1^3}, \dots, \frac{h_n'^2}{\phi_n^3}\}$, $\Lambda_8 = \text{diag}\{\frac{h_1''}{\phi_1^2}, \dots, \frac{h_n''}{\phi_n^2}\}$, $\Lambda_9 = \text{diag}\{\frac{h_1'^4}{\phi_1^4}, \dots, \frac{h_n'^4}{\phi_n^4}\}$, $\Lambda_{12} = \text{diag}\{\frac{h_1''^2}{\phi_1^2}, \dots, \frac{h_n''^2}{\phi_n^2}\}$, $\Lambda_{13} = \text{diag}\{\frac{h_1'h_1'''}{\phi_1^2}, \dots, \frac{h_n'h_n'''}{\phi_n^2}\}$ e $\Lambda_{14} = \text{diag}\{\frac{h_1'^2 h_1''}{\phi_1^3}, \dots, \frac{h_n'^2 h_n''}{\phi_n^3}\}$. As quantidades N 's são dadas por:

$$\begin{aligned}
N_1 &= -\frac{9}{2}((\delta_{(0,1,0,0,2)})^2 - 2\delta_{(0,1,0,0,2)} + 1), \\
N_2 &= \frac{1}{64}(15(\delta_{(0,1,0,0,2)})^2 - \delta_{(0,1,0,0,2)}\delta_{(0,0,1,0,3)} - 32\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)} + 17), \\
N_3 &= -\frac{3}{64}(16\delta_{(0,1,0,0,2)} - 9(\delta_{(0,1,0,0,2)})^2 + \delta_{(0,0,1,0,3)} - \delta_{(0,1,0,0,2)}\delta_{(0,0,1,0,3)} - 7), \\
N_4 &= \frac{1}{384}((\delta_{(0,0,1,0,3)})^2 - 135(\delta_{(0,1,0,0,2)})^2 - 6\delta_{(0,1,0,0,2)}\delta_{(0,0,1,0,3)} \\
&+ 258\delta_{(0,1,0,0,2)} + 10\delta_{(0,0,1,0,3)} - 119), \\
N_5 &= -\frac{15}{64}((\delta_{(0,1,0,0,2)})^2 - 2\delta_{(0,1,0,0,2)} + 1), \\
N_6 &= \frac{3}{128}(7(\delta_{(0,1,0,0,2)})^2 - \delta_{(0,1,0,0,2)}\delta_{(0,0,1,0,3)} - 16\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)} + 9), \\
N_7 &= -\frac{5}{128}(16\delta_{(0,1,0,0,2)} - 9(\delta_{(0,1,0,0,2)})^2 + \delta_{(0,0,1,0,3)} - \delta_{(0,1,0,0,2)}\delta_{(0,0,1,0,3)} - 7), \\
N_8 &= \frac{1}{256}((\delta_{(0,0,1,0,3)})^2 - 63(\delta_{(0,1,0,0,2)})^2 + 2\delta_{(0,0,1,0,3)} + 130\delta_{(0,1,0,0,2)} \\
&+ 2\delta_{(0,1,0,0,2)}\delta_{(0,0,1,0,3)} - 63), \\
N_9 &= \frac{1}{4}(\delta_{(0,1,0,0,2)} - 1)^2, \\
N_{10} &= -\frac{3}{2}(\delta_{(0,1,0,0,2)} - 1)\delta_{(0,0,1,0,1)},
\end{aligned}$$

$$\begin{aligned}
N_{11} &= -\frac{3}{16}(\delta_{(0,1,0,0,2)} - 1)\delta_{(0,1,0,0,0)}, \\
N_{12} &= -\frac{1}{64}\delta_{(0,0,1,0,1)}(7 - \delta_{(0,0,1,0,3)} - 9\delta_{(0,1,0,0,2)}), \\
N_{13} &= -\frac{1}{32}\delta_{(0,1,0,0,0)}(7 - \delta_{(0,0,1,0,3)} - 9\delta_{(0,1,0,0,2)}), \\
N_{14} &= \frac{5}{32}\delta_{(0,0,1,0,1)}(\delta_{(0,1,0,0,2)} - 1), \\
N_{15} &= -\frac{\delta_{(0,0,1,0,1)}}{64}(7\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)} - 9), \\
N_{16} &= \frac{5}{16}\delta_{(0,1,0,0,0)}(\delta_{(0,1,0,0,2)} - 1), \\
N_{17} &= -\frac{1}{32}\delta_{(0,1,0,0,0)}(7\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)} - 9), \\
N_{18} &= \left(\frac{N_{10} + N_{11}}{\delta_{(0,1,0,0,0)}} \right), \quad N_{19} = \left(\frac{N_{12} + N_{13}}{\delta_{(0,1,0,0,0)}} \right), \\
N_{20} &= -\left(\frac{N_{14} + N_{16}}{\delta_{(0,1,0,0,0)}} \right), \quad N_{21} = -\left(\frac{N_{15} + N_{17}}{\delta_{(0,1,0,0,0)}} \right), \\
N_{22} &= \left(\frac{1}{24} \left(\frac{\delta_{(0,0,1,0,1)}}{\delta_{(0,1,0,0,0)}} \right)^2 + \frac{\delta_{(0,0,1,0,1)}}{6\delta_{(0,1,0,0,0)}} + \frac{1}{6} \right), \\
N_{23} &= \left(\frac{1}{16} \left(\frac{\delta_{(0,0,1,0,1)}}{\delta_{(0,1,0,0,0)}} \right)^2 + \frac{\delta_{(0,0,1,0,1)}}{4\delta_{(0,1,0,0,0)}} + \frac{1}{4} \right) e \\
N_{24} &= \left(\frac{1}{24} \left(\frac{\delta_{(0,0,1,0,1)}}{\delta_{(0,1,0,0,0)}} \right)^2 - \frac{\delta_{(0,0,1,0,1)}}{3\delta_{(0,1,0,0,0)}} - \frac{5}{6} \right).
\end{aligned}$$

A estatística da razão de verossimilhanças perfiladas modificadas, definida por

$$LR_m = 2\{l_{CR}(\hat{\gamma}) - l_{CR}(\gamma^{(0)})\},$$

tem, sob certas condições de regularidade e sob H_0 , distribuição assintótica χ_q^2 . Mesmo com a modificação, o erro de aproximação da distribuição da estatística LR_m por χ_q^2 é de ordem $O(n^{-1})$. Visando melhorar tal aproximação, DiCiccio e Stern (1994) obtiveram uma correção de Bartlett para a estatística LR_m . Incorporando o fator de correção proposto por DiCiccio e Stern (1994) à estatística LR_m , deriva-se a estatística da razão de verossimilhanças

perfiladas modificadas corrigida (LR_m^*), expressa por

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

sendo c_m uma constante de ordem $O(n^{-1})$ tal que, sob H_0 , $E(LR_m^*) = q + O(n^{-3/2})$. Sob condições gerais de regularidade e sob H_0 , LR_m^* tem distribuição χ_q^2 com erro de ordem $O(n^{-2})$. Dessa forma, obtém-se uma estatística da razão de verossimilhanças corrigida baseada na verossimilhança perfilada modificada cuja distribuição tem uma melhor aproximação da distribuição χ^2 de referência do que a estatística LR_m .

A fim de representar a constante c_m do fator de correção para a estatística LR_m proposto por DiCiccio e Stern (1994), introduziremos as notações a seguir: os índices $a, b, \dots$ indexam os q parâmetros de interesse, os índices $i, j, \dots$ indexam os p parâmetros de perturbação e os índices $r, s, t, \dots$ variam sobre os $p + q$ parâmetros do modelo. Os cumulantes conjuntos das derivadas do logaritmo da função de verossimilhança são definidos como $\lambda_{rs} = E(\partial^2 l / \partial \theta_r \partial \theta_s)$, $\lambda_{rst} = E(\partial^3 l / \partial \theta_r \partial \theta_s \partial \theta_t)$, *etc.* Denotamos as derivadas dos cumulantes com respeito aos componentes do vetor $\boldsymbol{\theta}$ por $(\lambda_{rs})_t = \partial \lambda_{rs} / \partial \theta_t$, $(\lambda_{rst})_u = \partial \lambda_{rst} / \partial \theta_u$, *etc.* A matriz de informação de Fisher é composta pelos elementos $-\lambda_{rs}$, onde $-\lambda^{rs}$ são os correspondentes elementos da sua inversa. Definimos ainda $\tau^{rs} = \lambda^{ra} \lambda^{sb} \sigma_{ab}$, onde σ_{ab} é a inversa da matriz λ^{ab} e $\nu^{rs} = \lambda^{rs} - \tau^{rs}$.

DiCiccio e Stern (1994, equação 25) definiram uma equação geral para c_m dada por

$$\begin{aligned} c_m &= \frac{1}{4} \tau^{ru} \tau^{st} \lambda_{rstu} - \lambda^{rs} \tau^{st} (\lambda_{rst})_u + (\lambda^{ru} \lambda^{st} - \nu^{ru} \nu^{st}) (\lambda_{rs})_{tu} \\ &- \left(\frac{1}{4} \lambda^{ru} \tau^{st} \tau^{vw} + \frac{1}{2} \lambda^{ru} \tau^{sw} \tau^{tv} - \frac{1}{3} \tau^{ru} \tau^{sw} \tau^{tv} \right) \lambda_{rst} \lambda_{uvw} \\ &+ (\lambda^{ru} \tau^{st} \tau^{vw} + \lambda^{ru} \lambda^{sw} \lambda^{tu} - \nu^{ru} \lambda^{sw} \nu^{tv}) \lambda_{rst} (\lambda_{uv})_w \\ &- (\lambda^{ru} \lambda^{st} \lambda^{vw} - \nu^{ru} \nu^{st} \nu^{vw} + \lambda^{ru} \lambda^{sw} \lambda^{tv} - \nu^{ru} \nu^{sw} \nu^{tv}) (\lambda_{rs})_t (\lambda_{uv})_w. \end{aligned}$$

Tipicamente, c_m depende de parâmetros de perturbação desconhecidos,

os quais devem ser substituídos por estimativas consistentes, mas isto não afeta a ordem da aproximação resultante.

Para a obtenção de um fator de correção de Bartlett proposto por DiCiccio e Stern (1994) para a estatística LR_m do teste de H_0 contra H_1 na classe dos MNLSH, temos que, por definição, $\nu^{ab} = \nu^{ai} = \nu^{ia} = 0$, $\tau^{ij} = \tau^{ai} = \tau^{ia} = 0$, $\nu^{ij} = \lambda^{ij}$ e $\tau^{ab} = \lambda^{ab}$, uma vez que os parâmetros do modelo em estudo, β e γ , são globalmente ortogonais. Após alguma álgebra, chegamos à seguinte expressão para o fator de correção c_m :

$$\begin{aligned}
c_m = & \frac{1}{4} \lambda^{ab} \lambda^{cd} \lambda_{abcd} - \lambda^{ab} \lambda^{cd} (\lambda_{acd})_b + \lambda^{ab} \lambda^{cd} (\lambda_{ac})_{db} - \lambda^{ij} \lambda^{ab} (\lambda_{iab})_j \\
& + (\lambda^{ij} \lambda^{ab} \lambda^{kl}) \lambda_{iab} (\lambda_{jk})_l - \left(\frac{1}{4} \lambda^{ij} \lambda^{ab} \lambda^{cd} + \frac{1}{2} \lambda^{ij} \lambda^{ad} \lambda^{bc} \right) \lambda_{iab} \lambda_{jcd} \\
& - \left(\frac{1}{4} \lambda^{ab} \lambda^{cd} \lambda^{ef} + \frac{1}{6} \lambda^{ab} \lambda^{cf} \lambda^{de} \right) \lambda_{acd} \lambda_{bef} + (\lambda^{ab} \lambda^{cd} \lambda^{ef} + \lambda^{ab} \lambda^{cf} \lambda^{de}) \lambda_{acd} (\lambda_{be})_f \\
& - (\lambda^{ab} \lambda^{cd} \lambda^{ef} + \lambda^{ab} \lambda^{cf} \lambda^{de}) (\lambda_{ac})_d (\lambda_{be})_f + (\lambda^{ab} \lambda^{ij} \lambda^{cd}) \lambda_{aic} (\lambda_{bd})_j. \tag{3.3}
\end{aligned}$$

É importante observar que a expressão (3.3) pode ser utilizada para obter expressões em forma fechada para um fator de correção de Bartlett para a estatística LR_m não somente na classe dos MNLSH, e sim para qualquer classe de modelos que siga a mesma partição de parâmetros dos MNLSH e que os vetores β e γ sejam globalmente ortogonais.

Para expressar as quantidades necessárias ao cálculo de (3.3), utilizaremos as notações definidas nas Seções 2.3 e 3.2. Adicionalmente, definimos as matrizes diagonais M_1 , M_2 , M_3 , M_4 e M_5 de dimensão $n \times n$, cujos elementos são dados, respectivamente, por:

$$\begin{aligned}
m_{1i} &= \frac{1}{16} \left(\frac{-16h'_i h''_i}{\phi_i^2} - \frac{12h_i''^2}{\phi_i^2} + \frac{84h_i'^2 h''_i}{\phi_i^3} - \frac{57h_i'^4}{\phi_i^4} \right), \\
&+ \frac{\delta_{(0,1,0,0,2)}}{16} \left(\frac{87h_i'^4}{\phi_i^4} - \frac{108h_i'^2 h''_i}{\phi_i^3} + \frac{12h_i'^2}{\phi_i^2} + \frac{16h'_i h''_i}{\phi_i^2} \right), \\
&+ \frac{\delta_{(0,0,1,0,3)}}{8} \left(\frac{9h_i'^4}{\phi_i^4} - \frac{6h_i'^2 h''_i}{\phi_i^3} \right) + \frac{\delta_{(0,0,0,1,4)}}{16} \frac{h_i'^4}{\phi_i^4}, \\
m_{2i} &= \frac{3}{4} (\delta_{(0,1,0,0,2)} - 1) \left(\frac{h_i''^2}{\phi_i^2} + \frac{h'_i h_i'''}{\phi_i^2} - \frac{2h_i'^2 h''_i}{\phi_i^3} \right), \\
&+ (7 - \delta_{(0,0,1,0,3)} - 9\delta_{(0,1,0,0,2)}) \left(\frac{3h_i'^2 h''_i}{\phi_i^3} - \frac{3h_i'^4}{\phi_i^4} \right), \\
m_{3i} &= \left(\frac{\delta_{(0,1,0,0,2)} - 1}{2} \right) \left(\frac{h_i''^2 + h'_i h_i'''}{\phi_i^2} - \frac{5h_i'^2 h''_i}{\phi_i^3} + \frac{3h_i'^4}{\phi_i^4} \right), \\
m_{4i} &= \frac{3}{4} (\delta_{(0,1,0,0,2)} - 1) \frac{h'_i h''_i}{\phi_i^2} + \frac{1}{8} (7 - \delta_{(0,0,1,0,3)} - 9\delta_{(0,1,0,0,2)}) \frac{h_i'^3}{\phi_i^3} \text{ e} \\
m_{5i} &= \left(\frac{\delta_{(0,1,0,0,2)} - 1}{2} \right) \left(\frac{h'_i h''_i}{\phi_i^2} - \frac{h_i'^3}{\phi_i^3} \right).
\end{aligned}$$

Os cumulantes do logaritmo da função de verossimilhança dos MNLSH e suas derivadas encontram-se no Apêndice A. Após algumas manipulações algébricas, obtemos:

$$\begin{aligned}
\lambda_{abcd} &= \sum_{i=1}^n m_{1i}(a, b, c, d)_i, \quad (\lambda_{acd})_b = \sum_{i=1}^n m_{2i}(a, b, c, d)_i, \\
(\lambda_{ac})_{db} &= \sum_{i=1}^n m_{3i}, \quad (\lambda_{bd})_j = (\lambda_{iab})_j = \lambda_{aic} = 0, \\
(\lambda_{jk})_l &= \delta_{(0,1,0,0,0)} \sum_{i=1}^n \frac{1}{\phi_i} \{(jk, l) + (j, kl)\}_i, \\
\lambda_{acd} &= \sum_{i=1}^n m_{4i}(a, c, d)_i, \quad (\lambda_{ac})_d = \sum_{i=1}^n m_{5i}(a, c, d)_i.
\end{aligned}$$

Sendo $(\lambda_{bd})_j = (\lambda_{iab})_j = \lambda_{aic} = 0$, podemos reescrever (3.3) como:

$$\begin{aligned}
c_m &= \lambda^{ab}\lambda^{cd} \left(\frac{1}{4}\lambda_{abcd} - (\lambda_{acd})_b + (\lambda_{ac})_{db} \right) \\
&- \lambda^{ab}\lambda^{cd}\lambda^{ef} \left(\frac{1}{4}\lambda_{acd}\lambda_{bef} - \lambda_{acd}(\lambda_{be})_f + (\lambda_{ac})_d(\lambda_{be})_f \right) \\
&- \lambda^{ab}\lambda^{cf}\lambda^{ed} \left(\frac{1}{6}\lambda_{acd}\lambda_{bef} - \lambda_{acd}(\lambda_{be})_f + (\lambda_{ac})_d(\lambda_{be})_f \right). \quad (3.4)
\end{aligned}$$

Substituindo os λ 's obtidos na expressão (3.4) e rearranjando os termos, obtemos

$$\begin{aligned}
c_m &= \frac{1}{4} \sum_{i=1}^n (a)_i \lambda^{ab}(b)_i m_{1i}(c)_i \lambda^{cd}(d)_i - \sum_{i=1}^n (a)_i \lambda^{ab}(b)_i m_{2i}(c)_i \lambda^{cd}(d)_i \\
&+ \sum_{i=1}^n (a)_i \lambda^{ab}(b)_i m_{3i}(c)_i \lambda^{cd}(d)_i \\
&- \frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n (c)_i \lambda^{cd}(d)_i m_{4i}(a)_i \lambda^{ab}(b)_j m_{4j}(e)_j \lambda^{ab}(f)_j \\
&+ \sum_{i=1}^n \sum_{j=1}^n (c)_i \lambda^{cd}(d)_i m_{4i}(a)_i \lambda^{ab}(b)_j m_{5j}(e)_j \lambda^{ab}(f)_j \\
&- \sum_{i=1}^n \sum_{j=1}^n (c)_i \lambda^{cd}(d)_i m_{5i}(a)_i \lambda^{ab}(b)_j m_{5j}(e)_j \lambda^{ab}(f)_j \\
&- \frac{1}{6} \sum_{i=1}^n \sum_{j=1}^n m_{4i}(a)_i \lambda^{ab}(b)_j (c)_i \lambda^{cf}(f)_j (d)_i \lambda^{de}(e)_j m_{4j} \\
&+ \sum_{i=1}^n \sum_{j=1}^n m_{4i}(a)_i \lambda^{ab}(b)_j (c)_i \lambda^{cf}(f)_j (d)_i \lambda^{de}(e)_j m_{5j} \\
&- \sum_{i=1}^n \sum_{j=1}^n m_{5i}(a)_i \lambda^{ab}(b)_j (c)_i \lambda^{cf}(f)_j (d)_i \lambda^{de}(e)_j m_{5j} \quad (3.5)
\end{aligned}$$

Considerando o caso em que $\phi_i = \exp\{\tau_i\}$, com $\tau_i = \mathbf{z}_i^\top \boldsymbol{\gamma}$, isto é, o caso de heteroscedasticidade multiplicativa, temos que $\partial\phi_i/\partial\tau_i = \phi_i$. Por definição, $\partial\phi_i/\partial\tau_i = h'_i$, logo, $\phi_i = h'_i$. Desse modo, $h''_i = \partial h'_i/\partial\tau_i = h'_i$. Prodecendo desta forma para as derivadas de terceira e quarta ordens com relação a τ_i ,

temos que $\phi_i = h'_i = h''_i = h'''_i = h''''_i$. Sendo assim, observamos que $m_{2i} = m_{3i} = m_{5i} = 0$, $\forall i = 1, \dots, n$. Definindo $H = \{h_{ij}\} = -\mathbf{Z}[\mathbf{Z}^\top V \mathbf{Z}]^{-1} \mathbf{Z}^\top$, com $\mathbf{Z} = (z_1, \dots, z_n)^\top$ e V como definido na Seção 2.3. Assim, podemos observar que $(a)_i \lambda^{ab} (b)_j = h_{ij}$. Substituindo estes resultados em (3.5), obtemos a seguinte expressão para c_m :

$$\begin{aligned} c_m &= \frac{1}{4} \sum_{i=1}^n h_{ii} m_{1i} h_{ii} + \frac{1}{4} \sum_{i=1}^n \sum_{j=1}^n h_{ii} m_{4i} h_{ij} m_{4j} h_{jj} \\ &+ \frac{1}{6} \sum_{i=1}^n \sum_{j=1}^n m_{4i} h_{ij} h_{ij} h_{ij} m_{4j}, \end{aligned}$$

em notação matricial, temos

$$c_m = \frac{1}{4} \text{tr}(M_1 H_d^{(2)}) + \frac{1}{4} \mathbf{1}^\top H_d M_4 H M_4 H_d \mathbf{1} + \frac{1}{6} \mathbf{1}^\top M_4 H^{(3)} \mathbf{1}, \quad (3.6)$$

onde $H_d = \text{diag}\{h_{11}, \dots, h_{nn}\}$, $H_d^{(2)} = \text{diag}\{h_{11}^2, \dots, h_{nn}^2\}$ e $H^{(3)} = (h_{ij})^3$.

Como podemos observar, os fatores de correção c e c_m , expressos, respectivamente, pelas equações (3.2) e (3.6), envolvem apenas operações simples de matrizes e podem ser facilmente implementados em pacotes de computação simbólica e linguagens de programação que permitam executar operações simples de álgebra linear, tais como **Ox**, **R**, **S-Plus**, **MAPLE**, etc. Além disso, temos que c_m envolve apenas a matriz de covariadas $\mathbf{Z}$ usadas para modelar a variância, do número de parâmetros desconhecidos que definem o comportamento da heteroscedasticidade e do número de observações. É válido salientar, ainda, que o fator de correção c_m não depende de parâmetros de perturbação, tampouco depende de parâmetros desconhecidos.

CAPÍTULO 4

Resultados Numéricos

Neste capítulo, desenvolvemos um estudo de simulação de Monte Carlo a fim de comparar o desempenho dos testes baseados nas seguintes estatísticas, a saber: estatística da razão de verossimilhanças usual, LR , sua versão corrigida via Bartlett, LR^* , estatística da razão de verossimilhanças perfiladas modificadas, LR_m , e sua versão corrigida via DiCiccio e Stern, LR_m^* . Os desempenhos dos testes são avaliados em função da proximidade das probabilidades de rejeição da hipótese nula, dado esta verdadeira (probabilidade do erro tipo I), aos respectivos níveis nominais dos testes. Também avaliamos os poderes dos testes em estudo sob algumas situações e a discrepância relativa de quantis.

As simulações baseiam-se no modelo de regressão

$$y_i = \beta_0 + \exp\{\beta_1 x_{i1}\} + \sum_{j=2}^{p-1} \beta_j x_{ij} + \epsilon_i, \quad i = 1, \dots, n,$$

onde ϵ_i são variáveis aleatórias independentes tais que $\epsilon_i \sim S(\mu_i, \phi_i, g)$, isto é, ϵ_i tem distribuição pertencente à classe das distribuições simétricas, sendo μ_i e ϕ_i definidos por (2.2) e (2.3), respectivamente. As hipóteses testadas são

$H_0 : \gamma = 0 \times H_1 : \gamma \neq 0$, sendo γ q -dimensional, indicando que o modelo é homoscedástico.

O estudo de simulação foi baseado em duas distribuições simétricas, a saber: t -Student com 4 graus de liberdade e exponencial potência com parâmetro de forma $k = 0, 3$. Para as simulações variando o tamanho amostral para o modelo com número de parâmetros de perturbação e parâmetros de interesse fixo variamos o valor dos graus de liberdade (ν) do modelo t -Student. Consideramos $\nu = 4, 10$ e 30 a fim de avaliar o desempenho dos testes para modelos próximos ao modelo normal, uma vez que conforme aumentamos os graus de liberdade da distribuição t -Student, esta se aproxima da distribuição normal. A escolha dos graus de liberdade para o modelo t -Student foi baseada em Lange *et al.* (1989), que observaram, a partir de diversas aplicações, que para pequenas amostras o modelo t -Student com $\nu = 4$ apresentou bom desempenho. Para o modelo Exponencial potência foi adotado o parâmetro de forma $k = 0, 3$ pois era do interesse do trabalho realizar o estudo de simulação para distribuições com caudas mais pesadas que as da distribuição normal ($k = 0$). Assumimos ainda outros valores para o parâmetro k , a saber: $k = 0, 4$ e $k = 0, 5$, no mais, conforme o parâmetro de forma assumia valores mais elevados, as taxas de falha de convergência aumentavam consideravelmente, desse modo, comprometendo as estimativas dos parâmetros. As variáveis respostas foram geradas assumindo os parâmetros de regressão como $\beta_0 = \dots = \beta_{p-1} = 1$, $\gamma_1 = \gamma_4 = \gamma_5 = 0, 1$; $\gamma_2 = 0, 3$ e $\gamma_3 = 0, 5$. As covariáveis $x_1, \dots, x_{p-1}$ e $z_1, \dots, z_q$ foram geradas como amostras aleatórias da distribuição $U(0, 1)$. O número de réplicas de Monte Carlo usadas foi de 10.000 e os níveis nominais considerados foram $\alpha = 0, 5\%, 1\%, 5\%$ e 10% . As simulações foram realizadas a partir da linguagem de programação matricial **0x** (Doornik, 2001). Como os estimadores de máxima verossimilhança de β restrito a γ e de γ não possuem forma fechada, suas respectivas estimativas foram obtidas a partir do algoritmo SQP, através da função **MaxSQP** da linguagem matricial de programação **0x**. Este comando maximiza funções não lineares sujeitas a restrições a partir de uma técnica de programação quadrá-

tica sequencial. Para maiores detalhes sobre o método SQP, ver Nosedal e Wright (1999).

Para cada tamanho amostral e nível nominal considerado, calculamos as taxas de rejeição de cada teste, isto é, estimamos via simulação de Monte Carlo $P(LR \geq x_{(\alpha;q)})$, $P(LR^* \geq x_{(\alpha;q)})$, $P(LR_m \geq x_{(\alpha;q)})$ e $P(LR_m^* \geq x_{(\alpha;q)})$, em que $x_{(\alpha;q)}$ é o percentil $(1 - \alpha)$ da distribuição χ_q^2 . Todas as entradas das tabelas apresentadas são porcentagens.

Na Tabela 4.1 apresentamos as taxas de rejeição dos testes para o modelo t -Student e Exponencial potência com $\nu = 4, 10$ e 30 , $q = 3$, diferentes tamanhos amostrais, a saber, $n = 20, 30, 40$ e 50 e $p = 3$. Podemos observar que para amostras de tamanho pequeno o teste usual é bastante liberal, isto é, rejeita a hipótese nula além do esperado. Em contrapartida, o teste baseado na estatística da razão de verossimilhanças perfiladas modificadas (LR_m) apresenta taxas de rejeição mais próximas do nível nominal considerado. Por exemplo, quando $n = 20$ e $\alpha = 10\%$, temos que as taxas de rejeição dos testes baseados em LR e LR_m são, respectivamente, $27,9\%$ e $11,3\%$, para o modelo t -Student, e $26,5\%$ e $8,4\%$, para o modelo Exponencial potência. Aplicando a correção de Bartlett a ambos os testes, notamos que o teste baseado na estatística LR^* ainda apresenta tendência liberal, enquanto o teste baseado em LR_m^* apresenta taxas de rejeição bem mais próximas dos níveis nominais considerados. A exemplo, ainda para $n = 20$ e $\alpha = 10\%$, temos que o tamanho dos testes baseados em LR^* e LR_m^* são, respectivamente, 13% e $10,2\%$, para o modelo t -Student, e $14,7\%$ e $8,0\%$, para o modelo Exponencial potência. Conforme aumenta o tamanho amostral, as taxas de rejeição de todos os quatro testes para o modelo t -Student ficam mais próximas dos níveis nominais, como era de se esperar. Para o modelo Exponencial potência, conforme aumenta-se o tamanho amostral, os testes baseados nas estatísticas LR_m e LR_m^* apresentam-se, em geral, ligeiramente conservativos. Como ilustração, para $n = 30$ e $\alpha = 5\%$, as taxas de rejeição dos testes baseados em LR , LR^* , LR_m e LR_m^* são iguais a $12,7\%$, $5,9\%$, $5,8\%$ e $5,3\%$, para o modelo t -Student, e $11,0\%$, $6,2\%$, $4,7\%$ e $4,5\%$, para o modelo

Exponencial potência, e quando aumentamos n para 50, na mesma situação, temos que as taxas de rejeição dos quatro testes supracitados são, respectivamente, de 8,1%, 4,8%, 5,3% e 5,2%, para o modelo t -Student, e 7,7%, 5,4%, 5,0% e 4,9%, para o modelo Exponencial potência. Observamos também que mesmo para um número de amostra grande, o teste usual apresentou-se liberal, por exemplo, considerando $\alpha = 10\%$ e $n = 50$, temos que o tamanho do teste usual é de 14,8%, para o modelo t -Student, e 13,9% para o modelo Exponencial potência. Notamos ainda que conforme aumenta o tamanho amostral, o impacto da correção de Bartlett sobre a estatística LR_m é menor, por exemplo, para $n = 30$, e $\alpha = 1\%$, temos que a taxa de rejeição dos testes baseados em LR_m e LR_m^* são, respectivamente, iguais a 1,4% e 1,2%, para o modelo t -Student, e 0,9% e 0,8%, para o modelo Exponencial potência, enquanto que para $n = 50$, na mesma situação, as respectivas taxas são de 1,0% e 0,9%, para o modelo t -Student, e 0,9% e 0,8% para o modelo Exponencial potência. Observa-se também que a maioria dos testes baseados na estatística LR_m e LR_m^* para o modelo Exponencial potência apresentam-se conservativos para diversos tamanhos amostrais considerados, isto é, rejeitam H_0 aquém do esperado.

Conforme aumentamos os graus de liberdade do modelo t -Student, os testes baseados na estatística LR para pequenas amostras apresentam, de forma geral, taxas de rejeição mais próximas dos níveis nominais adotados, como exemplo, quando $n = 20$ e $\alpha = 10\%$, temos que as taxas de rejeição dos testes baseados em LR para o modelo t -Student com $\nu = 4, 10$ e 30 foram, respectivamente, iguais a 27,9%, 27,7% e 27,1%. A medida que o tamanho amostral aumenta, podemos observar que as taxas de rejeição assumem valores bastante próximos, a exemplo, considerando $n = 50$ e $\alpha = 10\%$, temos que as taxas de rejeição são iguais a 14,8%, 14,7% e 15% para $\nu = 4, 10$ e 30, respectivamente. Os testes baseados na estatística LR^* corrigem a tendência liberal dos testes baseado na estatística LR e conforme aumentamos os graus de liberdade, temos, para pequenas amostras, que os testes baseados em LR^* apresentam, em geral, taxas de rejeição mais distantes do nível nomi-

nal considerado. Os testes baseados nas estatísticas LR_m e LR_m^* apresentam taxas de rejeição bastante próximas quando $\nu = 4$ e 10, tanto para amostras de tamanho pequeno quanto para tamanhos amostrais grandes. Quando $\nu = 30$, os testes baseados em LR_m e LR_m^* para pequenas amostras apresentam, de forma geral, caráter liberal e conforme o tamanho amostral cresce, os testes baseados em LR_m e LR_m^* apresentam caráter conservativo, exceto para quando $n = 50$, quando os testes baseados nas estatísticas supracitadas apresentam, de forma geral, caráter liberal. Podemos observar ainda que, em geral, para o modelo Exponencial potência, os testes baseados em LR e LR^* são bastante liberais, enquanto que os testes baseados em LR_m e LR_m^* são conservativos, no mais, apresentam taxas de rejeição mais próximas dos níveis nominais adotados.

A fim de analisar a influência do número de parâmetros de perturbação no desempenho dos testes, fixamos o tamanho amostral em $n = 35$, o número de parâmetros de interesse em $q = 2$ e variamos o número de parâmetros de perturbação (p). Na Tabela 4.2, encontram-se os resultados das simulações para os cenários descritos acima para os modelos sob estudo, isto é, modelo t -Student com 4 graus de liberdade e modelo Exponencial potência com $k = 0,3$. Podemos observar que os testes baseados em LR são liberais e a medida que o número de parâmetros de perturbação aumenta, a distorção do tamanho dos testes aumenta consideravelmente, por exemplo, para o modelo t -Student quando $p = 5$ e $\alpha = 5\%$, temos que a taxa de rejeição do teste baseado em LR é 14%, ou seja, maior que o dobro do valor nominal considerado. Ainda para a mesma situação, considerando o modelo Exponencial potência, temos que a taxa de rejeição excede 13%. A correção de Bartlett aplicada ao teste usual tende a atenuar sua tendência liberal, de modo que o teste da razão de verossimilhanças corrigido apresenta menor distorção de tamanho do que o teste baseado em LR . Considerando a mesma situação exemplificada acima, temos que as taxas de rejeição do teste corrigido para os modelos t -Student e Exponencial potência são de 5,7% e 6,4%, respectivamente. Assim como o teste usual, podemos observar que a distorção do

tamanho dos testes corrigidos aumentam conforme o número de parâmetros de perturbação aumenta. O teste da razão de verossimilhanças perfiladas modificadas tende a corrigir a tendência liberal do teste usual, apresentando, em geral, taxas de rejeição ligeiramente superiores aos níveis nominais, para o modelo t -Student. Para o modelo Exponencial potência, temos que o teste baseado em LR_m é, em geral, ligeiramente conservador. Novamente para $p = 5$ e $\alpha = 5\%$, temos as taxas de rejeição do teste baseado em LR_m para os modelos t -Student e Exponencial potência são de 5,1% e 4,9%, respectivamente. A correção de Bartlett aplicada a este teste, de forma geral, aproxima ainda mais as taxas de rejeição aos tamanhos nominais do teste, para o modelo t -Student, enquanto que para o modelo Exponencial potência apresenta menor impacto, uma vez que, em geral, o teste baseado em LR_m para este modelo é ligeiramente conservativo. Ainda quando $p = 5$ e $\alpha = 5\%$, temos que a taxa de rejeição do teste baseado em LR_m^* é 5%, para o modelo t -Student, e 4,8%, para o modelo Exponencial potência.

Em suma, temos que, de forma geral, o teste usual apresenta tendência liberal e conforme o número de parâmetros de perturbação aumenta, a taxa de rejeição deste teste aumenta. A correção de Bartlett aplicada a este teste apresenta taxas de rejeição próximas aos níveis nominais, e assim como o teste usual, conforme aumenta-se o número de parâmetros de perturbação, as taxas de rejeição da hipótese nula deste teste aumenta. Em contrapartida, temos que os testes baseados em LR_m e LR_m^* não são tão sensíveis ao aumento do número de parâmetros de perturbação, isto é, as taxas de rejeição dos referidos testes permanecem estáveis conforme aumenta-se o número de parâmetros de perturbação. Para o modelo t -Student pode-se observar que entre os testes corrigidos aquele que apresentou, de forma geral, taxas de rejeição mais próximas dos níveis nominais adotados foi o teste baseado na estatística LR_m^* . Para o modelo Exponencial potência, o teste baseado na estatística LR^* apresentou caráter ligeiramente liberal, enquanto que o teste baseado na estatística LR_m^* apresentou-se ligeiramente conservativo.

Na Tabela 4.3 avaliamos a influência do número de parâmetros de inte-

resse nos modelos em estudo, isto é, apresentamos as taxas de rejeição dos testes sob investigação para os modelos em estudo quando $n = 35$, $p = 3$ e q variando entre 2 e 6. Notamos que os comportamentos dos quatros testes foram bastante similares aos observados quando aumentamos o número de parâmetros de perturbação, sendo o teste mais afetado com o aumento do número de parâmetros que define a heteroscedasticidade o teste da razão de verossimilhanças usual. Os demais testes apresentaram tamanhos mais próximos dos respectivos níveis nominais considerados, entretanto, o teste baseado em LR_m apresentou-se ligeiramente liberal conforme o número de parâmetros de interesse aumenta para o modelo t -Student. Para o modelo Exponencial potência, os testes baseados nas estatísticas LR_m e LR_m^* apresentam-se ligeiramente conservativos e a correção de Bartlett aplicada à estatística LR_m teve, em geral, menor impacto, isto é, as taxas de rejeição dos testes baseados nas estatísticas LR_m e LR_m^* apresentaram diferença pequena. As conclusões obtidas a partir da Tabela 4.2 também se verificam aqui. Como ilustração, considerando $q = 4$ e $\alpha = 1,0\%$, temos que as taxas de rejeição dos testes baseados nas estatísticas LR , LR^* , LR_m e LR_m^* são, respectivamente, 4,1%, 1,3%, 1,1% e 0,9%, para o modelo t -Student, e 3,3%, 1,4%, 0,9% e 0,9%, para o modelo Exponencial potência.

Nas Tabelas 4.4 e 4.5 são apresentados os resultados obtidos a partir da hipótese alternativa $H_1 : \gamma_1 = \gamma_2 = \gamma_3 \neq 0$, isto é, heteroscedasticidade, para o modelo t -Student e Exponencial potência, respectivamente, quando $n = 35$, $q = 3$, $p = 5$, $\alpha = 10\%$ e diferentes valores de $\gamma_1 = \gamma_2 = \gamma_3 = \gamma$, com γ variando entre 0,1 e 1,0. Tendo em vista a tendência liberal dos testes baseados nas estatísticas LR e LR^* , as simulações foram realizadas com os valores críticos estimados em vez dos valores tabelados para que todos os testes pudessem ter o mesmo tamanho. A partir destes resultados, observamos para todos os modelos considerados o poder dos testes baseados nas estatísticas LR_m e LR_m^* são semelhantes e superiores aos testes baseados em LR e LR^* . A exemplo, considerando o modelo t -Student quando $p = 5$ e $\delta = 0,5$, as estimativas dos poderes dos testes baseados em LR , LR^* , LR_m

e LR_m^* são, respectivamente, 49,2%, 49,4%, 79,7% e 79,7%. Considerando o mesmo cenário para o modelo Exponencial potência, temos que as estimativas dos poderes para os testes baseados em LR , LR^* , LR_m e LR_m^* são, respectivamente, 76,8%, 76,7%, 86,8% e 86,9%.

A fim de avaliar a aproximação assintótica dos testes baseados em LR , LR^* , LR_m e LR_m^* , apresentamos nas Figuras 4.1 e 4.2 os gráficos das discrepâncias relativas entre os quantis assintóticos e os quantis assintóticos das estatísticas dos quatro testes sob estudo contra os quantis assintóticos para os modelos t -Student e Exponencial potência, respectivamente. Definimos discrepância relativa como

$$\frac{ST(1 - \alpha) - \chi_3^2(1 - \alpha)}{\chi_3^2(1 - \alpha)},$$

onde ST denota a estatística do teste (LR , LR^* , LR_m ou LR_m^* , conforme o caso), $ST(1 - \alpha)$ denota o quantil amostral de ordem $(1 - \alpha)$ do conjunto de valores simulados da estatística de teste e $\chi_3^2(1 - \alpha)$ representa o correspondente quantil da distribuição χ_3^2 . Nesse contexto, quão mais próxima da ordenada nula estiver a curva, melhor é a aproximação assintótica utilizada no teste.

Analisando a Figura 4.1, constatamos que a curva de discrepância de quantis da estatística LR está bem distante da ordenada nula e conforme o valor do quantil assintótico aumenta, a discrepância relativa tende a se estabilizar em torno de 57%, ratificando a tendência liberal do teste baseado na estatística LR . A curva de discrepância relativa da estatística LR^* apresenta-se mais próxima da ordenada nula, se comparada com a curva da estatística LR . Notamos também que a curva de LR^* apresenta-se estável em torno de 11%, confirmando que o teste baseado nesta estatística é liberal. Podemos observar também que as curvas de discrepâncias de quantis das estatísticas LR_m e LR_m^* são bem mais próximas do eixo horizontal de ordenada nula e que conforme aumenta-se o valor dos quantis, tais discrepâncias alteram-se muito pouco, com as curvas fluando ao redor da ordenada nula.

Na Figura 4.2, podemos observar que a curva de discrepância de quantis

da estatística LR , assim como para o modelo t -Student, está bem distante e acima da ordenada nula. Diferentemente do observado na Figura 4.1, a discrepância de quantis da estatística LR aumenta ligeiramente conforme aumenta-se o valor do quantil assintótico. A curva de discrepância relativa da estatística LR^* apresenta-se estável, em torno de 14%, conforme aumenta-se o valor do quantil assintótico. As curvas das estatísticas LR e LR^* ratificam que os testes baseados nestas estatísticas são liberais. As curvas de discrepância de quantis das estatísticas LR_m e LR_m^* praticamente se sobrepõem e para valores de quantis entre 0 e 2 as curvas de discrepância relativa das estatísticas LR e LR^* apresentam um crescimento acelerado, enquanto que para valores de quantis acima de 2 essas curvas apresentam-se mais estáveis. Verificamos ainda que essas curvas estão abaixo da ordenada nula, confirmando que estes testes são conservativos.

Figura 4.1: Gráfico das discrepâncias relativas de quantis - Modelo *t*-Student com $\nu = 4$, $n = 35$, $p = 5$ e $q = 3$.

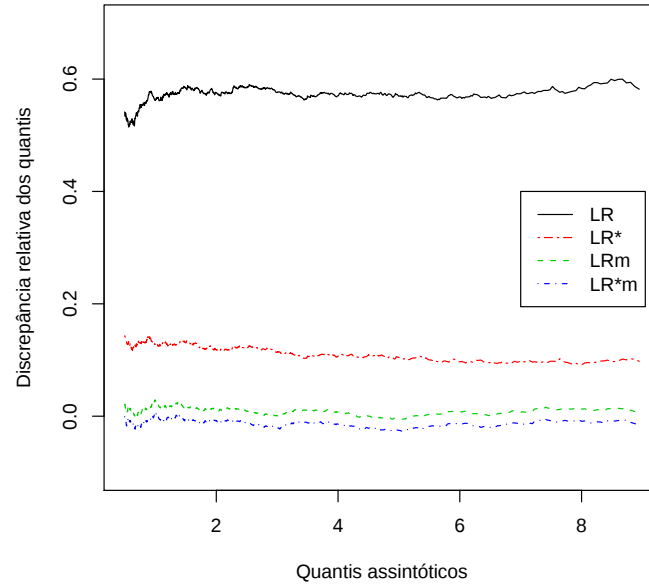


Figura 4.2: Gráfico das discrepâncias relativas de quantis - Modelo Exponencial potência com $k = 0, 3$; $n = 35$, $p = 5$ e $q = 3$.

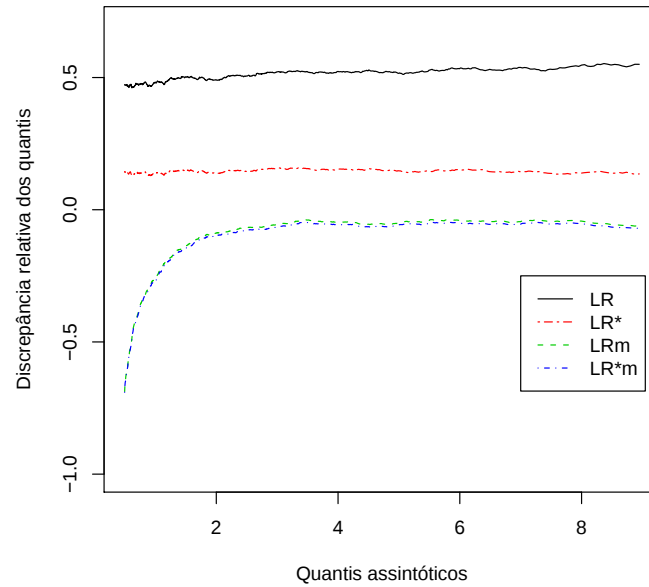


Tabela 4.1: Tamanho dos testes - Modelo t -Student com $\nu = 4, 10$ e 30 e Exponencial potência com $k = 0, 3$; $p = 3$, $q = 3$ e diversos valores para n .

n	$\alpha(\%)$	t -Student $\nu = 4$			t -Student $\nu = 10$			t -Student $\nu = 30$			Exponencial potência		
		LR	LR^*	LR_m	LR	LR^*	LR_m	LR	LR^*	LR_m	LR	LR^*	LR_m
20	0,5	5,1	1,0	0,6	4,6	0,7	0,6	3,9	0,8	0,6	4,7	1,1	0,3
	1,0	7,4	1,7	1,2	7,0	1,5	1,1	6,2	1,6	1,2	6,8	2,1	0,7
	5,0	18,5	6,9	5,9	18,8	7,2	6,2	17,7	7,2	5,7	17,6	8,3	4,0
	10,0	27,9	13,0	11,3	27,7	13,3	11,7	27,1	13,6	11,3	26,5	14,7	8,4
30	0,5	2,9	0,7	0,7	2,1	0,6	0,6	1,6	0,4	0,5	1,8	0,6	0,4
	1,0	4,2	1,3	1,4	3,6	1,0	1,2	2,8	0,9	0,9	3,1	1,3	0,8
	5,0	12,7	5,9	5,8	11,1	5,3	5,6	10,2	4,9	4,8	11,0	6,2	4,7
	10,0	20,0	11,6	11,1	18,6	10,4	11,2	17,7	10,1	9,9	18,4	11,7	9,5
40	0,5	2,2	0,6	0,6	1,5	0,4	0,5	1,2	0,5	0,6	1,5	0,6	0,5
	1,0	3,1	1,0	1,1	2,5	0,9	1,2	2,3	1,0	1,2	2,6	1,3	1,1
	5,0	9,7	4,9	5,1	8,7	5,0	5,3	9,0	4,9	5,0	9,0	5,8	4,6
	10,0	16,8	9,5	10,7	15,7	9,5	10,4	15,3	10,4	10,2	16,1	11,0	10,3
50	0,5	1,1	0,4	0,5	1,1	0,5	0,6	1,3	0,5	0,6	0,9	0,4	0,3
	1,0	2,2	0,9	1,0	1,9	0,9	0,9	2,3	1,2	1,2	2,0	0,9	0,8
	5,0	8,1	4,8	5,3	8,2	5,0	5,1	8,2	5,4	5,5	7,7	5,4	5,0
	10,0	14,8	9,9	10,6	14,7	10,0	10,4	15,0	10,4	10,8	13,9	10,2	9,9

Tabela 4.2: Tamanho dos testes para os Modelo t-Student com $\nu = 4$ e Exponencial potência com $k = 0, 3$; ambos com $n = 35$, $q = 2$ e diversos valores para p .

α	p	Modelo T-student 4 gl				Modelo Exponencial Potência			
		LR	LR^*	LR_m	LR_m^*	LR	LR^*	LR_m	LR_m^*
0,5%	2	1,1	0,5	0,6	0,5	1,0	0,6	0,5	0,5
	3	1,5	0,5	0,5	0,5	1,2	0,4	0,4	0,4
	4	2,5	0,4	0,5	0,5	2,0	0,5	0,5	0,5
	5	3,2	0,4	0,5	0,5	2,8	0,6	0,3	0,3
	6	4,0	0,8	0,6	0,5	3,8	0,8	0,5	0,5
1%	2	1,9	0,9	1,2	1,1	1,7	1,0	1,0	0,9
	3	2,6	0,8	0,9	0,8	2,3	0,9	0,8	0,8
	4	3,7	0,9	0,9	0,9	3,2	1,1	1,0	1,0
	5	5,0	1,1	1,1	0,9	4,3	1,8	0,9	0,8
	6	6,0	1,3	1,1	1,0	5,7	1,7	0,9	0,9
5%	2	7,0	4,4	5,0	4,7	7,1	4,9	5,2	5,1
	3	9,2	4,6	5,1	4,8	8,2	4,9	4,6	4,5
	4	11,1	5,1	5,5	5,2	10,3	5,4	4,7	4,5
	5	14,0	5,7	5,1	5,0	13,1	6,4	4,9	4,8
	6	16,0	5,9	5,2	5,0	16,2	7,3	5,1	5,0
10%	2	12,9	8,9	9,9	9,5	12,4	9,7	10,0	9,8
	3	16,0	9,7	10,6	10,2	14,5	10,0	9,7	9,5
	4	19,0	10,0	10,2	10,0	17,4	10,8	10,0	9,8
	5	22,0	11,2	10,3	10,0	21,2	12,2	10,3	10,0
	6	24,9	11,6	10,2	9,8	24,6	14,0	10,5	10,3

Tabela 4.3: Tamanho dos testes para os Modelo t-Student com $\nu = 4$ e Exponencial potência com $k = 0,3$; ambos com $n = 35$, $p = 3$ e diversos valores para q .

α	q	Modelo T-student 4 gl				Modelo Exponencial Potência			
		LR	LR^*	LR_m	LR_m^*	LR	LR^*	LR_m	LR_m^*
0,5%	2	1,5	0,5	0,5	0,5	1,2	0,4	0,4	0,4
	3	1,6	0,5	0,6	0,5	1,6	0,6	0,4	0,3
	4	2,7	0,5	0,6	0,5	2,2	0,7	0,4	0,3
	5	3,2	0,6	0,7	0,5	2,2	0,7	0,3	0,3
	6	3,1	0,6	0,5	0,4	2,6	0,8	0,4	0,4
1%	2	2,6	0,8	0,9	0,8	2,3	0,9	0,8	0,8
	3	2,8	1,0	1,0	1,0	2,8	1,1	0,9	0,9
	4	4,1	1,3	1,1	0,9	3,3	1,4	0,9	0,9
	5	5,2	1,4	1,3	1,1	3,8	1,4	0,7	0,7
	6	5,1	1,3	1,2	0,9	4,2	1,6	0,7	0,6
5%	2	9,2	4,6	5,1	4,8	8,2	4,9	4,6	4,5
	3	9,4	4,6	5,0	4,6	9,9	5,7	5,3	5,0
	4	12,0	5,9	5,6	5,1	10,7	6,0	4,3	4,1
	5	14,7	6,1	6,3	5,5	11,8	6,4	4,4	4,2
	6	14,8	6,5	5,8	4,9	13,9	7,3	4,0	3,8
10%	2	16,0	9,7	10,6	10,2	14,5	10,0	9,7	9,5
	3	16,5	9,4	10,4	9,9	17,2	10,9	10,1	9,8
	4	19,7	10,8	10,8	10,0	18,2	11,2	8,9	8,6
	5	21,9	11,8	11,9	10,6	19,7	12,0	9,0	8,5
	6	22,9	11,9	11,5	10,0	22,1	13,6	8,7	8,1

Tabela 4.4: Poder dos testes - Modelo t -Student com $\nu = 4$, $p = 5$, $q = 3$, $n = 35$ e $\alpha = 10\%$.

γ	$p = 5$			
	LR	LR^*	LR_m	LR_m^*
0,1	8,3	8,5	13,3	13,3
0,2	10,9	10,5	24,5	24,6
0,3	17,9	18,8	43,6	43,6
0,4	31,0	31,0	63,4	63,4
0,5	49,2	49,4	79,7	79,7
0,6	68,5	68,5	91,3	91,3
0,7	80,7	80,7	95,8	95,8
0,8	92,6	92,9	99,1	99,1
0,9	97,1	97,3	99,7	99,7
1,0	99,4	99,4	99,9	99,9

Tabela 4.5: Poder dos testes - Modelo Exponencial potência com $k = 0,3$; $p = 5$, $q = 3$, $n = 35$ e $\alpha = 10\%$.

γ	$p=5$			
	LR	LR^*	LR_m	LR_m^*
0,1	10,9	10,6	10,5	10,6
0,2	16,4	16,0	23,2	23,2
0,3	28,5	28,5	42,2	42,2
0,4	59,7	59,9	75,1	75,1
0,5	76,8	76,7	86,8	86,9
0,6	88,6	89,0	93,8	93,8
0,7	97,8	97,8	98,9	98,9

CAPÍTULO 5

Considerações finais

A principal contribuição teórica desta dissertação concentra-se no Capítulo 3, onde obtivemos uma expressão geral para o aperfeiçoamento do teste da razão de verossimilhanças perfiladas modificadas com base no ajuste proposto por Cox e Reid (1987) para qualquer classe de modelos em que os parâmetros que o compõem sejam ortogonais e que siga a mesma partição do vetor de parâmetros adotada nesta dissertação e ainda derivamos um fator de correção de Bartlett para um teste de heteroscedasticidade baseado na estatística da razão de verossimilhanças perfiladas modificadas (LR_m) nos modelos não lineares simétricos heteroscedásticos.

No Quinto Capítulo, realizamos um estudo de simulação de Monte Carlo a fim de avaliar o desempenho dos testes baseados na estatística da razão de verossimilhanças (LR), em sua versão corrigida via Bartlett (LR^*), na estatística da razão de verossimilhanças perfiladas modificadas (LR_m) e em sua respectiva versão corrigida via Bartlett (LR_m^*). O estudo de simulação para o teste baseado na estatística LR^* também é contribuição desta dissertação, uma vez que na literatura apenas foi proposto o fator de correção para a estatística em questão. A partir dos resultados obtidos verificamos que para todos os modelos considerados o teste da razão de verossimilhanças (baseado

na estatística LR) é bastante liberal, conduzindo a taxas de rejeição da hipótese nula de homoscedasticidade superiores aos níveis nominais adotados. O teste da razão de verossimilhanças corrigido via DiCiccio e Stern (baseado em LR^*) tende a atenuar a tendência liberal do teste usual, trazendo a taxa de rejeição de H_0 para um valor mais próximo do nível nominal adotado. Entretanto, conforme aumentamos o número de parâmetros de perturbação ou de parâmetros de interesse, o teste baseado em LR^* apresentou taxas de rejeição de H_0 superiores aos níveis nominais adotados, assim, mostrando-se sensível ao aumento do número de parâmetros de interesse e de perturbação, além disso, de caráter liberal.

O teste da razão de verossimilhanças perfiladas modificadas (baseado em LR_m) apresentou-se eficiente, corrigindo a tendência liberal do teste baseado na estatística LR . Conforme aumentamos os tamanhos amostrais, o teste baseado na estatística LR_m apresentou, de forma geral, taxas de rejeição bastante próximas aos níveis nominais adotados. À medida que aumentamos o número de parâmetros de perturbação ou de interesse, as taxas de rejeição do teste baseado em LR_m não se alteraram, mostrando que o teste baseado em LR_m não é sensível a estes aumentos. No mais, em geral, o teste baseado em LR_m ainda mostrou-se ligeiramente liberal, principalmente para o modelo t -Student. O teste da razão de verossimilhanças perfiladas modificadas corrigido (baseado na estatística LR_m^*) conduziu à taxas de rejeição da hipótese ainda mais próximas dos níveis nominais adotados, para o modelo t -Student. Assim como no teste baseado em LR_m , conforme aumentamos o tamanho amostral, o teste baseado em LR_m^* apresentou, de forma geral, taxas próximas aos níveis nominais adotados. Conforme aumentamos o número de parâmetros de interesse ou de perturbação, o teste baseado em LR_m^* apresentou comportamento similar ao apresentado pelo teste baseado em LR_m . Para o modelo exponencial potência, temos que, de modo geral, quando aumentamos o tamanho amostral e também quando aumentamos o número de parâmetros de interesse ou de perturbação, os testes baseados nas estatísticas LR_m e LR_m^* mostraram-se conservadores, apresentando taxas de

rejeição da hipótese nula inferiores aos níveis nominais considerados. Foi possível observar também que para o modelo Exponencial potência, o impacto da correção de Bartlett foi, de modo geral, reduzido, uma vez que o teste baseado na estatística LR_m já apresentava caráter conservativo. Em relação aos poderes dos testes, os testes baseados em LR_m e LR_m^* apresentaram poder semelhante e mostraram-se mais poderosos do que os testes baseados em LR e LR^* .

Em suma, dentre as estatísticas não corrigidas apresentadas nesta dissertação, temos que a estatística LR_m produziu um melhor teste de heteroscedasticidade, dentre as estatísticas corrigidas, a estatística LR_m^* produziu um melhor teste de heteroscedasticidade e dentre as estatísticas LR_m e LR_m^* , a estatística LR_m^* produziu um melhor teste de heteroscedasticidade.

APÊNDICE A

Cálculo dos cumulantes

Considere o modelo não linear simétrico heteroscedástico apresentado na Seção 2.2. Neste apêndice, apresentamos a obtenção de algumas derivadas do logaritmo da função de verossimilhança dos parâmetros do modelo em estudo, alguns cumulantes da log-verossimilhança e suas respectivas derivadas, necessários para o cálculo do fator de correção de Bartlett proposto por DiCiccio e Stern (1994). Para isto, faremos uso das notações definidas nas Seções 2.3 e 3.2.

A.1 Derivadas do logaritmo da função de verossimilhança

Considere o logaritmo da função de verossimilhança $l(\theta)$ definido em (2.5). Diferenciando $l(\theta)$ em relação aos componentes do vetor de parâmetros β , obtemos:

$$\begin{aligned} U_j &= \frac{\partial l(\theta)}{\partial \beta_j} = \sum_i \left\{ \frac{\partial t(z_i)}{\partial z_i} \frac{\partial z_i}{\partial \mu_i} \frac{\mu_i}{\beta_j} \right\} \\ &= \sum_i t_{(z_i)}^{(1)} \left(-\frac{1}{\sqrt{\phi_i}} \right) (j)_i = - \sum_i \frac{t_{(z_i)}^{(1)}}{\sqrt{\phi_i}} (j)_i. \end{aligned}$$

Procedendo de forma análoga, podemos obter as derivadas de segunda, terceira e quarta ordem de $l(\boldsymbol{\theta})$ com relação a $\boldsymbol{\beta}$ da seguinte maneira:

$$\begin{aligned} U_{jl} &= \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l} = - \sum_i t_{(z_i)}^{(2)} \left(-\frac{1}{\phi_i} \right) (j, l)_i - \sum_i t_{(z_i)}^{(1)} \frac{1}{\sqrt{\phi_i}} (jl)_i \\ &= \sum_i t_{(z_i)}^{(2)} \frac{1}{\phi_i} (j, l)_i - \sum_i t_{(z_i)}^{(1)} \frac{1}{\sqrt{\phi_i}} (jl)_i, \end{aligned}$$

$$\begin{aligned} U_{jlm} &= \frac{\partial^3 l(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l \partial \beta_m} = \sum_i t_{(z_i)}^{(3)} \left(-\frac{1}{\sqrt{\phi_i}} \right) (m)_i \frac{1}{\phi_i} (j, l)_i + \sum_i t_{(z_i)}^{(2)} \frac{1}{\phi_i} \{ (jm, l)_i + (j, lm)_i \} \\ &\quad - \sum_i t_{(z_i)}^{(2)} \left(-\frac{1}{\sqrt{\phi_i}} \right) (m)_i \frac{1}{\sqrt{\phi_i}} (jl)_i - \sum_i t_{(z_i)}^{(1)} \frac{1}{\sqrt{\phi_i}} (jlm)_i \\ &= - \sum_i t_{(z_i)}^{(3)} \frac{1}{\phi_i^{3/2}} (j, l, m)_i + \sum_i t_{(z_i)}^{(2)} \frac{1}{\phi_i} \{ (jm, l)_i + (j, lm)_i + (jl, m)_i \} \\ &\quad - \sum_i t_{(z_i)}^{(1)} \frac{1}{\sqrt{\phi_i}} (jlm)_i, \end{aligned}$$

e

$$\begin{aligned} U_{jlmn} &= \frac{\partial^4 l(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l \partial \beta_m \partial \beta_n} \\ &= \sum_i t_{(z_i)}^{(4)} \frac{1}{\phi_i^2} (j, l, m, n)_i - \sum_i t_{(z_i)}^{(3)} \frac{1}{\phi_i^{3/2}} \{ (jn, l, m)_i + (j, ln, m)_i + (j, l, mn)_i \} \\ &\quad - \sum_i t_{(z_i)}^{(3)} \frac{1}{\phi_i^{3/2}} \{ (jm, l, n)_i + (j, lm, n)_i + (jl, m, n)_i \} + \sum_i t_{(z_i)}^{(2)} \frac{1}{\phi_i} \{ (jmn, l)_i \\ &\quad + (jm, ln)_i + (jn, lm)_i + (j, lmn)_i + (jln, m)_i + (jl, mn)_i \} \\ &\quad + \sum_i t_{(z_i)}^{(2)} \frac{1}{\phi_i} (jlm, n)_i - \sum_i t_{(z_i)}^{(1)} \frac{1}{\sqrt{\phi_i}} (jlmn)_i. \end{aligned}$$

Agora derivando $l(\boldsymbol{\theta})$ com relação aos componentes do vetor $\boldsymbol{\gamma}$, obtemos:

$$\begin{aligned} U_a &= \frac{\partial l(\boldsymbol{\theta})}{\partial \gamma_a} = -\frac{1}{2} \sum_i \left\{ \frac{\partial \log \phi_i}{\partial \phi_i} \frac{\partial \phi_i}{\partial \tau_i} \frac{\partial \tau_i}{\partial \gamma_a} \right\} + \sum_i \left\{ \frac{\partial t(z_i)}{\partial z_i} \frac{\partial z_i}{\partial \phi_i} \frac{\partial \phi_i}{\partial \tau_i} \frac{\partial \tau_i}{\partial \gamma_a} \right\} \\ &= -\frac{1}{2} \sum_i \frac{h'_i}{\phi_i} (a)_i - \frac{1}{2} \sum_i t_{(z_i)}^{(1)} z_i \frac{h'_i}{\phi_i} (a)_i. \end{aligned}$$

Procedendo de forma análoga, podemos obter as derivadas de segunda, terceira e quarta ordem de $l(\boldsymbol{\theta})$ com relação a $\boldsymbol{\gamma}$ da seguinte maneira:

$$\begin{aligned}
U_{ab} &= \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \gamma_a \partial \gamma_b} \\
&= -\frac{1}{2} \sum_i \frac{h_i''}{\phi_i} (a, b)_i + \frac{1}{2} \sum_i \frac{h_i'^2}{\phi_i^2} (a, b)_i + \frac{1}{4} \sum_i t_{(z_i)}^{(2)} z_i^2 \frac{h_i'^2}{\phi_i^2} (a, b)_i \\
&\quad + \frac{1}{4} \sum_i t_{(z_i)}^{(1)} z_i \frac{h_i'^2}{\phi_i^2} (a, b)_i - \sum_i t_{(z_i)}^{(1)} z_i \frac{h_i''}{\phi_i} (a, b)_i + \frac{1}{2} \sum_i t_{(z_i)}^{(1)} z_i \frac{h_i'^2}{\phi_i^2} (a, b)_i,
\end{aligned}$$

$$\begin{aligned}
U_{abc} &= \frac{\partial^3 l(\boldsymbol{\theta})}{\partial \gamma_a \partial \gamma_b \partial \gamma_c} \\
&= -\frac{1}{2} \sum_i \frac{h_i'''}{\phi_i} (a, b, c)_i + \frac{3}{2} \sum_i \frac{h_i' h_i''}{\phi_i^2} (a, b, c)_i - \sum_i \frac{h_i'^3}{\phi_i^3} (a, b, c)_i \\
&\quad - \frac{1}{8} \sum_i t_{(z_i)}^{(3)} z_i^3 \frac{h_i'^3}{\phi_i^3} (a, b, c)_i - \frac{9}{8} \sum_i t_{(z_i)}^{(2)} z_i^2 \frac{h_i'^3}{\phi_i^3} + \frac{3}{4} \sum_i t_{(z_i)}^{(2)} z_i^2 \frac{h_i' h_i''}{\phi_i^2} (a, b, c)_i \\
&\quad - \frac{15}{8} \sum_i t_{(z_i)}^{(1)} z_i \frac{h_i'^3}{\phi_i^3} (a, b, c)_i + \frac{9}{4} \sum_i t_{(z_i)}^{(1)} z_i \frac{h_i' h_i''}{\phi_i^2} (a, b, c)_i \\
&\quad - \frac{1}{2} \sum_i t_{(z_i)}^{(1)} z_i \frac{h_i'''}{\phi_i} (a, b, c)_i
\end{aligned}$$

e

$$\begin{aligned}
U_{abcd} &= \frac{\partial^4 l(\boldsymbol{\theta})}{\partial \gamma_a \partial \gamma_b \partial \gamma_c \partial \gamma_d} \\
&= -\frac{1}{2} \sum_i \frac{h_i''''}{\phi_i} (a, b, c, d)_i + 2 \sum_i \frac{h_i' h_i'''}{\phi_i^2} (a, b, c, d)_i + \frac{3}{2} \sum_i \frac{h_i''^2}{\phi_i^2} (a, b, c, d)_i \\
&\quad - 6 \sum_i \frac{h_i'^2 h_i''}{\phi_i^3} (a, b, c, d)_i + 3 \sum_i \frac{h_i'^4}{\phi_i^4} (a, b, c, d)_i + \frac{1}{16} \sum_i t_{(z_i)}^{(4)} z_i^4 \frac{h_i'^4}{\phi_i^4} (a, b, c, d)_i \\
&\quad + \frac{9}{8} \sum_i t_{(z_i)}^{(3)} z_i^3 \frac{h_i'^4}{\phi_i^4} (a, b, c, d)_i - \frac{3}{4} \sum_i t_{(z_i)}^{(3)} z_i^3 \frac{h_i'^2 h_i''}{\phi_i^3} (a, b, c, d)_i \\
&\quad + \frac{87}{16} \sum_i t_{(z_i)}^{(2)} z_i^2 \frac{h_i'^4}{\phi_i^4} (a, b, c, d)_i - \frac{27}{4} \sum_i t_{(z_i)}^{(2)} z_i^2 \frac{h_i'^2 h_i''}{\phi_i^3} (a, b, c, d)_i \\
&\quad + \frac{3}{4} \sum_i t_{(z_i)}^{(2)} z_i^2 \frac{h_i''^2}{\phi_i^2} (a, b, c, d)_i + \sum_i t_{(z_i)}^{(2)} z_i^2 \frac{h_i' h_i''}{\phi_i^2} (a, b, c, d)_i
\end{aligned}$$

$$\begin{aligned}
& + \frac{105}{16} \sum_i t_{(z_i)}^{(1)} z_i \frac{h_i'^4}{\phi_i^4} (a, b, c, d)_i - \frac{45}{4} \sum_i t_{(z_i)}^{(1)} z_i \frac{h_i'^2 h_i''}{\phi_i^3} (a, b, c, d)_i \\
& + \frac{9}{4} \sum_i t_{(z_i)}^{(1)} z_i \frac{h_i''^2}{\phi_i^2} (a, b, c, d)_i + 3 \sum_i t_{(z_i)}^{(1)} z_i \frac{h_i' h_i'''}{\phi_i^2} (a, b, c, d)_i \\
& - \frac{1}{2} \sum_i t_{(z_i)}^{(1)} z_i \frac{h_i''''}{\phi_i} (a, b, c, d)_i.
\end{aligned}$$

Para obter alguns cumulantes mistos, calculamos as seguintes derivadas mistas:

$$\begin{aligned}
U_{ja} &= \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \beta_j \partial \gamma_a} = \frac{1}{2} \sum_i t_{(z_i)}^{(2)} z_i \frac{h_i'}{\phi_i^{3/2}} (j, a)_i + \frac{1}{2} \sum_i t_{(z_i)}^{(1)} \frac{h_i'}{\phi_i^{3/2}} (j, a)_i, \\
U_{abj} &= \frac{\partial^3 l(\boldsymbol{\theta})}{\partial \gamma_a \partial \gamma_b \partial \beta_j} = -\frac{1}{4} \sum_i t_{(z_i)}^{(3)} z_i^2 \frac{h_i'^2}{\phi_i^{3/2}} (a, b, j)_i \\
& - \frac{1}{2} \sum_i t_{(z_i)}^{(2)} z_i \frac{h_i'^2}{\phi_i^{5/2}} (a, b, j)_i - \frac{1}{4} \sum_i t_{(z_i)}^{(2)} z_i \frac{h_i'^2}{\phi_i^{5/2}} (a, b, j)_i \\
& - \frac{1}{4} \sum_i t_{(z_i)}^{(1)} \frac{h_i'^2}{\phi_i^{5/2}} (a, b, j)_i + \frac{1}{2} \sum_i t_{(z_i)}^{(2)} z_i \frac{h_i''}{\phi_i^{5/2}} (a, b, j)_i \\
& + \frac{1}{2} \sum_i t_{(z_i)}^{(1)} \frac{h_i''}{\phi_i^{5/2}} (a, b, j)_i - \frac{1}{2} \sum_i t_{(z_i)}^{(2)} z_i \frac{h_i'^2}{\phi_i^{5/2}} (a, b, j)_i \\
& - \frac{1}{2} \sum_i t_{(z_i)}^{(1)} \frac{h_i'^2}{\phi_i^{5/2}} (a, b, j)_i, \\
U_{jla} &= \frac{\partial^3 l(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l \partial \gamma_a} = -\frac{1}{2} \sum_i t_{(z_i)}^{(3)} z_i \frac{h_i'}{\phi_i^2} (j, l, a)_i - \sum_i t_{(z_i)}^{(2)} \frac{h_i'}{\phi_i^2} (j, l, a)_i \\
& + \frac{1}{2} \sum_i t_{(z_i)}^{(2)} z_i \frac{h_i'}{\phi_i^{3/2}} (jl, a)_i + \frac{1}{2} \sum_i t_{(z_i)}^{(1)} \frac{h_i'}{\phi_i^{3/2}} (jl, a)_i,
\end{aligned}$$

e

$$\begin{aligned}
U_{jlab} &= \frac{\partial^4 l(\boldsymbol{\theta})}{\partial \beta_j \partial \beta_l \partial \gamma_a \partial \gamma_b} = \frac{1}{4} \sum_i t_{(z_i)}^{(4)} z_i^2 \frac{h_i'^2}{\phi_i^3} (j, l, a, b)_i \\
& + \frac{7}{4} \sum_i t_{(z_i)}^{(3)} z_i \frac{h_i'^2}{\phi_i^3} (j, l, a, b)_i - \frac{1}{2} \sum_i t_{(z_i)}^{(3)} z_i \frac{h_i''}{\phi_i^2} (j, l, a, b)_i
\end{aligned}$$

$$\begin{aligned}
& - \frac{1}{2} \sum_i t_{(z_i)}^{(2)} \frac{h_i''}{\phi_i^2} (j, l, a, b)_i + 2 \sum_i t_{(z_i)}^{(2)} \frac{h_i'^2}{\phi_i^3} (j, l, a, b)_i \\
& - \frac{1}{4} \sum_i t_{(z_i)}^{(3)} z_i^2 \frac{h_i'^2}{\phi_i^{5/2}} (jl, a, b)_i - \frac{5}{4} \sum_i t_{(z_i)}^{(2)} z_i \frac{h_i'^2}{\phi_i^{5/2}} (jl, a, b)_i \\
& + \frac{1}{2} \sum_i t_{(z_i)}^{(2)} z_i \frac{h_i''}{\phi_i^{3/2}} (jl, a, b)_i + \frac{1}{2} \sum_i t_{(z_i)}^{(1)} \frac{h_i''}{\phi_i^{3/2}} (jl, a, b)_i \\
& - \frac{3}{4} \sum_i t_{(z_i)}^{(1)} \frac{h_i'^2}{\phi_i^{5/2}} (jl, a, b)_i.
\end{aligned}$$

A.2 Cálculo dos Cumulantes

Tomando a esperança das derivadas de segunda, terceira e quarta ordem com relação aos parâmetros β e γ calculadas acima, obtemos os seguintes cumulantes:

$$\begin{aligned}
\lambda_{jl} &= \delta_{(0,1,0,0,0)} \sum_i \frac{1}{\phi_i} (j, l)_i, \\
\lambda_{jlm} &= \delta_{(0,1,0,0,0)} \sum_i \frac{1}{\phi_i} \{ (lm, l)_i + (j, lm)_i + (jl, m)_i \}, \\
\lambda_{jlmn} &= \delta_{(0,0,0,1,0)} \sum_i \frac{1}{\phi_i^2} (j, l, m, n)_i \{ (jm, l)_i + (j, lm)_i + (jl, m)_i \}, \\
\lambda_{ab} &= \frac{\delta_{(0,1,0,0,2)} - 1}{4} \sum_i \frac{h_i'^2}{\phi_i^2} (a, b)_i, \\
\lambda_{abc} &= \frac{3(\delta_{(0,1,0,0,2)} - 1)}{4} \sum_i \frac{h_i' h_i''}{\phi_i^2} (a, b, c)_i, \\
&+ \frac{(7 - \delta_{(0,0,1,0,3)} - 9\delta_{(0,1,0,0,2)})}{8} \sum_i \frac{h_i'^3}{\phi_i^3} (a, b, c)_i,
\end{aligned}$$

$$\begin{aligned}
\lambda_{abcd} &= -\sum_i \frac{h'_i h''_i}{\phi_i^2} (a, b, c, d)_i - \frac{3}{4} \sum_i \frac{h''_i{}^2}{\phi_i^2} (a, b, c, d)_i, \\
&+ \frac{21}{4} \sum_i \frac{h_i'^2 h''_i}{\phi_i^3} (a, b, c, d)_i - \frac{57}{16} \sum_i \frac{h_i'^4}{\phi_i^4} (a, b, c, d)_i, \\
&+ \frac{1}{16} \delta_{(0,0,0,4,4)} \sum_i \frac{h_i'^4}{\phi_i^4} (a, b, c, d)_i + \frac{9}{8} \delta_{(0,0,1,0,3)} \sum_i \frac{h_i'^4}{\phi_i^4} (a, b, c, d)_i, \\
&- \frac{3}{4} \delta_{(0,0,1,0,3)} \sum_i \frac{h_i'^2 h''_i}{\phi_i^3} (a, b, c, d)_i + \frac{87}{16} \delta_{(0,1,0,0,2)} \sum_i \frac{h_i'^4}{\phi_i^4} (a, b, c, d)_i, \\
&- \frac{27}{4} \delta_{(0,1,0,0,2)} \sum_i \frac{h_i'^2 h''_i}{\phi_i^3} (a, b, c, d)_i + \frac{3}{4} \delta_{(0,1,0,0,2)} \sum_i \frac{h''_i{}^2}{\phi_i^2} (a, b, c, d)_i, \\
&+ \delta_{(0,1,0,0,2)} \sum_i \frac{h'_i h''_i}{\phi_i^2} (a, b, c, d)_i, \\
\lambda_{ja} &= \lambda_{abj} = 0, \\
\lambda_{jla} &= -\frac{1}{2} \delta_{(0,0,1,0,1)} \sum_i \frac{h'_i}{\phi_i^2} (j, l, a)_i - \delta_{(0,1,0,0,0)} \sum_i \frac{h'_i}{\phi_i^2} (j, l, a)_i \text{ e} \\
\lambda_{jlab} &= \frac{1}{4} \delta_{(0,0,0,1,2)} \sum_i \frac{h_i'^2}{\phi_i^3} (j, l, a, b)_i + \frac{7}{4} \delta_{(0,0,1,0,1)} \sum_i \frac{h_i'^2}{\phi_i^3} (j, l, a, b)_i \\
&- \frac{1}{2} \delta_{(0,0,1,0,1)} \sum_i \frac{h''_i}{\phi_i^2} (j, l, a, b)_i - \frac{1}{2} \delta_{(0,1,0,0,0)} \sum_i \frac{h''_i}{\phi_i^2} (j, l, a, b)_i \\
&+ 2 \delta_{(0,1,0,0,0)} \sum_i \frac{h_i'^2}{\phi_i^3} (j, l, a, b)_i
\end{aligned}$$

A.3 Derivadas dos Cumulantes

Nesta seção encontram-se as derivadas dos cumulantes necessárias ao cálculo do fator de correção c_m com relação aos componentes de β e γ .

$$\begin{aligned}
(\lambda_{jl})_m &= \delta_{(0,1,0,0,0)} \sum_i \frac{1}{\phi_i} \{ (jl, m)_i + (j, lm)_i \}, \\
(\lambda_{ab})_j &= 0, \\
(\lambda_{ab})_c &= \left(\frac{\delta_{(0,1,0,0,2)} - 1}{2} \right) \sum_i \left\{ \frac{h'_i h''_i}{\phi_i^2} - \frac{h_i'^3}{\phi_i^3} \right\} (a, b, c)_i, \\
(\lambda_{ab})_{cd} &= \left(\frac{\delta_{(0,1,0,0,2)} - 1}{2} \right) \sum_i \left\{ \frac{h_i''^2}{\phi_i^2} + \frac{h'_i h''_i}{\phi_i^2} - \frac{5h_i'^2 h''_i}{\phi_i^3} + \frac{3h_i'^4}{\phi_i^4} \right\} (a, b, c, d)_i, \\
(\lambda_{abc})_d &= \frac{3(\delta_{(0,1,0,0,2)} - 1)}{4} \sum_i \left\{ \frac{h_i''^2}{\phi_i^2} + \frac{h'_i h_i'''}{\phi_i^2} - \frac{2h_i'^2 h''_i}{\phi_i^3} \right\} (a, b, c, d)_i, \\
&\quad + \frac{(7 - \delta_{(0,0,1,0,3)} - 9\delta_{(0,1,0,0,2)})}{8} \sum_i \left\{ \frac{3h_i'^2 h''_i}{\phi_i^3} - \frac{3h_i'^4}{\phi_i^4} \right\} (a, b, c, d)_i \quad \text{e} \\
(\lambda_{abj})_l &= 0.
\end{aligned}$$

APÊNDICE B

Valores dos δ' s

Neste apêndice encontram-se os valores dos δ' s associados à correção de Bartlett para algumas distribuições pertencentes à classe dos modelos simétricos. Estes valores foram obtidos por Uribe-Opazo (1997). As notações $t_{(z_i)}^k$ e $\delta_{(a,b,c,d,e)}$ estão definidas na Seção 2.3 e neste estudo consideramos $z_i = \frac{y_i - \mu_i}{\sqrt{\phi_i}}$, com $\mu_i = f(x_i; \beta)$.

B.1 Distribuição Normal

Seja $y_i \sim N(\mu_i, \phi_i)$ com função de densidade da forma

$$\pi(y_i, \beta_i, \phi_i) = \frac{1}{\sqrt{\phi_i} 2\pi} \exp \left\{ -\frac{1}{2} \frac{(y_i - \mu_i)^2}{\phi_i} \right\}, \quad y \in \mathbb{R},$$

para $\beta \in \mathbb{R}^p$ e $\phi > 0$. O logaritmo da função de verossimilhança é expresso por

$$\ell(\beta, \gamma) = -\frac{1}{2} \sum_{i=1}^n \log \phi_i + \sum_{i=1}^n t(z_i),$$

com $t(z_i) = -z_i^2/2$. Por definição $t^{(k)} = \partial^k t / \partial z^k$, logo,

$$t_{(z_i)}^{(1)} = -z, \quad t_{(z_i)}^{(2)} = -1, \quad t_{(z_i)}^{(3)} = t_{(z_i)}^{(4)} = 0.$$

Como $z \sim N(0, 1)$, temos que

$$\begin{aligned} \delta_{(2,0,0,0,0)} &= 1, & \delta_{(0,0,0,1,0)} &= 0, & \delta_{(2,1,0,0,4)} &= -16, \\ \delta_{(2,1,0,0,0)} &= -1, & \delta_{(0,0,0,1,2)} &= 0, & \delta_{(3,0,0,0,3)} &= -15, \\ \delta_{(0,0,1,0,1)} &= 0, & \delta_{(0,1,0,0,0)} &= -1, & \delta_{(2,1,0,0,2)} &= -3, \\ \delta_{(0,0,1,0,3)} &= 0, & \delta_{(1,1,0,0,1)} &= 1, & \delta_{(1,0,1,0,0)} &= 0, \\ \delta_{(4,0,0,0,2)} &= 15, & \delta_{(2,0,0,0,2)} &= 3, & \delta_{(0,0,0,1,4)} &= 0, \\ \delta_{(4,0,0,0,4)} &= 105, & \delta_{(3,0,0,0,1)} &= -3, & \delta_{(0,1,0,0,2)} &= -1. \end{aligned}$$

B.2 Distribuição de Cauchy

Seja $y_i \sim C(\mu_i, \phi_i)$ com função densidade dada da forma

$$\pi(y_i, \phi_i) = \frac{1}{\pi\sqrt{\phi_i}} \left[1 + \frac{(y_i - \mu_i)^2}{\phi_i} \right]^{-1}, \quad y_i \in \mathbb{R},$$

para $\beta \in \mathbb{R}^p$ e $\phi > 0$. O logaritmo da função de verossimilhança é dado por

$$\ell(\beta, \gamma) = -\frac{1}{2} \sum_{i=1}^n \log \phi_i + \sum_{i=1}^n t(z_i),$$

onde $t(z_i) = \log \left(\frac{1}{\pi} [1 + z^2]^{-1} \right)$. Assim, temos que

$$\begin{aligned} t_{(Z_i)}^{(1)} &= 2zs, \\ t_{(Z_i)}^{(2)} &= 2s + 4z^2s^2, \\ t_{(Z_i)}^{(3)} &= 12zs^2 + 16z^3s^3, \\ t_{(Z_i)}^{(4)} &= 12s^2 + 96z^2s^2 + 96z^4s^4, \end{aligned}$$

onde s é dado na Tabela 2.2. Considerando que $z \sim \text{Cauchy}(0, 1)$, então

$z^2 \sim F(1, 1)$, portanto:

$$\begin{aligned} \delta_{(2,0,0,0,0)} &= 1/2, & \delta_{(0,0,0,1,0)} &= 3/4, & \delta_{(2,1,0,0,4)} &= 15/8, \\ \delta_{(2,1,0,0,0)} &= -1/8, & \delta_{(0,0,0,1,2)} &= -3/4, & \delta_{(3,0,0,0,3)} &= -5/2, \\ \delta_{(0,0,1,0,1)} &= 1/2, & \delta_{(0,1,0,0,0)} &= -1/2, & \delta_{(2,1,0,0,2)} &= -3/4, \\ \delta_{(0,0,1,0,3)} &= -1/2, & \delta_{(1,1,0,0,1)} &= 0, & \delta_{(1,0,1,0,0)} &= 1/8, \\ \delta_{(4,0,0,0,2)} &= 5/8, & \delta_{(2,0,0,0,2)} &= 3/2, & \delta_{(0,0,0,1,4)} &= 3/4, \\ \delta_{(4,0,0,0,4)} &= 35/8, & \delta_{(3,0,0,0,1)} &= -1/2, & \delta_{(0,1,0,0,2)} &= 1/2. \end{aligned}$$

B.3 Distribuição t-Student

Seja $y_i \sim t(\mu_i, \phi_i, \nu)$ com a função densidade da forma

$$\pi(y_i, \mu_i, \phi_i, \nu) = \frac{\nu^{\nu/2}}{\sqrt{\phi_i} B(1/2, \nu/2)} \left[\nu + \frac{(y_i - \mu_i)^2}{\phi_i} \right]^{-\frac{\nu+1}{2}}, \quad y_i \in \mathbb{R},$$

onde $\beta \in \mathbb{R}^p$, $\phi > 0$, $\nu > 0$ e $B(\cdot)$ é a função beta. O logaritmo da função de verossimilhança tem a forma

$$\ell(\beta, \gamma) = -\frac{1}{2} \sum_{i=1}^n \log \phi_i + \sum_{i=1}^n t(z_i),$$

com $t(z_i) = \log \left(\frac{\nu^{\frac{\nu}{2}}}{B(1/2, \nu/2)} [\nu + z_i^2]^{\frac{-(\nu+1)}{2}} \right)$. Derivando t_i com relação à z_i , temos

$$\begin{aligned} t_{(Z_i)}^{(1)} &= 2zs, \\ t_{(Z_i)}^{(2)} &= 2s + 4 \frac{8z^2 s^2}{\nu + 1}, \\ t_{(Z_i)}^{(3)} &= \frac{24}{\nu + 1} z s^2 + \frac{64}{\nu + 1} z^3 s^3, \\ t_{(Z_i)}^{(4)} &= \frac{24}{\nu + 1} s^2 + \frac{384}{(\nu + 1)^2} z^2 s^3 + \frac{768}{(\nu + 1)^3} z^4 s^4, \end{aligned}$$

onde s é dado na Tabela 2.2.

Considerando que $z \sim t(0, 1, \nu)$, temos $z^2 \sim F(1, \nu)$, portanto:

$$\begin{aligned} \delta_{(2,0,0,0,0)} &= \frac{(\nu+1)}{(\nu+3)}, & \delta_{(0,0,0,1,0)} &= \frac{6(\nu+1)(\nu+2)}{\nu(\nu+5)(\nu+7)}, \\ \delta_{(2,1,0,0,0)} &= \frac{-(\nu+1)^3(\nu+2)}{\nu(\nu+5)(\nu+7)(\nu+3)}, & \delta_{(0,0,0,1,2)} &= \frac{6}{(\nu+3)} \left(\frac{\nu-19}{(\nu+5)} + \frac{120}{(\nu+5)(\nu+7)} \right), \\ \delta_{(0,0,1,0,1)} &= \frac{6(\nu+1)(\nu+2)}{\nu(\nu+5)(\nu+7)}, & \delta_{(0,1,0,0,0)} &= -\frac{\nu+1}{\nu+3}, \\ \delta_{(0,0,1,0,3)} &= \frac{6(3\nu-5)}{(\nu+5)(\nu+3)}, & \delta_{(1,1,0,0,1)} &= \frac{(\nu+1)(\nu-1)}{(\nu+5)(\nu+3)}, \\ \delta_{(4,0,0,0,2)} &= \frac{15(\nu+1)^3}{(\nu+5)(\nu+7)(\nu+3)}, & \delta_{(2,0,0,0,2)} &= 3 \frac{\nu+1}{\nu+3}, \\ \delta_{(1,0,1,0,0)} &= \frac{-6(\nu+1)(\nu+2)}{\nu(\nu+5)(\nu+7)}, & \delta_{(0,1,0,0,2)} &= \frac{3-\nu}{\nu+3}, \\ \delta_{(2,1,0,0,2)} &= \frac{3(\nu+1)^2(3-\nu)}{(\nu+5)(\nu+7)(\nu+3)}, & \delta_{(3,0,0,0,1)} &= \frac{-3(\nu+1)^2}{(\nu+5)(\nu+3)}, \\ \delta_{(3,0,0,0,3)} &= \frac{-15(\nu+1)^2}{(\nu+5)(\nu+3)}, & \delta_{(4,0,0,0,4)} &= \frac{105(\nu+1)^3}{(\nu+5)(\nu+3)(\nu+7)}, \\ \delta_{(0,0,0,1,4)} &= \frac{18(\nu^2-28\nu+35)}{(\nu+3)(\nu+5)(\nu+7)} \end{aligned}$$

B.4 Distribuição t-Student generalizada

Seja $y_i \sim tG(y_i; \mu_i, \sqrt{\phi_i}, r, w)$ com função densidade da forma

$$\pi(y_i; \mu_i, \sqrt{\phi_i}, r, w) = \frac{w^{r/2}}{\sqrt{\phi_i} B(1/2, r/2)} \left[w + \frac{(y_i - \mu_i)^2}{\phi_i} \right]^{-\frac{r+1}{2}}, \quad y_i \in \mathbb{R},$$

onde $r, w > 0$, $\beta \in \mathbb{R}^p$ e $\phi > 0$. O logaritmo da função de verossimilhança tem a forma

$$\ell(\beta, \gamma) = -\frac{1}{2} \sum_{i=1}^n \log \phi_i + \sum_{i=1}^n t(z_i),$$

onde $t(z_i) = \log \left(\frac{w^{\frac{r}{2}}}{B(1/2, r/2)} [w + z_i^2]^{\frac{-(r+1)}{2}} \right)$. As derivadas de t com relação à z_i são

$$\begin{aligned} t_{(z_i)}^{(1)} &= 2zs, \\ t_{(z_i)}^{(2)} &= 2s + \frac{8z^2 s^2}{r+1}, \\ t_{(z_i)}^{(3)} &= \frac{24}{r+1} z s^2 + \frac{64}{r+1} z^3 s^3, \\ t_{(z_i)}^{(4)} &= \frac{24}{r+1} s^2 + \frac{384}{(r+1)^2} z^2 s^3 + \frac{768}{(r+1)^3} z^4 s^4, \end{aligned}$$

onde s é dado na Tabela 2.2.

Considerando que $z \sim tG(0, 1, r, w)$, temos

$$u = \frac{\sqrt{r}}{\sqrt{w}} z \sim t(0, 1, w) \quad v = \frac{r}{w} z^2 \sim F(1, r).$$

Logo,

$$E[z^{2k} w^m] = \frac{\left(-\frac{r+1}{2}\right)^m}{w^{m-k} B(1/2, r/2)} B\left(\frac{2k+1}{2}, \frac{r+k(m-k)}{2}\right), \quad m = 0, 1, 2, \dots$$

Assim, temos que os δ' s são dados por

$$\begin{aligned}
\delta_{(2,0,0,0,0)} &= \frac{r(r+1)}{w(r+3)}, & \delta_{(0,0,0,1,0)} &= \frac{6(r+1)(r+2)}{w^2(r+5)(r+7)}, \\
\delta_{(2,1,0,0,0)} &= \frac{-r(r+1)^3(r+2)}{3w^2(r+5)(r+7)(r+3)}, & \delta_{(0,0,0,1,2)} &= \frac{6r((r-19)(r+7)+120)}{w(r+3)(r+5)(r+7)}, \\
\delta_{(0,0,1,0,1)} &= \frac{6r(r+1)(\nu+2)}{\nu(\nu+5)(\nu+7)}, & \delta_{(0,1,0,0,0)} &= -\frac{r(r+1)}{w(r+3)}, \\
\delta_{(0,0,1,0,3)} &= \frac{6(3r-5)}{(r+5)(r+3)}, & \delta_{(1,1,0,0,1)} &= \frac{r(r^2-1)}{w(\nu+5)(\nu+3)}, \\
\delta_{(4,0,0,0,2)} &= \frac{15r(r+1)^3}{w(r+5)(r+7)(r+3)}, & \delta_{(2,0,0,0,2)} &= 3\frac{(r+1)}{(r+3)}, \\
\delta_{(1,0,1,0,0)} &= \frac{-6(r+1)(r+2)}{w^2(r+5)(r+7)}, & \delta_{(0,1,0,0,2)} &= \frac{3-r}{r+3}, \\
\delta_{(2,1,0,0,2)} &= \frac{-3r(r+1)^2(3-r)}{(r+5)(r+7)(r+3)}, & \delta_{(3,0,0,0,1)} &= \frac{-3r(r+1)^2}{w(r+5)(r+3)}, \\
\delta_{(3,0,0,0,3)} &= \frac{-15(r+1)^2}{(r+5)(r+3)}, & \delta_{(4,0,0,0,4)} &= \frac{105(r+1)^3}{(r+5)(r+3)(r+7)}, \\
\delta_{(0,0,0,1,4)} &= \frac{18(r^2-28r+35)}{(r+3)(r+5)(r+7)}.
\end{aligned}$$

B.5 Distribuição Logística I

Seja $y_i \sim LI(\mu_i, \phi_i)$ com função de densidade dada por

$$\pi(y_i, \mu_i, \phi_i) = \frac{c}{\sqrt{\phi_i}} \frac{\exp\left\{-\frac{(y_i - \mu_i)^2}{\phi_i}\right\}}{\left(1 + \exp\left\{-\frac{(y_i - \mu_i)^2}{\phi_i}\right\}\right)^2}, \quad y_i \in \mathbb{R},$$

onde $c \approx 1,484300029$, $\beta \in \mathbb{R}^p$ e $\phi > 0$. O logaritmo da função de verossimilhança é dado por

$$\ell(\beta, \gamma) = -\frac{1}{2} \sum_{i=1}^n \log \phi_i + \sum_{i=1}^n t(z_i),$$

com $t(z_i) = \log\left(\frac{e^{-z_i^2}}{1+e^{-z_i^2}}\right)$. As derivadas de t com relação à z_i são dadas por

$$\begin{aligned}
t_{(z_i)}^{(1)} &= 2zs, \\
t_{(z_i)}^{(2)} &= 2s + 2z^2(s^2 - 1), \\
t_{(z_i)}^{(3)} &= 6z(s^2 - 1) + 4z^3s(s^2 - 1), \\
t_{(z_i)}^{(4)} &= 6(s^2 - 1) + 24sz^2(s^2 - 1) + 4z^4(s^2 - 1)(3s^2 - 1),
\end{aligned}$$

onde s é dado na Tabela 2.2.

Considerando que $z \sim LI(0, 1)$ e

$$E [z^{2r} s^m] = (-1)^m \frac{c}{2} \int_0^1 \left(\log \left(\frac{1+w}{1-w} \right) \right)^{\frac{2r-1}{2}} w^m dw, \text{ para } m = 0, 1, 2, \dots,$$

onde $w = 1 - \frac{e^{z^2}}{(1+e^{-z^2})^2}$.

Assim, os $\delta's$ são dados por

$$\begin{array}{llll} \delta_{(2,0,0,0,0)} & \approx & 1,477240176, & \delta_{(0,0,0,1,0)} \approx 4,259052264, \\ \delta_{(2,1,0,0,0)} & \approx & 4,153806544, & \delta_{(0,0,0,1,2)} \approx 2,65931983, \\ \delta_{(0,0,1,0,1)} & \approx & -1,27916363, & \delta_{(0,1,0,0,0)} \approx -1,477240176, \\ \delta_{(0,0,1,0,3)} & \approx & -0,508877866, & \delta_{(1,1,0,0,1)} \approx 2,756409976, \\ \delta_{(4,0,0,0,2)} & \approx & 46,76577389, & \delta_{(2,0,0,0,2)} \approx 4,013783934, \\ \delta_{(1,0,1,0,0)} & \approx & -4,259052264, & \delta_{(0,1,0,0,2)} \approx -2,013783934, \\ \delta_{(2,1,0,0,2)} & \approx & -10,92854975, & \delta_{(3,0,0,0,1)} \approx 46,76577386, \\ \delta_{(3,0,0,0,3)} & \approx & -25,12989577, & \delta_{(4,0,0,0,4)} \approx 206,1514675, \\ \delta_{(0,0,0,1,4)} & \approx & 4,135810. & \end{array}$$

B.6 Distribuição Logística II

Seja $y_i \sim LII(\mu_i, \phi_i)$ com função de densidade dada por

$$\pi(y_i, \mu_i, \phi_i) = \frac{1}{\sqrt{\phi_i}} \frac{\exp \left\{ -\frac{(y_i - \mu_i)^2}{\phi_i} \right\}}{\left(1 + \exp \left\{ -\frac{(y_i - \mu_i)^2}{\phi_i} \right\} \right)^2}, \quad y_i \in \mathbb{R},$$

com $\beta \in \mathbb{R}^p$ e $\phi > 0$. O logaritmo da função de verossimilhança é dado por

$$\ell(\beta, \gamma) = -\frac{1}{2} \sum_{i=1}^n \log \phi_i + \sum_{i=1}^n t(z_i),$$

onde $t(z_i) = \log \left(\frac{e^{z_i}}{(1+e^{-z_i^2})^2} \right)$. As derivadas de t com relação à z_i são

$$\begin{aligned} t_{(z_i)}^{(1)} &= \frac{1 - e^z}{1 + e^z}, \\ t_{(z_i)}^{(2)} &= \frac{-2e^z}{(1 + e^z)^2}, \\ t_{(z_i)}^{(3)} &= \frac{2(e^{2z} - e^z)}{(1 + e^z)^3}, \\ t_{(z_i)}^{(4)} &= \frac{-2(e^{3z} - 4e^{2z} + e^z)}{(1 + e^z)^4}. \end{aligned}$$

As expressões para os $\delta's$ podem ser obtidas dos resultados desenvolvidos em Uribe-opazo (1997) para a distribuição logística generalizada com $m = 1$ e $\alpha = 1$. Temos que

$$\begin{array}{llll}
\delta_{(2,0,0,0,0)} & = & 1/3, & \delta_{(0,0,0,1,0)} & = & 1/15, \\
\delta_{(2,1,0,0,0)} & = & -1/15, & \delta_{(0,0,0,1,2)} & = & -0,11401, \\
\delta_{(0,0,1,0,1)} & = & 1/6, & \delta_{(0,1,0,0,0)} & = & -1/3, \\
\delta_{(0,0,1,0,3)} & = & 0,64493, & \delta_{(1,1,0,0,1)} & \approx & 1/6, \\
\delta_{(4,0,0,0,2)} & \approx & 1,99131, & \delta_{(2,0,0,0,2)} & \approx & 2,42996, \\
\delta_{(1,0,1,0,0)} & = & -1/5, & \delta_{(0,1,0,0,2)} & \approx & -0,42996, \\
\delta_{(2,1,0,0,2)} & \approx & -0,21932, & \delta_{(3,0,0,0,1)} & = & -2/3, \\
\delta_{(3,0,0,0,3)} & \approx & -8,57974, & \delta_{(4,0,0,0,4)} & \approx & 38,61046, \\
\delta_{(0,0,0,1,4)} & \approx & -1,14923.
\end{array}$$

B.7 Exponencial potência

Seja $y_i \sim EP(\mu_i, \phi_i, k)$ com função densidade dada por

$$\pi(y_i, \mu_i, \phi_i) = \frac{C(k)}{\sqrt{\phi_i}} \exp \left\{ -\frac{1}{2} \left[\frac{(y_i - \mu_i)^2}{\phi_i} \right]^{\frac{1}{1+k}} \right\}, \quad y_i \in \mathbb{R},$$

em que $C(k)^{-1} = \Gamma(1 + \frac{1+k}{2}) 2^{1+(1+k)/2}$, $\beta \in \mathbb{R}^p$, $\phi_i = h(z_i^\top \gamma)$, com $\phi_i > 0$ e $-1 < k \leq 1$. O logarítmo da função de verossimilhança é dado por

$$l(\beta, \gamma) = n \log C(k) - \frac{1}{2} \sum_{i=1}^n \log \phi_i + i + \sum_{i=1}^n t(z_i),$$

com $t(z_i) = -\frac{1}{2}|z_i|^{2/(1+k)}$. Deste modo, as quatro primeiras derivadas de t_i com relação a z_i são dadas por

$$\begin{aligned}
t_{(z_i)}^{(1)} &= -\frac{1}{(1+k)} z_i^{\frac{1-k}{1+k}}, \text{ se } -1 < k \leq 1, \\
t_{(z_i)}^{(2)} &= -\frac{1-k}{(1+k)^2} z_i^{-\frac{2k}{1+k}}, \text{ se } k > 0, \\
t_{(z_i)}^{(3)} &= \frac{2(1-k)}{(1+k)^3} z_i^{-\frac{1+3k}{1+k}}, \text{ se } k < -1/3, \\
t_{(z_i)}^{(4)} &= -\frac{2(1-k)(1+3k)^k}{(1+k)^4} z_i^{-\frac{2(1+2k)}{1+k}}, \text{ se } k < -1/2.
\end{aligned}$$

Os $\delta's$, para $-1 < k < 1/2$, são dados por:

$$\begin{aligned}
\delta_{(2,0,0,0,0)} &= \frac{\Gamma\left(\frac{3-k}{2}\right)}{2^{k-1}(1+k)^2\Gamma\left(\frac{k+1}{2}\right)}, & \delta_{(0,0,0,1,0)} &= \frac{k(1-k)\Gamma\left(\frac{3-k}{2}\right)}{2^{2k-1}(1+k)^4\Gamma\left(\frac{k+1}{2}\right)}, \\
\delta_{(2,1,0,0,0)} &= -\frac{(1-k)\Gamma\left(\frac{3(1-k)}{2}\right)}{2^{2k-1}(1+k)^4\Gamma\left(\frac{k+1}{2}\right)}, & \delta_{(0,0,0,1,2)} &= -\frac{k(1+3k)\Gamma\left(\frac{3-k}{2}\right)}{2^{k-2}(1+k)^4\Gamma\left(\frac{k+1}{2}\right)}, \\
\delta_{(0,0,1,0,1)} &= \frac{k\Gamma\left(\frac{3-k}{2}\right)}{2^{k-2}(1+k)^3\Gamma\left(\frac{k+1}{2}\right)}, & \delta_{(0,1,0,0,0)} &= -\delta_{(2,0,0,0,0)}, \\
\delta_{(0,0,1,0,3)} &= \frac{2k(1-k)}{(1+k)^2}, & \delta_{(1,1,0,0,1)} &= \frac{(1-k)\Gamma\left(\frac{3-k}{2}\right)}{2^{k-1}(1+k)^3\Gamma\left(\frac{k+1}{2}\right)}, \\
\delta_{(4,0,0,0,2)} &= \frac{\Gamma\left(\frac{7-k}{2}\right)}{2^{k-3}(1+k)^4\Gamma\left(\frac{k+1}{2}\right)}, & \delta_{(2,0,0,0,2)} &= \frac{k+3}{k+1}, \\
\delta_{(1,0,1,0,0)} &= \frac{k(k-1)\Gamma\left(\frac{1-3k}{2}\right)}{2^{k-1}(1+k)^4\Gamma\left(\frac{k+1}{2}\right)}, & \delta_{(0,1,0,0,2)} &= -\frac{1-k}{k+1}, \\
\delta_{(2,1,0,0,2)} &= \frac{(k-1)\Gamma\left(\frac{5-k}{2}\right)}{2^{k-2}(1+k)^4\Gamma\left(\frac{k+1}{2}\right)}, & \delta_{(3,0,0,0,1)} &= -\frac{\Gamma\left(\frac{5-k}{2}\right)}{2^{k-2}(1+k)^3\Gamma\left(\frac{k+1}{2}\right)}, \\
\delta_{(3,0,0,0,3)} &= -\frac{8\Gamma\left(\frac{7+k}{2}\right)}{(1+k)^3\Gamma\left(\frac{k+1}{2}\right)}, & \delta_{(4,0,0,0,4)} &= \frac{16\Gamma\left(\frac{9+k}{2}\right)}{(1+k)^4\Gamma\left(\frac{k+1}{2}\right)}, \\
\delta_{(0,0,0,1,4)} &= -\frac{2(1-k)k(1+3k)}{(1+k)^3}.
\end{aligned}$$

Bibliografia

- [1] Arellano-Valle, R. B. (1994). Distribuições Elípticas: Propriedades, Inferência e Aplicações a Modelos de Regressão. Tese de Doutorado, Programa de pós-graduação em Estatística, Universidade de São Paulo-*USP*.
- [2] Barndorff-Nielsen, O. E. (1983). On a formula to the distribution of the maximum likelihood estimator. *Biometrika*, **70**, 343–365.
- [3] 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.
- [4] Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London A*, **160**, 268–282.
- [5] 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*.

- [6] Cordeiro, G. M. (1983). Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistical Society B*, **45**, 404–413.
- [7] Cordeiro, G. M. (1987). On the corrections to the likelihood ratio statistics. *Biometrika*, **74**, 265–274.
- [8] 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 & Probability*. **46**, 317–328.
- [9] Cordeiro, G. M. (2004). Bartlett adjustments for overdispersed generalised linear models. *Communications in Statistics - Theory and Methods* **35**, 973–952.
- [10] Cox, D. R., Hinkley, D. V. (1974). Theoretical Statistics. *Chapman and Hall*, London.
- [11] Cox, D.R., Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*, **49**, 1–39.
- [12] Cox, D. R., Reid, N. (1992). A note on the difference between profile and modified profile likelihood. *Biometrika*, **79**, 408–411.
- [13] Cribari-Neto, F., Cordeiro, G. M. (1996). On Bartlett and Bartlett-type corrections. *Econometric reviews*, **15**, 339–367.
- [14] Cribari-Neto, F., Ferrari, S. L. P. (1995). Bartlett corrected tests for heteroskedastic linear models. *Economics Letters*, **48**, 113–118.
- [15] Cribari-Neto, F., Zarkos, S. G. (1995). Improved test statistics for multivariate regression. *Economics Letters*, **49**, 113–120.

- [16] Cysneiros, A. H. M. A., Ferrari, S. L. P. (2006). An improved likelihood ratio test for varying dispersion in exponential family nonlinear models. *Statistics & Probability Letters*, **76**, 255–265.
- [17] Cysneiros, A.H.M.A., Rodrigues, K. S. P., Cordeiro, G. M., Ferrari, S. L. P. (2010). Three-Bartlett-type corrections for Score statistics in Symmetric nonlinear regression models. *Statistical papers*, **51**, 273–284.
- [18] 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.
- [19] Cysneiros, F. J. A., Paula, G. A., Galea, M. (2005). Modelos simétricos aplicados. *ABE*, XI Escola de Modelos de Regressão, São Paulo.
- [20] 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.
- [21] Doornik, J. A. (2006). Ox: An Object-Oriented Matrix Programming Language. *Timberlake Consultants Ltd*, London, 4 ed..
- [22] Fang, K. T., Anderson, T. W. (1990). Statistical Inference in Elliptical Contoured and Related Distributions. *Allerton Press*, New York.
- [23] Fang, K. L., Kotz, S., Ng K. W. (1990). Symmetric Multivariate and Related Distributions, *Chapman and Hall*, London.
- [24] Ferrari, S. L. P., Cribari-Neto, F. (2002). Corrected modified profile likelihood heteroskedasticity tests. *Statistics & Probability*, **57**, 353–361.
- [25] Ferrari, S. L. P., Cysneiros, A. H. M. A., Cribari-Neto, F. (2004). An improved test for heteroskedasticity using adjusted modified profile likelihood inference. *Journal of Statistical Planning and Inferences*, **57**, 353–361.

- [26] Ferrari, S. L. P., Lucambio, F., Cribari-Neto, F. (2005). Improve Profile Likelihood Inference. *Journal of Statistical Planning and Inference*, **134**, 2, 373–391.
- [27] 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.
- [28] Lange, K. L., Little, R. J., Taylor, J. (1989). Robust statistical modeling using the t-distribution. *Journal of the American Statistical Association*, **84**, 881–896.
- [29] Lawley, D. N. (1956). A general method for approximating to the distribution of the likelihood ratio criteria. *Biometrika*, **71**, 233–244.
- [30] Little, R. J. (1988). Robust estimation of the mean e covariance matrix from data with missing values. *Applied Statistics*, **37**, 23–39.
- [31] McCullagh, P., Tibishirani, R. (1990). A simple method for the adjustment of profile likelihood. *Journal of the Royal Statistical Society B*, **52**, 325–344.
- [32] 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*.
- [33] Nocedal, J., Wright, S. J. (1999). Numerical Optimization. *Springer Verlag*, New York.
- [34] Pace, L., Savan, A. (1997). Principles of Statistical Inference from a Neo-Fisherian Perspective. *World Scientific*, Singapore.
- [35] Rao, C. R. (1973). Linear Statistical Inference and its Applications. 2.ed. *Wiley*, New York

- [36] Severini, T. A. (2000). Likelihood Methods in Statistics. *Oxford University Press*, Oxford.
- [37] 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.
- [38] Uribe-Opazo, M. A. (1997). Aperfeiçoamento de testes estatísticos em várias famílias de distribuições. Tese de Doutorado, Programa de pós-graduação em Estatística, Universidade de São Paulo-*USP*.
- [39] Yamaguchi, K. (1990). Generalized EM algorithm for model with contaminated error term. *Proceedings of the Seven Japan and Korea Joint Conference of Statistics*, 107–114.