
APERFEIÇOAMENTO DE TESTES DE HIPÓTESES PARA MODELOS NÃO-LINEARES SIMÉTRICOS

KATYA SILENE PORTO RODRIGUES BRAGA

Orientadora: Prof. Dra. Audrey Helen Mariz de Aquino Cysneiros

Co-orientador: Prof. Dr. Francisco Cribari Neto

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 2007

Braga, Katya Silene Porto Rodrigues.

Aperfeiçoamento de testes de hipóteses para modelos não-lineares simétricos / Kátya Silene Porto Rodrigues Braga – Recife : O autor, 2007.

v,152 folhas: il., fig., tab., quadros.

Dissertação (mestrado) – Universidade Federal de Pernambuco. CCEN. Estatística, 2007.

Inclui bibliografia e apêndices.

1. Bootstrap. 2. Correção de Bartlett 3. Correção tipo Bartlett 4. Modelos não-lineares simétricos I. Título.

519.54

CDD (22.ed.)

MEI2007-012

Universidade Federal de Pernambuco
Pós-Graduação em Estatística

08 de fevereiro de 2007

(data)

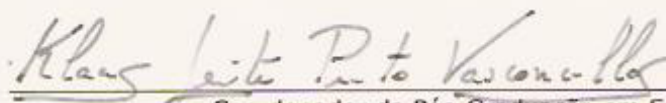
Nós recomendamos que a dissertação de mestrado de autoria de

Katya Silene Porto Rodrigues Braga

intitulada

"Aperfeiçoamento de testes de hipóteses em modelos de regressão não-lineares simétricos"

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



Lúcia Pereira Barroso (USP)



Francisco José de Azevedo Cysneiros

Este documento será anexado à versão final da dissertação.

Agradecimentos

A Deus, pelo dom da vida, por sempre iluminar meus caminhos e permitir a conclusão dessa etapa.

Ao meu esposo, Braga Júnior, pelo amor, apoio, compreensão e disponibilidade, que foram indispensáveis nesta fase da minha vida.

Aos meus pais, que me deram princípios básicos e fundamentais para a minha formação moral e tudo que uma filha precisa para viver feliz.

Ao meu irmão, Osny, pela alegria e confiança que sempre depositou em mim.

À minha orientadora, Audrey Helen Mariz de Aquino Cysneiros, pela confiança, orientação segura e disponibilidade, sempre que precisei.

Ao meu co-orientador, Francisco Cribari Neto, pelo profissionalismo e dedicação por aquilo que faz.

Aos professores do Programa de Mestrado em Estatística da UFPE, pela ajuda e incentivo.

À Valéria Bittencourt, pela amizade, extrema eficiência e organização.

Aos colegas da minha turma de mestrado, Ênio, Hemílio, Geraldo, Leonardo, Marcelo, Marina, Murilo, Rejane e Themis, que, de alguma forma, com carinho, com incentivo, alegria, me ajudaram nesta fase de minha vida.

Às minhas amigas, Lilian e Kerúcia, que sempre estiveram do meu lado, dando apoio, “brincas”, ou um sorriso amigo.

Aos colegas da turma 2006, pela agradável convivência e momentos de descontração.

A FACEPE, pelo apoio financeiro.

Resumo

Neste trabalho, tratamos de refinamentos para testes de hipóteses em modelos de regressão não-lineares simétricos, assumindo que o parâmetro de escala é desconhecido. Apresentamos, em notação matricial, fatores de correção de Bartlett e tipo-Bartlett para melhorar as estatísticas da razão de verossimilhanças e escore, respectivamente, nesta classe de modelos. Apresentamos, também, versões bootstrap dos testes da razão de verossimilhanças e escore. Os resultados de simulação mostram que os testes corrigidos e de bootstrap têm melhores desempenhos que os testes da razão de verossimilhanças e o teste escore originais em amostras pequenas.

Palavras-chave: Bootstrap, Correção de Bartlett, Correção tipo-Bartlett, Modelos de não-lineares simétricos, Teste Escore, Teste da razão de verossimilhanças.

Abstract

We consider the issue of performing inference through hypothesis tests in the class of symmetric nonlinear regression models when the scale parameter is unknown. We present, in matrix form, closed-form expressions for Bartlett and Bartlett-type correction factors for likelihood ratio and score tests, respectively. We also consider bootstrap-based inference. The Monte Carlo simulation results show that both the analytically corrected tests and the bootstrap tests outperform the standard likelihood ratio and score tests original in small samples.

Keywords: Bartlett correction, Bartlett-type correction, Bootstrap, likelihood ratio test, score test, symmetric nonlinear models.

Índice

1. Introdução	01
1.1. Introdução	01
1.2. Organização da Dissertação	02
1.3. Estimação por máxima verossimilhança	02
1.4. Otimização não-linear em estatística	04
1.4.1. Algoritmo de Newton-Raphson	04
1.4.2. Algoritmo Escore de Fisher	05
1.4.3. Método BFGS	05
1.5. Identidades de Bartlett	06
1.6. Suporte computacional	07
2. Modelos de regressão não-lineares simétricos	09
2.1. Introdução	09
2.2. Definição	10
2.3. Estimação por máxima verossimilhança	22
2.4. Testes da razão de verossimilhanças e escore em modelos não-lineares da família simétrica	26
2.4.1. Testes de hipótese sobre o parâmetro β com ϕ conhecido	26
2.4.2. Testes de hipótese sobre o parâmetro β com ϕ desconhecido	28
2.5. Testes da razão de verossimilhanças e escore bootstrap	29
3. Ajuste para o teste da razão de verossimilhanças e para o teste escore ..	31
3.1. Introdução	31
3.2. Correção de Bartlett para a estatística da razão de verossimilhanças	34
3.2.1. Correção de Bartlett em modelos não-lineares simétricos com ϕ conhecido	36
3.2.2. Correção de Bartlett em modelos não-lineares simétricos com ϕ desconhecido	36
3.3. Correção tipo Bartlett para a estatística escore	47
3.3.1. Correção tipo-Bartlett sobre o parâmetro β com ϕ conhecido	49
3.3.2. Correção tipo-Bartlett sobre o parâmetro β com ϕ desconhecido	51
4. Resultados Numéricos	57
5. Conclusão	82
o Apêndice	
o Referências	

Introdução

1.1 Introdução

O modelo de regressão linear é amplamente utilizado em aplicações práticas com o objetivo de relacionar uma variável resposta a uma ou mais variáveis explicativas. Esse modelo supõe que os erros, que representam seu componente aleatório, têm variância constante, propriedade conhecida como homoscedasticidade. A suposição de normalidade sempre foi muito atrativa para os modelos de regressão com resposta contínua e, mesmo quando não era alcançada, tentava-se algum tipo de transformação no sentido de obter a normalidade procurada. É conhecido, na literatura, que a modelagem sob a suposição de erros normalmente distribuídos pode ser altamente influenciada por observações extremas, comumente chamadas de observações aberrantes. Em muitas situações da modelagem estatística há a necessidade da procura de modelos menos sensíveis a observações aberrantes, incentivando o desenvolvimento de metodologias robustas à presença de tais observações. Na linha de modelos robustos, alternativas à suposição de erros normais têm sido propostas na literatura. Uma dessas alternativas é assumir para os erros distribuições com caudas mais pesadas do que a normal, a fim de reduzir a influência de pontos aberrantes. Na última década, diversos resultados de natureza teórica e aplicada surgiram como alternativa à modelagem com erros normais como, por exemplo, o uso de distribuições simétricas. Grande parte desses resultados podem ser encontrados em Fang, Kotz e Ng (1990) e Fang e Anderson (1990).

A idéia desta dissertação é a obtenção do fator de correção tipo-Bartlett para a estatística score nos modelos não-lineares simétricos e comparar os desempenhos dos testes score, sua versão corrigida com os testes da razão de verossimilhanças e sua versão corrigida via correção de Bartlett publicado em Cordeiro (2004). No decorrer do estudo, quando começamos a desenvolver os cálculos dos cumulantes que são valores esperados de produtos de derivadas do logaritmo da função de verossimilhança, percebemos que alguns cumulantes estavam diferentes. Em comunicação com o autor, ele nos afirmou que realmente alguns cumulantes continham erro e que concordava com os nossos resultados. Esta dissertação tem três objetivos principais. O primeiro é obter a ex-

pressão correta para o fator de correção de Bartlett para a estatística da razão de verossimilhanças nos modelos não-lineares simétricos com parâmetro de escala desconhecido. O segundo objetivo é obter a expressão para o fator de correção tipo-Bartlett para a estatística escore nos modelos não-lineares simétricos com parâmetro de escala desconhecido. O terceiro objetivo é avaliar numericamente o comportamento dos testes da razão de verossimilhanças e escore com suas respectivas versões corrigidas via correção de Bartlett e tipo-Bartlett e versões bootstrap nos modelos de regressão não-lineares simétricos.

1.2 Organização da Dissertação

A presente dissertação está dividida em cinco capítulos e cinco apêndices. O segundo capítulo é dedicado à inferência estatística em modelos de regressão não-lineares com erros tendo distribuição simétrica. Apresentamos algumas propriedades e resultados para modelos simétricos em geral. Em particular, mostramos como ficam os testes da razão de verossimilhanças e escore e suas respectivas versões bootstrap nesta classe de modelos.

No terceiro capítulo, apresentamos as fórmulas das correções de Bartlett para a estatística da razão de verossimilhanças e tipo-Bartlett para a estatística escore sobre os parâmetros de regressão em modelos não-lineares simétricos, considerando os casos em que o parâmetro de dispersão é conhecido e desconhecido. As fórmulas das correções são dadas em notação matricial e podem ser facilmente implementadas em sistemas de computação algébrica para se obter expressões em forma fechada quando aplicadas a modelos especiais.

No quarto capítulo, apresentamos resultados de simulação de Monte Carlo para avaliar os desempenhos dos testes da razão de verossimilhanças e escore, de suas respectivas versões corrigidas via correção de Bartlett e tipo-Bartlett e das versões bootstrap nos modelos de regressão não-lineares simétricos em amostras finitas. Finalmente, no quinto capítulo, são feitas algumas considerações finais.

1.3 Estimação por Máxima Verossimilhança

Seja $y = (y_1, \dots, y_n)^\top$ o valor observado de uma variável aleatória $Y = (Y_1, \dots, Y_n)$ caracterizada por uma função de probabilidade ou densidade com forma analítica $f(y; \theta)$ conhecida mas envolvendo um vetor $\theta = (\theta_1, \dots, \theta_p)^\top$ de

parâmetros desconhecidos. Seja $\Theta \subseteq \mathbb{R}^p$ o espaço paramétrico representando o conjunto de valores possíveis do vetor θ . A função $f(y; \theta)$ é denominada função do modelo estatístico e define alguma família $\mathcal{F}$ de distribuições de probabilidade.

A função de verossimilhança $L(\theta)$ é definida como sendo igual à função do modelo, embora seja interpretada diferentemente como a função de θ para y conhecido. Isto é, a função de verossimilhança para θ baseada na observação $Y = y$ é dada por

$$L(\theta) = L(\theta, y) = f(y; \theta), \quad \theta \in \Theta.$$

Muito freqüentemente, as componentes de Y são mutuamente independentes para todas as distribuições em $\mathcal{F}$ e a função de verossimilhança de θ pode ser escrita como

$$L(\theta) = \prod_{l=1}^n f_l(y_l; \theta),$$

em que f_l corresponde à densidade individual da l -ésima observação. Usualmente, trabalha-se com o logaritmo da função de verossimilhança:

$$\ell(\theta) = \ell(\theta, y) = \log(L(\theta)) = \sum_{l=1}^n \log\{f_l(y_l; \theta)\}.$$

Dada a função $L(\cdot)$, a estimativa de máxima verossimilhança (EMV) $\hat{\theta}$ de θ é o valor que maximiza $L(\theta)$ em Θ , isto é, $L(\hat{\theta}) \geq L(\theta)$ para todo $\theta \in \Theta$.

Como a função logaritmo é monótona crescente, maximizar $L(\theta)$ e $\ell(\theta)$ em Θ são processos equivalentes. Então, a EMV é definida de modo que, para todo $\theta \in \Theta$, $\ell(\hat{\theta}) \geq \ell(\theta)$.

A função escore $U(\theta) = (U(\theta_1), \dots, U(\theta_p))^T$ é definida como

$$U_r(\theta) = \frac{\partial \ell(\theta)}{\partial \theta_r}, \quad r = 1, \dots, p,$$

e descreve como o logaritmo da função de verossimilhança varia em Θ .

Se Θ é um conjunto discreto, calcula-se $\ell(\theta)$ para os diversos valores de θ 's e escolhe-se $\hat{\theta}$ como aquele valor de θ correspondente ao máximo $\ell(\theta)$. Quando $\ell(\theta)$ é contínua e diferenciável em Θ , a EMV $\hat{\theta}$ pode ser obtida resolvendo-se o sistema de equações simultâneas $\frac{\partial \ell(\theta)}{\partial \theta_r} = 0$ para $r = 1, \dots, p$, desde que θ não se encontre na fronteira do espaço paramétrico.

A matriz de primeiras derivadas da função escore com sinal negativo

$$J(\theta) = -\frac{\partial U(\theta)^T}{\partial \theta} = -\frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^T}$$

é denominada matriz de informação observada (Fisher, 1925). A matriz hessiana é simplesmente $-J(\theta)$ e sob condições gerais de regularidade, ver Cox e Hinkley (1974, Cap. 9) e Lehmann (1999, Cap. 7), temos que $E_\theta[J(\theta)] = K(\theta)$, em que o subscrito θ indica que a esperança matemática é calculada supondo que θ é o verdadeiro valor do parâmetro. A $K(\theta)$ é chamada de matriz de informação (esperada) de Fisher para θ . Para $\hat{\theta}$ ser o máximo local as condições $\hat{U} = U(\hat{\theta}) = 0$ e $\hat{J} = J(\hat{\theta}) \geq 0$ ($\hat{J}$ positiva semi-definida) são necessárias, enquanto que $\hat{U} = 0$ e $\hat{J} > 0$ ($\hat{J}$ positiva definida) são suficientes.

Lehmann e Casela (1998, Cap. 6) mostram que, sob condições de regularidade, o estimador de máxima verossimilhança $\hat{\theta}$, obtido como raiz única da equação, apresenta as seguintes propriedades:

- (i) É consistente;
- (ii) É assintoticamente eficiente, isto é, dentre todos os estimadores consistentes de θ , o estimador de máxima verossimilhança possui variância assintótica mínima;
- (iii) É invariante sob transformação biunívoca, ou seja, se $\hat{\theta}$ é estimador de máxima verossimilhança de θ e g é função bijetora, então $g(\hat{\theta})$ é estimador de máxima verossimilhança de $g(\theta)$;
- (iv) Tem distribuição assintótica normal p -variada, onde p é a dimensão do vetor θ , da forma

$$\hat{\theta} \sim N_p(\theta, K(\theta)^{-1});$$

- (v) É, em geral, viesado, embora seja assintoticamente não-viesado.

1.4 Otimização Não-Linear em Estatística

Um dos problemas mais freqüentes na inferência estatística é encontrar o estimador de máxima verossimilhança de algum parâmetro desconhecido. Em geral, o estimador de máxima verossimilhança $\hat{\theta}$ de θ não apresenta forma analítica fechada e as estimativas devem ser encontradas através da maximização numérica do logaritmo da função de verossimilhança. Para isto utilizamos algum procedimento iterativo.

1.4.1 Algoritmo de Newton-Raphson

Este procedimento de maximização é obtido expandindo a função escore $U(\theta)$ em série de Taylor até a primeira ordem em torno de um valor inicial $\theta^{(0)}$, tem-se aproximadamente

$$U(\theta) = U(\theta^{(0)}) + H(\theta^{(0)})(\theta - \theta^{(0)}),$$

em que

$$H(\theta^{(0)}) = \left. \frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^\top} \right|_{\theta=\theta^{(0)}}$$

é a matriz de segundas derivadas do logaritmo da função de verossimilhança. A seguir, fazendo-se $U(\theta) = 0$ e repetindo o procedimento acima, chega-se à seguinte equação:

$$\theta^{(k+1)} = \theta^{(k)} - \lambda^{(k)} H(\theta^{(k)})^{-1} U(\theta^{(k)}),$$

podendo ser generalizado da forma

$$\theta^{(k+1)} = \theta^{(k)} - \lambda^{(k)} H(\theta^{(k)})^{-1} U(\theta^{(k)}), \quad k = 0, 1, 2, \dots \quad (1.1)$$

em que o termo $\lambda^{(k)}$ é um escalar determinado por algum procedimento de busca linear a partir de $\theta^{(k)}$ na direção $-H(\theta^{(k)})^{-1}U(\theta^{(k)})$, normalmente toma-se $\lambda^{(k)} = 1$. Este processo é repetido até se obter uma distância entre $\theta^{(k+1)}$ e $\theta^{(k)}$ que seja menor que a tolerância especificada.

1.4.2 Algoritmo Escore de Fisher

Este procedimento foi sugerido por Sir Ronald Fisher e baseia-se em substituir a matriz $-H$ por seu valor esperado $E(-H) = K$, a matriz de informação esperada. Assim, a equação (1.1) é reescrita na forma

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

em que o termo $\lambda^{(k)}$ é um escalar determinado por algum procedimento de busca linear a partir de $\theta^{(k)}$ na direção $K(\theta^{(k)})^{-1} U(\theta^{(k)})$.

1.4.3 Método BFGS

Este algoritmo foi desenvolvido por Broyden, Fletcher, Goldfarb e Shanno. Este método se assemelha ao método de Newton-Raphson, diferenciando-se pelo fato de substituir a matriz $-H^{-1}$ por uma sequência de matrizes simétricas e positivas definidas por $B^{(k)}$ tal que $\lim_{k \rightarrow \infty} B^{(k)} = -H^{-1}$.

A forma recursiva para se obter tais matrizes é dada pela expressão

$$B^{(k+1)} = B^{(k)} - \frac{B^{(k)}(g^{(k)})^\top B^{(k)}}{(g^{(k)})^\top B^{(k)} g^{(k)}} + \frac{h^{(k)}(h^{(k)})^\top}{(h^{(k)})^\top g^{(k)}},$$

para $k = 0, 1, \dots$, em que $g^{(k)} = \theta^{(k+1)} - \theta^{(k)}$ e $h^{(k)} = U(\theta^{(k+1)}) - U(\theta^{(k)})$. Comumente, toma-se como matriz inicial, $B^{(0)}$, a matriz identidade de mesma ordem, pois ela é positiva definida e simétrica, conduzindo assim a aproximações $B^{(k)}$ positivas definidas e simétricas. Assim, a equação (1.1) fica

$$\theta^{(k+1)} = \theta^{(k)} - \lambda^{(k)} B^{(k)} U(\theta^{(k)}), \quad k = 0, 1, \dots,$$

em que o termo $\lambda^{(k)}$ é um escalar determinado por algum procedimento de busca linear a partir de $\theta^{(k)}$ na direção $-B^{(k)} U(\theta^{(k)})$.

1.5 Identidades de Bartlett

Utilizaremos a notação de Lawley (1956) para as derivadas do logaritmo da função de verossimilhança, em que todos os índices variam de $1, \dots, p$, onde p é a dimensão do vetor θ : $U_r = \partial \ell(\theta) / \partial \theta_r$, $U_{rs} = \partial^2 \ell(\theta) / \partial \theta_r \partial \theta_s$, $U_{r,s} = \partial \ell(\theta) / \partial \theta_r \partial \ell(\theta) / \partial \theta_s$, etc. Os momentos conjuntos de derivadas de $\ell(\theta)$

são denotados por: $\mu_r = E[U_r]$, $\mu_{rs} = E[U_{rs}]$, $\mu_{r,s} = E[U_{r,s}]$, $\mu_{r,st} = E[U_{r,st}]$ e assim por diante. Como $\mu_r = 0$, os correspondentes cumulantes conjuntos expressos em termos dos momentos são $\kappa_{r,s} = \mu_{r,s}$, $\kappa_{rs} = \mu_{rs}$, $\kappa_{rs,t} = \mu_{rs,t}$, $\kappa_{rs,tu} = \mu_{rs,tu} - \mu_{rs}\mu_{tu}$, $\kappa_{r,s,t} = \mu_{r,s,t}$ e $\kappa_{r,s,t,u} = \mu_{r,s,t,u} - \sum_{(3)} \kappa_{r,s}\kappa_{t,u}$ onde $\sum_{(k)}$ representa o somatório sobre todas as k combinações de índices, etc. Ainda, $-\kappa^{rs} = \kappa^{r,s}$ representa o elemento (r, s) da inversa da matriz de informação esperada $K(\theta)^{-1}$. As derivadas dos cumulantes são representadas com sobrescritos: $\kappa_{rs}^{(t)} = \partial \kappa_{rs} / \partial \theta_t$, $\kappa_{rs}^{(tu)} = \partial^2 \kappa_{rs} / \partial \theta_t \partial \theta_u$, $\kappa_{rst}^{(u)} = \partial \kappa_{rst} / \partial \theta_u$, etc. Os momentos e cumulantes acima não são independentes, porém satisfazem certas equações que facilitam seus cálculos. Estas equações, que representam condições de regularidade, são denominadas de identidades de Bartlett. As mais importantes são $\kappa_r = 0$ e $\kappa_{rs} + \kappa_{r,s} = 0$. Os cumulantes κ 's referem-se a um total sobre a amostra e, em geral, são de ordem $O(n)^1$.

A idéia central na dedução das identidades de Bartlett é a validade em problemas regulares da fórmula

$$\frac{\partial}{\partial \theta} E[T(Y)] = \int T(y) \frac{\partial}{\partial \theta} f(y; \theta) dy, \quad (1.2)$$

para qualquer estatística $T(Y)$, isto é, pode-se inverter a ordem das operações de diferenciação em relação a θ e integração com respeito a y . As identidades de Bartlett são obtidas expressando a equação (1.2) em termos de momentos e diferenciando-a sucessivamente com relação às componentes de θ . Da definição da função escore tem-se $U_r = L_r/L$, onde $L_r = \partial L / \partial \theta_r$. Diferenciando $\int L dy = 1$ em relação a θ_r vem $\int L_r dy = 0$ e, então, $\kappa_r = E(U_r) = 0$. Diferenciando a última integral em relação a θ_s , encontra-se $\int (U_{rs}L + U_r U_s L) dy = 0$, ou seja, $\kappa_{rs} + \kappa_{r,s} = 0$. Outras identidades são deduzidas de forma análoga. Temos

$$\begin{aligned} \kappa_{r,st} + \kappa_{rst} - \kappa_{st}^{(r)} &= 0, \\ \kappa_{r,s,t} - 2\kappa_{rst} + \sum_{(3)} \kappa_{rs}^{(t)} &= 0, \\ \kappa_{r,s,t}^{(u)} &= \kappa_{rst,u}, \\ \kappa_{r,stu} + \kappa_{rstu} - \kappa_{stu}^{(r)} &= 0, \end{aligned}$$

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

$$\begin{aligned}
\kappa_{rstu} + \sum_{(4)} \kappa_{r,stu} + \sum_{(3)} \kappa_{rs,tu} + \sum_{(6)} \kappa_{r,s,tu} + \kappa_{r,s,t,u} &= 0, \\
\kappa_{r,s,t,u} &= -3\kappa_{rstu} + 2\sum_{(4)} \kappa_{rst}^{(u)} + \sum_{(6)} \kappa_{rs}^{(tu)} + \sum_{(3)} \kappa_{rs,tu}, \\
\kappa_{r,s,tu} &= \kappa_{rstu} - 2\kappa_{rtu}^{(s)} + \kappa_{tu}^{(rs)} - \kappa_{rs,tu}, \text{ etc.}
\end{aligned}$$

A grande vantagem das identidades de Bartlett é facilitar a obtenção dos cumulantes κ 's, já que sob determinada parametrização pode conduzir a um cálculo simples de alguns cumulantes, sendo os demais feitos indiretamente através destas identidades. Esses cumulantes têm grande aplicabilidade no cálculo de correções de Bartlett para a estatística da razão de verossimilhanças e tipo Bartlett para a estatística score como será visto adiante.

1.6 Suporte computacional

Todas as simulações apresentadas na dissertação foram desenvolvidas a partir de programas construídos usando a linguagem de programação matricial **Ox** em sua versão 4.02 para sistema operacional Windows. Esta linguagem de programação foi criada por Jurgen Doornik em 1994 na Universidade de Oxford, Inglaterra. Esta é uma linguagem bastante flexível, tendo sido desenvolvida com base na linguagem de programação C. **Ox** é distribuída gratuitamente para uso acadêmico e se encontra disponível em <http://www.doornik.com>. Maiores detalhes sobre esta linguagem de programação podem ser encontrados em Doornik (2001).

As maximizações não-lineares necessárias para o cálculo das estimativas de máxima verossimilhança foram feitas utilizando o algoritmo quasi-Newton BFGS (ver Nocedal e Wright, 1999, Capítulo 8) e o algoritmo SQP (ver Nocedal e Wright, 1999, Capítulo 18), disponível em funções pré-definidas da linguagem **Ox**.

Para a elaboração do texto desta dissertação utilizamos o sistema de tipografia (Plain) **T_EX**. A opção por esta linguagem fundamenta-se principalmente em dois pontos: flexibilidade e qualidade da apresentação. A versão mais atual do **T_EX** para a plataforma Windows pode ser obtida gratuitamente em <http://www.miktex.org>. Para maiores detalhes, ver Knuth (1986).

Os gráficos apresentados foram produzidos utilizando o ambiente de programação **R** em sua versão 2.4.1 para o sistema operacional Windows. Esta linguagem foi criada por Ross Ihaka e Robert Gentleman na Universidade de Auck-

land com o objetivo de produzir um ambiente de programação parecido com o S. O R se encontra disponível gratuitamente em <http://www.r-project.org>. Para maiores detalhes, ver Ihaka e Gentleman (1996).

Modelo de regressão não-linear usando uma família simétrica

2.1 Introdução

Em nosso trabalho, nos concentramos na família de distribuições simétricas com suporte na reta real. Esta família forma uma classe geral de distribuições com a mesma simetria que a distribuição normal padrão. As distribuições *t*-Student, Cauchy, *t*-Student generalizada, Kotz, Kotz generalizada, Exponencial potência, Logística tipos I e II, entre outras, pertencem a esta classe como veremos a seguir.

A classe de distribuições simétricas tem recebido crescente atenção na literatura estatística nos últimos anos pois ela assume para os erros distribuições com caudas mais pesadas do que a normal, a fim de reduzir a influência de pontos aberrantes. Uma revisão de diferentes áreas em que distribuições simétricas são aplicadas pode ser encontrada em Chmielewski (1981), Fang, Kotz e Ng (1990), Fang e Zhang (1990), Fang e Anderson (1990) e Gupta e Varga, (1993). Nessa linha podemos citar Lange, Little e Taylor (1989), que propõem o modelo baseado na suposição de erros *t*-Student enquanto Little (1988) e Yamaguchi (1990) utilizam a distribuição normal contaminada. Em ambos os modelos incorporam-se parâmetros adicionais, os quais permitem ajustar a curtose da distribuição dos dados. No caso da *t*-Student, os graus de liberdade são usados para controlar a curtose. Taylor (1992) propõe o ajuste de um modelo de regressão linear supondo erros distribuídos como exponencial potência com um parâmetro extra de forma. Arellano-Valle (1994) apresenta vários resultados sobre regressão usando a distribuição *t*-Student. Ferrari e Uribe-Opazo (2001) estendem esses resultados para modelos de regressão lineares simétricos. Cordeiro, Ferrari, Uribe-Opazo e Vasconcellos (2000) obtiveram a correção do viés do estimador de máxima verossimilhança na classe de modelos não-lineares simétricos. Cordeiro (2004) desenvolveu correção de Bartlett para os modelos de regressão não-lineares simétricos, generalizando assim os resultados de Ferrari e Uribe-Opazo (2001).

Neste capítulo, apresentamos a classe de modelos de regressão não-lineares

simétricos e quatro testes de hipóteses são apresentados: da razão de verossimilhanças e escore e suas respectivas versões bootstrap. Na Subseção 2.2, apresentamos a definição, propriedades destes modelos e características de algumas distribuições simétricas. Posteriormente, na Subseção 2.3, apresentamos o método de estimação dos parâmetros desconhecidos. Na Subseção 2.4, apresentamos os testes baseados nas estatísticas da razão de verossimilhança e escore supondo inicialmente que o parâmetro de escala é conhecido e posteriormente supondo que o mesmo é desconhecido. Finalmente, na Subseção 2.5, versões de bootstrap para os testes da razão de verossimilhanças e escore, as quais denominaremos teste da razão de verossimilhanças bootstrap e teste escore bootstrap, são apresentados.

2.2 Definição

Dizemos que a variável aleatória Y tem distribuição simétrica, com suporte em $\mathbb{R}$, com parâmetros de locação $\mu \in \mathbb{R}$ e de escala $\phi > 0$, se sua função densidade é da forma

$$\pi(y; \mu, \phi) = \frac{1}{\phi} h\left(\left(\frac{y - \mu}{\phi}\right)^2\right), \quad y \in \mathbb{R}, \quad (2.1)$$

para alguma função $h(\cdot)$ denominada função geradora de densidade, com $h(u) > 0$, para $u > 0$ e $\int_0^\infty u^{-1/2} h(u) du = 1$. Essa condição é necessária para que $f(y; \mu, \phi)$ seja uma função densidade de probabilidade. Denotamos $Y \sim S(\mu, \phi^2)$ e denominaremos Y de variável aleatória simétrica.

Vale ressaltar que algumas propriedades da distribuição normal podem ser estendidas para a classe simétrica de distribuições.

Podemos ver que, se $Y \sim S(\mu, \phi^2)$, então a sua função característica tem a forma

$$\psi_y(t) = e^{it\mu} \varphi(t^2 \phi^2), \quad t \in \mathbb{R},$$

para alguma função φ , com $\varphi(u) \in \mathbb{R}$, para $u > 0$. Quando existem, $E(Y_i) = \mu_i$ e $Var(Y_i) = \xi \phi^2$, em que $\xi > 0$ é uma constante dada por $\xi = -2\varphi'(0)$, com $\varphi'(0) = \frac{d\varphi(u)}{du}|_{u=0}$ e que não depende dos parâmetros μ e ϕ (Fang, Kotz e Ng, 1990, pg.43). Kelker (1970) nota que, se $u^{\frac{1}{2}(k+1)} h(u)$ é integrável, então o k -ésimo momento de Y existe. Temos também, que se $Y \sim S(\mu, \phi^2)$ então $a + bY \sim S(a + b\mu, b^2 \phi^2)$, onde a e $b \in \mathbb{R}$, com $b \neq 0$, ou seja, a distribuição de qualquer transformação linear de uma variável aleatória com distribuição

simétrica é também simétrica. Como exemplo, se $Y \sim S(\mu, \phi^2)$, então $Z = \frac{(Y-\mu)}{\phi} \sim S(0, 1)$ com função densidade $f(z) = f(z; 0, 1) = h(z^2)$, $z \in \mathbb{R}$; e chamaremos Z de simétrica padrão.

Considerando o caso em que $Z \sim S(0, 1)$ e assumindo que seus momentos existem, Berkane e Bentler (1986) mostram que a função característica de z pode ser expandida como

$$\psi_z(t) = \sum_{k=0}^{\infty} i^k \mu'_k \frac{t^k}{k!},$$

em que $\mu'_k = E(y^k) = i^{-k} \psi_z^{(k)}(0)$, sendo que $\psi_z^{(k)}(0)$ é a k -ésima derivada de $\psi_z(t)$ avaliada em $t = 0$. Então,

$$\mu'_k = \begin{cases} 0, & \text{para } k \text{ ímpar;} \\ \frac{(2m)!}{2^m m!} (\mu'_2)^m (k(m) + 1), & \text{para } k = 2m, \quad m = 1, 2, \end{cases}$$

em que

$$k(m) = \frac{\varphi^{(m)}(0)}{(\varphi^{(1)}(0))^m} - 1,$$

sendo $\varphi^{(r)}(0)$ a r -ésima derivada da função φ avaliada em zero. Os coeficientes $k(m)$, $m = 1, 2, \dots$, são conhecidos como parâmetros de momentos e generalizam o coeficiente de curtose $\gamma_2 = 3(k(2) + 1)$ de uma distribuição $S(\mu, \phi^2)$ (Muirhead, 1982). Cambanis, Huang e Simons (1981) observaram que a família de distribuições simétricas coincide com a classe de distribuições elípticas univariadas. A partir dos trabalhos de Kelker (1970) vários artigos foram publicados sobre distribuições elípticas univariadas e multivariadas. Podemos citar alguns trabalhos que discutem propriedades dessas distribuições: Berkane e Bertler (1986), Muirhead (1980, 1982), Rao (1990), Cambanis, Huang e Simons (1981) e Anderson e Fang (1987). A classe de distribuições simétricas tem recebido crescente atenção na literatura estatística nos últimos anos, ver por exemplo, os livros de Fang, Kotz e Ng (1990), Fang e Anderson (1990), Fang e Zhang (1990).

Sejam $Y_1, Y_2, \dots, Y_n$ variáveis aleatórias independentes, mas não necessariamente identicamente distribuídas, tais que $Y_l \sim S(\mu_l, \phi^2)$, para $l = 1, 2, \dots, n$. Assumimos que cada Y_l tem função densidade na forma dada por (2.1) onde o parâmetro de escala é comum para todas as variáveis, mas o parâmetro de locação μ_l não é necessariamente igual à média de Y_i , por exemplo podemos

citar a distribuição Cauchy. Considere a estrutura de regressão não-linear que relaciona o parâmetro de locação μ_l com as variáveis x_l de interesse da forma

$$\mu_l = f(x_l; \beta), \quad (2.2)$$

onde $x_l = (x_{l1}, \dots, x_{lm})^\top$ é um vetor de valores conhecidos de variáveis explicativas associadas à l -ésima observação, $\beta = (\beta_1, \dots, \beta_p)^\top$ é um vetor $p \times 1$ de parâmetros desconhecidos a serem estimados, tal que $\beta \in \Omega_\beta \subset \mathbb{R}^p$, e $f(\cdot; \cdot)$ é uma função, possivelmente não-linear no segundo argumento, contínua, diferenciável com respeito aos componentes de β tal que a matriz $n \times p$ de derivadas $\tilde{X} = \tilde{X}(\beta) = \frac{\partial \mu}{\partial \beta^\top}$ com $\mu = (\mu_1, \dots, \mu_n)^\top$ tenha posto p para todo β . A matriz $\tilde{X}$ tem elementos que são em geral, funções do vetor de parâmetros β desconhecidos.

O modelo de regressão não-linear simétrico é, então, definido por (2.1) e (2.2). Esta classe de modelos de regressão estende a classe dos modelos normais lineares em duas direções. Primeiro, inclui importantes distribuições não-normais e, segundo, permite, uma estrutura possivelmente não-linear no parâmetro de locação μ . Expressando a componente sistemática na forma de uma função linear, ou seja $\mu = f(x_l; \beta) = x_l^\top \beta$, temos um modelo de regressão linear simétrico. Este caso, onde o parâmetro μ se associa linearmente com β , é bem conhecido e inclui análise de variância, regressão linear, etc, como casos particulares.

A seguir, descrevemos algumas distribuições especiais de (2.1) que são de grande importância para aplicações práticas em diversas áreas do conhecimento: ver Fang, Kotz e Ng (1990).

2.2.1 Distribuição Normal

A normal é a distribuição pertencente à classe simétrica mais utilizada, devido a todo desenvolvimento teórico e aplicado estabelecido no decorrer dos anos. Os primeiros trabalhos consideram a distribuição somente como uma aproximação conveniente para a distribuição binomial. No início do século XIX, o reconhecimento da sua importância teórica foi propagado por Laplace e por Gauss. A distribuição normal possui várias propriedades que permitem caracterizá-la dentro da classe das distribuições simétricas nas chamadas distribuições normais compostas. Alguns resultados nesse sentido podem ser vistos em Muirhead (1982) e Devlin, Gnanadesikan e Kettenring (1976). Se $Y \sim S(\mu, \phi)$ e a função

geradora de densidade $h(\cdot)$ é da forma

$$h(u) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\}, u > 0,$$

então Y tem distribuição normal denotada por $Y \sim N(\mu, \phi^2)$, e sua função característica é dada por

$$\psi_y(t) = e^{it\mu} \exp\left\{-\frac{1}{2}t^2\phi^2\right\}, t \in \mathbb{R}.$$

Se $Y \sim N(\mu, \phi^2)$, então $E(Y) = \mu$, $Var(Y) = \phi^2$, e os momentos centrais de ordem r são

$$\mu_r = E\{(Y - \mu)^r\} = \begin{cases} 0, & r \text{ ímpar}, \\ \phi^r \frac{r!}{2^{r/2}(r/2)!}, & r \text{ par}; \end{cases}$$

portanto, o coeficiente de curtose é $\gamma_2 = 3$.

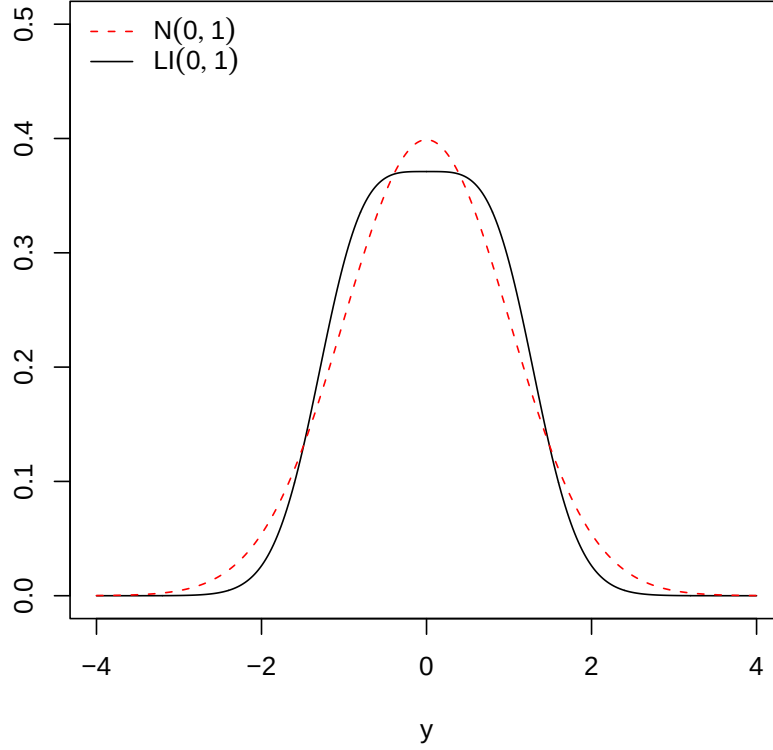
2.2.2 Distribuição Logística I

Dizemos que a variável aleatória $Y \sim S(\mu, \phi^2)$ tem distribuição logística I se a função geradora de densidade $h(\cdot)$ tem a forma

$$h(u) = c \frac{e^{-u}}{[1 + e^{-u}]^2}, u > 0,$$

em que c é a constante normalizadora obtida da relação $\int_0^\infty u^{-1/2} h(u) du = 1$; logo $c \approx 1,484300029$, e é denotada por $Y \sim LI(\mu, \phi^2)$. Temos que $E(Y) = \mu$, $Var(Y) \approx 0,79569\phi^2$ e o coeficiente de curtose $\gamma_2 \approx 2,385165$. Note que o coeficiente de curtose desta distribuição é menor do que o coeficiente de curtose da distribuição normal. Na Figura 2.1, temos os gráficos das densidades normal e logística I. Podemos observar a forma achatada no patamar da distribuição logística I que difere da distribuição normal.

Figura 2.1 Gráficos de funções densidade logística I e normal.



2.2.3 Distribuição Logística II

Dizemos que a variável aleatória $Y \sim S(\mu, \phi^2)$ tem distribuição logística II se a função geradora de densidade $h(\cdot)$ tem a forma

$$h(u) = \frac{\exp\{-u^{1/2}\}}{[1 + \exp(-u^{1/2})]^2}, u > 0,$$

e a denotamos por $Y \sim LII(\mu, \phi^2)$. A função característica é dada por

$$\psi_y(t) = \frac{2(e^{it\mu}\pi\phi t)}{e^{\pi\phi t} - e^{-\pi\phi t}}, t \in \mathbb{R}.$$

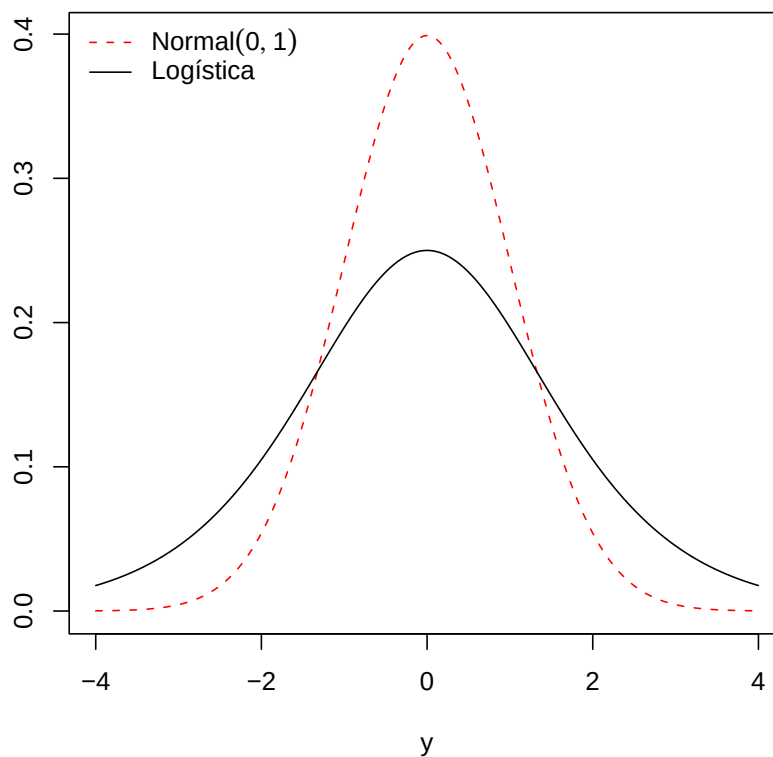
Temos que $E(Y) = \mu$, $Var(Y) = \pi^2\phi^2/3$ e o coeficiente de curtose é $\gamma_2 = 4,2$, que é maior que as curtoses das distribuições normal e logística I. Adicionalmente, a mediana e a moda são iguais à média.

Uma relação muito útil para gerar amostras aleatórias é dada por Hastings e Peacock (1975). Seja $u \sim U(0, 1)$ e $Y = \mu + \phi \log(\frac{u}{1-u})$, então $Y \sim LII(\mu, \phi^2)$.

Esta distribuição foi utilizada por Verhulst (1838, 1845) para o ajuste de curvas de crescimento demográfico. Pearl e Reed (1920, 1924), Pearl, Reed e Kish (1940) e Schultz (1930) aplicaram o modelo logístico como modelo de crescimento em populações humanas e em alguns organismos biológicos. Para aplicação em análise de sobrevivência, ver Plackett (1959) e para aplicação em modelagem de distribuição de renda, ver Fisk (1961).

Na Figura 2.2 observamos a forma da distribuição logística II e constatamos que ela tem caudas mais pesadas do que a distribuição normal.

Figura 2.2 Gráficos de funções densidade logística e normal.



2.2.4 Distribuição Cauchy

Dizemos que a variável aleatória $Y \sim S(\mu, \phi^2)$ tem distribuição de Cauchy se sua função geradora da densidade $h(\cdot)$ tem a forma

$$h(u) = \frac{1}{\pi}(1+u)^{-1}, u > 0.$$

Essa distribuição também é conhecida como distribuição de *Pearson tipo VII*. Nós a denotamos por $Y \sim C(\mu, \phi^2)$. A sua função característica é da forma

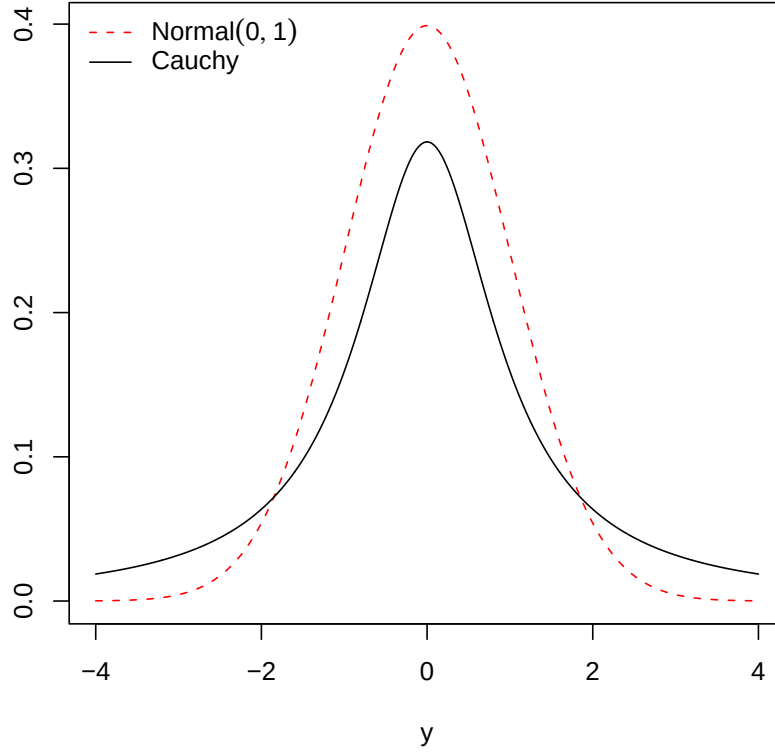
$$\psi_y(t) = \exp\{it\mu - |t|\phi\}, t \in \mathbb{R}.$$

A distribuição é simétrica em torno de μ e os pontos de inflexão da função densidade são $\mu \pm \phi\sqrt{3}$, os valores da função de distribuição acumulada nos pontos de inflexão são 0,273 e 0,723, que podem ser comparados com os correspondentes valores, 0,159 e 0,841, da distribuição normal. Em particular, a distribuição não tem momentos finitos e, portanto, não tem valor esperado nem desvio-padrão. Ela tem mediana e moda iguais a μ e os quartis superior e inferior são $\mu \pm \phi$. Note que a distribuição Cauchy possui caudas mais pesadas do que a distribuição normal; ver Figura 2.3.

Se $Y_1, Y_2, \dots, Y_n$ são variáveis aleatórias independentes onde cada Y_l possui distribuição Cauchy, então $S = \sum_{l=1}^n Y_l$ tem distribuição de Cauchy com parâmetros de locação $\mu = \sum_{l=1}^n \mu_l$ e escala $\phi = \sum_{l=1}^n \phi_l$. Um resultado interessante é que para $a_l \neq 0$, $\sum_{l=1}^n a_l Y_l$ tem distribuição de Cauchy com parâmetros de locação $\sum_{l=1}^n a_l \mu_l$ e escala $\sum_{l=1}^n |a_l| \phi_l$. Em particular, se os Y_l são i.i.d (independentes identicamente distribuídos) com $Y_l \sim C(\mu, \phi^2)$, então $\bar{Y} = \frac{1}{n} \sum_{l=1}^n Y_l \sim C(\mu, \phi^2)$. Quando os valores dos parâmetros de locação e escala são iguais a 0 e 1, respectivamente, a distribuição Cauchy passa a ser chamada como Cauchy padrão ou *t*-Student central com um grau de liberdade.

Temos, ainda, a relação $Y = \mu + \phi N_1/N_2$ em que $N_i \sim N(0, 1)$ para $i = 1, 2$, sendo N_1 e N_2 independentes, que pode ser usada para definir um gerador de números aleatórios para a distribuição de Cauchy com parâmetros (μ, ϕ^2) .

Figura 2.3 Gráficos de funções densidade cauchy e normal.



2.2.5 Distribuição t -Student

A variável aleatória $Y \sim S(\mu, \phi^2)$ tem distribuição t -Student com ν graus de liberdade se sua função geradora de densidade $h(\cdot)$ é dada por

$$h(u) = \frac{\nu^{\frac{\nu}{2}}}{B\left(\frac{1}{2}, \frac{\nu}{2}\right)} [\nu + u]^{-\frac{(\nu+1)}{2}}, u > 0, \nu > 0,$$

em que $B(\cdot, \cdot)$ é a função Beta; denotamos por $Y \sim t(\mu, \phi^2, \nu)$. A função característica em Fang, Kotz e Ng (1990)

A distribuição t é simétrica em torno de μ . Quando $\nu \rightarrow \infty$, a distribuição de y tende para a distribuição normal padrão. Quando $\nu = 1$, a distribuição reduz-se à distribuição Cauchy. A distribuição t com ν graus de liberdade foi originada através da razão $t_\nu = U(\frac{\chi_\nu^2}{\nu})^{-1/2}$, onde U é a variável aleatória normal padrão e χ_ν^2 é uma variável aleatória qui-quadrado com ν graus de liberdade, sendo ambas independentes.

Todos os momentos ordinários são da forma

$$\mu_r = \begin{cases} \frac{\nu^{\frac{r}{2}} \Gamma(\frac{1}{2}(r+1)) \Gamma(\frac{1}{2}(\nu-r))}{\Gamma(\frac{1}{2}) \Gamma(\frac{\nu}{2})}, & r \text{ par}, \\ 0, & r \text{ ímpar}, \end{cases}$$

onde $\Gamma(\cdot)$ denota a função gama. Assim, $E(Y) = 0$, se $\nu > 1$ e $Var(Y) = \frac{\nu}{\nu-2}$, se $\nu > 2$. Se $\nu \leq r$ e r é par, o momento de ordem r é infinito. O desvio médio é dado por

$$E[|Y|] = \frac{\nu^{1/2} \Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{1}{2}) \Gamma(\frac{\nu}{2})}.$$

O coeficiente de assimetria é $\gamma_3 = 0$ e o coeficiente de curtose é $\gamma_2 = 3 + \frac{6}{(\nu-4)}$ para $\nu > 4$, sendo este coeficiente maior do que o coeficiente de curtose da distribuição normal. A função densidade de Y tem pontos de inflexão em $y = \pm \sqrt{\nu/(\nu+2)}$.

A distribuição t -Student é utilizada para modelar o comportamento de dados que provêm de uma distribuição com caudas mais pesadas que a normal, permitindo reduzir a influência de observações aberrantes. Lange, Little e Taylor (1989) propõem o modelo t -Student como uma extensão paramétrica robusta do modelo normal, já que a t -Student é uma distribuição que permite ajustar a curtose da distribuição dos dados através do parâmetro ν .

Na Figura 2.4, observamos a forma da distribuição t -Student para alguns valores de ν .

2.2.6 Distribuição t -Student generalizada

Se uma variável aleatória $Y \sim S(\mu, \phi^2)$ e a função geradora de densidade $h(\cdot)$ tem a forma

$$h(u) = \frac{s^{\frac{r}{2}}}{B\left(\frac{1}{2}, \frac{r}{2}\right)} (s+u)^{-\frac{(r+1)}{2}}, u > 0, s, r > 0,$$

então, Y possui uma distribuição t -Student generalizada, o que denotamos por $Y \sim tG(\mu, \phi^2, s, r)$.

Na Figura 2.5, observamos a forma da t -Student generalizada para vários valores de r, s .

Quando $r = s = \nu$, a distribuição t -Student generalizada coincide com a distribuição t -Student com parâmetros μ, ϕ e r graus de liberdade. Quando $r = s = 1$, temos a distribuição Cauchy.

Supondo que $(Y|v = \nu) \sim N(\mu, \nu\phi^2)$ em que $v \sim GI(r/2, s/2)$, independentes com $s > 0$ e $r > 0$ podem ser não-inteiros. Temos as seguintes propriedades: $Y \sim tG(\mu, \phi^2, s, r)$, $E(Y) = \mu$, para $r > 1$, $Var(Y) = (s/(r-2))\phi^2$, para $r > 2$, e o coeficiente de curtose é $\gamma_2 = 3 + 6/(r-4)$, para $r > 4$, que é maior que o coeficiente de curtose da distribuição normal, e não depende do valor do parâmetro s .

Se definimos $y = v^{-1/2}z$, com z e v variáveis aleatórias independentes, onde $z \sim N(0, 1)$ e $v \sim GI(r/2, s/2)$, então $y \sim tG(0, 1, s, r)$. Esta propriedade é importante para gerar observações de uma distribuição t -Student generalizada.

2.2.7 Distribuição Exponencial Potência

Se uma variável aleatória $Y \sim S(\mu, \phi^2)$ e sua função geradora de densidade tem a forma distribuição exponencial potência é dada por

$$h(u) = C(k)\exp\{-u^{1/(1+k)}/2\}, -1 < k \leq 1, u > 0,$$

onde $C(k)^{-1} = \Gamma(1 + (1+k)/2)2^{1+(1+k)/2}$, então Y possui uma distribuição exponencial potência e denotamos por $Y \sim EP(\mu, \phi^2, k)$. Temos ainda que $E[Y] = \mu$ e $Var(Y) = 2^{(1+k)}\{\Gamma(\frac{3(1+k)}{2})\Gamma(\frac{1+k}{2})\}\phi^2$.

Podemos ver o parâmetro k como uma medida de curtose, ou mesmo, como uma medida de não normalidade pois quando $k = 0$ temos a distribuição normal. Em particular, quando $k = 1$ temos a distribuição exponencial dupla. Se k tende a -1, a distribuição tende a uma uniforme no intervalo $(\mu - \sqrt{(3\phi)}, \mu + \sqrt{(3\phi)})$.

2.2.8 Distribuição de Kotz

Uma variável aleatória Y tem distribuição de Kotz se a função geradora de densidade $h(\cdot)$ tem a forma

$$h(u) = \frac{r^{(2N-1)/2}}{\Gamma(\frac{2N-1}{2})}u^{N-1}\exp\{-ru\}, r > 0, N > 1, u > 0.$$

e denotamos por $Y \sim k(\mu, \phi^2, N, r)$. Se Y tem distribuição de Kotz, então a média de Y é igual a μ e a variância é $\{(2n-1)/2\}$. O coeficiente de curtose é dado por $\gamma_2 = (2n+1)/(2n-1)$ e os momentos centrais de ordem $2m$ são da forma

$$\mu_{2m} = \frac{\Gamma(\frac{2m+2n-1}{2})}{r^m\Gamma(\frac{2n-1}{2})}\phi^{2m}.$$

Quando $N = 1$, a distribuição de Kotz se reduz à distribuição normal de parâmetros μ e $\phi^2/(2r)$. Ainda se $N > 1$, a distribuição é bimodal com modas em $y = \mu \pm \sqrt{(N-1)/r\phi}$. Temos que $E(Y) = \mu$ e $Var(Y) = \{(2N-1)/(2r)\}\phi^2$. Além disto, a variável $Z^2 = (Y - \mu)^2/\phi^2$ tem distribuição Gama($(2N-1)/2, r$). Em particular, se $N = 1$ e $r = 1/2$, então $Z^2 \sim \chi_1^2$.

2.2.9 Distribuição de Laplace

Se Y tem uma função densidade de uma distribuição de Laplace ou Exponencial dupla, então a função geradora de densidades $h(\cdot)$ é da forma

$$h(u) = \frac{1}{2} \exp\{-\sqrt{u}\}, u > 0.$$

e escrevemos $Y \sim ED(\mu, \phi^2)$. Sua função característica tem a forma

$$\Psi_y(t) = \frac{e^{it\mu}}{1 + t^2\phi^2}, \quad t \in \mathbb{R}.$$

A mediana, a moda e a média da variável y são iguais a μ , $Var(y) = 2\phi^2$ e o coeficiente de curtose é $\gamma_2 = 6$. Os quartis superior e inferior são $\mu \pm 0.534\phi$. Esta distribuição surge a partir da diferença entre duas variáveis independentes com distribuição exponencial.

Figura 2.4 Gráficos de funções densidade t -Student e normal.

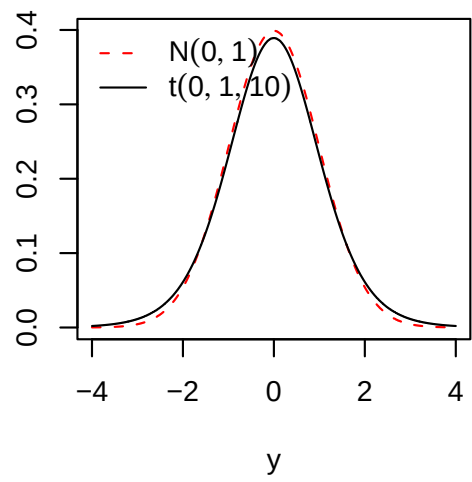
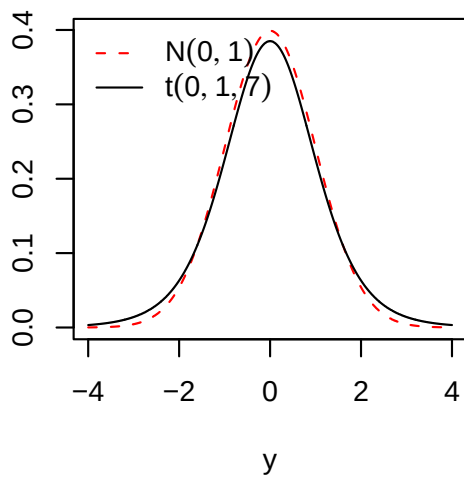
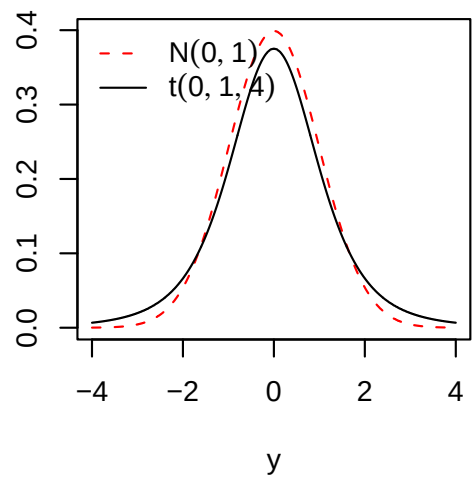
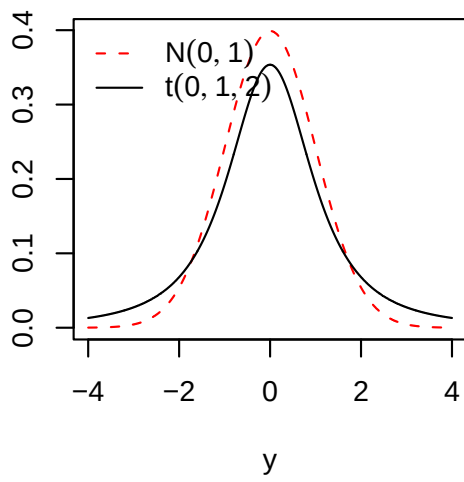
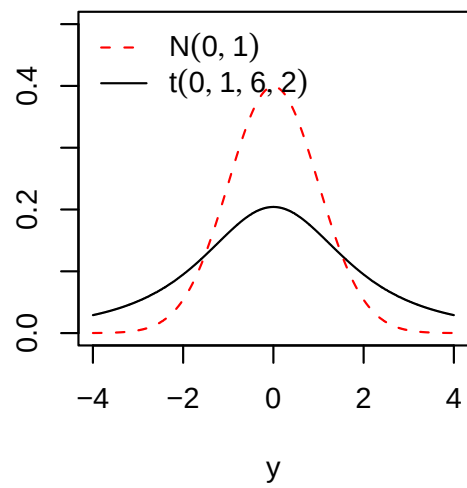
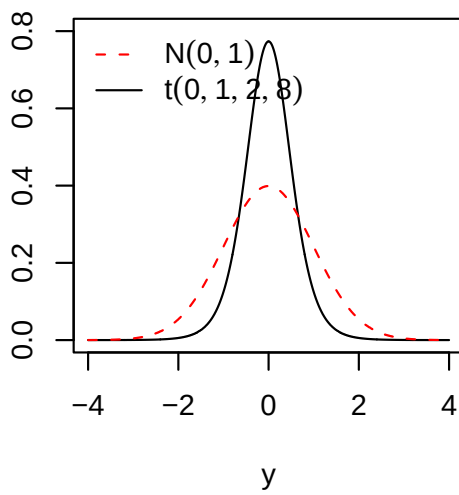
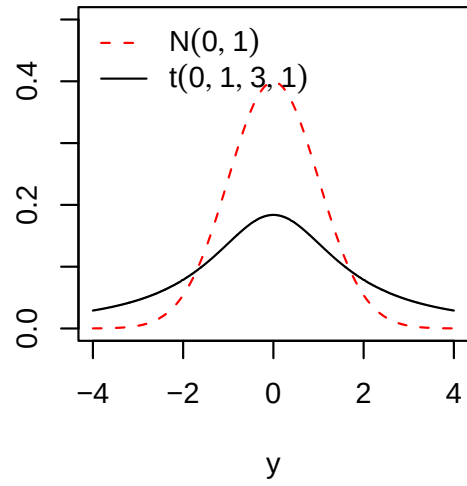
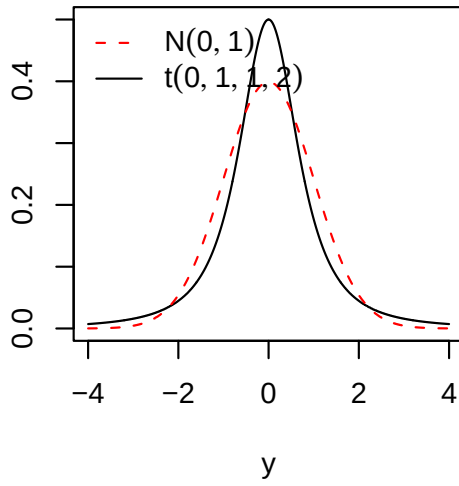


Figura 2.5 Gráficos de funções densidade t -Student generalizada e normal.



2.3 Estimação por Máxima Verossimilhança

Considere n variáveis aleatórias independentes $Y_1, Y_2, \dots, Y_n$, onde cada Y_l tem função densidade contínua na forma (2.1). Seja $y = (y_1, \dots, y_n)'$ uma amostra aleatória de tamanho n da família de distribuições simétricas, onde os y_l 's são valores observados de cada Y_l . Seja $L(\theta)$, onde $\theta = (\beta^\top, \phi)^\top$, a função de verossimilhança do modelo definido em (2.1) e (2.2), dados $y_1, \dots, y_n$. Temos

$$\begin{aligned} L(y; \theta) &= \prod_{l=1}^n \pi(y_l; \mu_l, \phi) \\ &= \frac{1}{\phi^n} \prod_{l=1}^n h\left(\left(\frac{y_l - \mu_l}{\phi}\right)^2\right). \end{aligned}$$

Seja $\ell(\theta)$ o logaritmo da função de verossimilhança, definido como

$$\begin{aligned} \ell(y; \theta) &= -n \log(\phi) + \sum_{l=1}^n \log h\left(\left(\frac{y_l - \mu_l}{\phi}\right)^2\right) \\ &= -n \log(\phi) + \sum_{l=1}^n g(z_l), \end{aligned} \tag{2.3}$$

onde

$$g(z_l) = \log h(z_l^2),$$

sendo

$$z_l = \frac{y_l - \mu_l}{\phi}, \quad \text{sendo } \mu_l = f(x_l; \beta).$$

Assumimos que $\ell(\theta)$ é regular com respeito às derivadas dos componentes de β e ϕ até a quarta ordem.

Para a obtenção de estimadores de máxima verossimilhança, bem como para a obtenção das estatísticas de teste e de suas respectivas correções de Bartlett ou tipo-Bartlett será necessário obter as derivadas do logaritmo da função de verossimilhança com relação aos parâmetros desconhecidos e também alguns momentos destas derivadas. Assumimos, portanto, no que segue, que tais derivadas e momentos existem.

Por outro lado, introduziremos a seguinte notação: $f_l^r = \partial \mu_l / \partial \beta_r$, $f_l^{rs} = \partial^2 \mu_l / \partial \beta_r \partial \beta_s$, $f_l^{r,s} = \partial \mu_l / \partial \beta_r \partial \mu_l / \partial \beta_s$, $f_l^{rs,t} = \partial^2 \mu_l / \partial \beta_r \partial \beta_s \partial \mu_l / \partial \beta_t$, etc.

A primeira derivada do logaritmo da função de verossimilhança dada em (2.3) é

$$\frac{\partial \ell(\theta)}{\partial \beta_r} = -\frac{1}{\phi} \sum_{l=1}^n g_l^{(1)} f_l^r \quad \text{com } r = 1, \dots, p, \quad (2.4)$$

$$\begin{aligned} \frac{\partial \ell(\theta)}{\partial \phi} &= -\frac{1}{\phi} \sum_{l=1}^n (1 + g_l^{(1)} z_l), \\ &= -\frac{n}{\phi} - \frac{1}{\phi} \sum_{l=1}^n g_l^{(1)} z_l \end{aligned} \quad (2.5)$$

onde

$$g^{(i)} = \frac{\partial^{(i)} g}{\partial z_l^{(i)}}, \quad l = 1, \dots, n.$$

Assumimos que esta derivada existe para todo $z_l \in \mathbb{R}$. Podemos escrever $g_l^{(1)}$ como $g_l^{(1)} = -z_l w_l$, onde

$$w_l = -\frac{2d \log h(u)}{du}.$$

A Quadro 2.1 apresenta os pesos ω_l para algumas distribuições .

Quadro 2.1 . Valor de ω para algumas distribuições simétricas.

Distribuição	ω
Normal	1
Cauchy	$2/(1 + z^2)$
t -Student	$(\nu + 1)/(\nu + z^2)$
t -Student generalizada	$(r + 1)/(s + z^2)$
Logística I	$2(1 - e^{-z^2})/(1 + e^{-z^2})$
Logística II	$(e^{ z } - 1)/(z e^{ z } + 1)$

A função escore total de θ $U = U(\theta)$ tem a forma $U = (U_\beta^\top, U_\phi)^\top$, tal que $U_\phi = \partial \ell(\theta)/\partial \phi$ e $U_\beta^\top = (\partial \ell(\theta)/\partial \beta_1, \partial \ell(\theta)/\partial \beta_2, \dots, \partial \ell(\theta)/\partial \beta_p)^\top$, com

$$U_\beta = \frac{1}{\phi^2} \tilde{X}^\top W(Y - \mu)$$

e

$$U_\phi = -\frac{n}{\phi} + \frac{1}{\phi^3} (Y - \mu)^\top W(Y - \mu),$$

sendo $W = \text{diag}(w_1, \dots, w_n)$ e $\tilde{X}$ é como definido na Subseção 2.2.

O estimador de máxima verossimilhança de θ é obtido da solução do sistema de equações dado por

$$\left. \frac{\partial \ell(\theta)}{\partial \beta_r} \right|_{\theta=\hat{\theta}} = 0 \quad \text{e} \quad \left. \frac{\partial \ell(\theta)}{\partial \phi} \right|_{\theta=\hat{\theta}} = 0, \quad r = 1, \dots, p.$$

As equações dadas acima são não-lineares e não podem ser resolvidas explicitamente. Elas devem ser resolvidas iterativamente usando, por exemplo, o algoritmo de Newton-Raphson ou escore de Fisher, que envolvem, respectivamente, as matrizes de informação observada e esperada que definimos a seguir.

As segundas derivadas do logaritmo da função da verossimilhança têm as formas

$$\begin{aligned}\frac{\partial^2 \ell(\theta)}{\partial \beta_r \partial \beta_s} &= \frac{1}{\phi^2} \sum_{l=1}^n g_l^{(2)} f_l^{r,s} - \frac{1}{\phi} \sum_{l=1}^n g_l^{(1)} f_l^{rs}, \\ \frac{\partial^2 \ell(\theta)}{\partial \beta_r \partial \phi} &= \frac{1}{\phi^2} \sum_{l=1}^n (g_l^{(1)} + g_l^{(2)} z_l) f_l^r, \\ \frac{\partial^2 \ell(\theta)}{\partial \phi^2} &= \frac{1}{\phi^2} \sum_{l=1}^n (1 + 2g_l^{(1)} z_l + g_l^{(2)} z_l^2).\end{aligned}$$

Assim, a matriz de informação total de Fisher para θ na classe dos modelos simétricos é dada por

$$K(\theta) = - \begin{bmatrix} E \left[\frac{\partial^2 \ell(\theta)}{\partial \beta_r \partial \beta_s} \right] & E \left[\frac{\partial^2 \ell(\theta)}{\partial \beta_r \partial \phi} \right] \\ E \left[\frac{\partial^2 \ell(\theta)}{\partial \phi \partial \beta_r} \right] & E \left[\frac{\partial^2 \ell(\theta)}{\partial \phi^2} \right] \end{bmatrix},$$

em que

$$\begin{aligned}E \left[\frac{\partial^2 \ell(\theta)}{\partial \beta_r \partial \beta_s} \right] &= -E \left[\frac{\partial \ell(\theta)}{\partial \beta_r} \frac{\partial \ell(\theta)}{\partial \beta_s} \right] = -\delta_{(0,1,0,0,0)} \frac{\tilde{X}^\top \tilde{X}}{\phi^2}, \\ E \left[\frac{\partial^2 \ell(\theta)}{\partial \phi^2} \right] &= -E \left[\frac{\partial \ell(\theta)}{\partial \phi} \frac{\partial \ell(\theta)}{\partial \phi} \right] = n \frac{[\delta_{(0,1,0,0,2)} - 1]}{\phi^2}, \\ E \left[\frac{\partial^2 \ell(\theta)}{\partial \beta_r \partial \phi} \right] &= 0,\end{aligned}$$

ou seja,

$$K(\theta) = -\frac{1}{\phi^2} \begin{bmatrix} \delta_{(0,1,0,0,0)} \tilde{X}^\top \tilde{X} & 0 \\ 0 & n[\delta_{(0,1,0,0,2)} - 1] \end{bmatrix},$$

onde usamos a notação

$$\delta_{(a,b,c,d,e)} = E\{g^{(1)a}g^{(2)b}g^{(3)c}g^{(4)d}z_l^e\}, \quad (2.6)$$

com $a, b, c, d, e = 1, 2, 3, 4$, $g^{(r)} = \partial g / \partial z^r$ (no Apêndice D são dados os δ 's necessários aos algoritmos para algumas distribuições).

O fato de que β e ϕ , são globalmente ortogonais (Cox e Reid, 1987), no sentido de que a matriz de informação K é bloco diagonal, implica que os estimadores de máxima verossimilhança de β e ϕ são assintoticamente não-correlacionados. Temos que

$$\hat{\beta} \sim AN_p(\beta, K_{\beta,\beta}^{-1}), \quad \hat{\phi} \sim AN(\phi, K_{\phi,\phi}^{-1})$$

e $\hat{\beta}$ e $\hat{\phi}$ são assintoticamente independentes.

É fácil ver, usando o algoritmo score de Fisher definido na Seção 1.4, que a estimação de θ por máxima verossimilhança pode ser feita através do seguinte esquema iterativo

$$\begin{aligned} \beta^{(m+1)} &= \beta^{(m)} + (\tilde{X}^{(m)\top} \tilde{X}^{(m)})^{-1} \tilde{X}^{(m)\top} W^{(m)} (y - \mu^{(m)}) \delta_{(2,0,0,0,0)}^{-1}, \\ \phi^{(m+1)} &= \phi^{(m)} + \phi^{(m)-1} (\delta_{(2,0,0,0,2)} - 1) \left\{ \frac{1}{n} (y - \mu^{(m)})^\top W^{(m)} (y - \mu^{(m)}) - \phi^{(m)} \right\}, \end{aligned}$$

em que $W = \text{diag}\{w_1, \dots, w_n\}$ com $w = -z^{-1} \frac{d \log h(z^2)}{dz}$, $m = 0, 1, 2, \dots$

Vale salientar que, quando existem outros parâmetros desconhecidos no modelo, tais como, graus de liberdade, é necessário obter a matriz de informação para todos os parâmetros desconhecidos e estimá-los. Outra alternativa, mais simples, é repetir o processo iterativo para uma gama de valores para o(s) parâmetro(s) extra(s) e escolher aquele(s) que produz(em) o maior valor para a verossimilhança.

2.4 Testes da Razão de Verossimilhanças e Escore em Modelos Não-Lineares Simétricos

2.4.1 Testes de hipóteses sobre o parâmetro β com ϕ conhecido

Considere o vetor $y = (y_1, \dots, y_n)^\top$ representando n observações independentes, onde cada y_i tem função densidade na família simétrica definida em (2.1). Assumimos que o parâmetro de escala é conhecido, que μ está relacionado com β através de (2.2) e que os outros parâmetros serão considerados fixados. Seja $\ell(\beta)$ o logaritmo da função de verossimilhança para o parâmetro β dado em (2.3). Assumimos que a função $\ell(\beta)$ é regular com respeito às derivadas em relação aos componentes de β até a quarta ordem. Pelas condições de regularidade (ver em Cox e Hinkley, 1974, Cap.9), temos $\delta_{(1,0,0,0,0)} = 0$ e $\delta_{(0,1,0,0,0)} = -\delta_{(2,0,0,0,0)}$. A função $\delta_{(a,b,c,d,e)}$ desempenha papel importante no cálculo das correções de Bartlett para a estatística da razão de verossimilhanças e tipo-Bartlett para a estatística escore na classe de distribuições simétricas.

O interesse é testar

$$H_0 : \beta_1 = \beta_1^{(0)} \quad \text{versus} \quad H_1 : \beta_1 \neq \beta_1^{(0)},$$

onde o vetor de parâmetros desconhecidos β com p componentes é particionado como $\beta = (\beta_1^\top, \beta_2^\top)^\top$, sendo $\beta_1 = (\beta_1, \beta_2, \dots, \beta_q)^\top$ o vetor de parâmetros de interesse e $\beta_2 = (\beta_{q+1}, \dots, \beta_p)^\top$ o vetor de dimensão $(p - q) \times 1$ de parâmetros de perturbação e $\beta_1^{(0)}$ é um vetor especificado de dimensão $q \times 1$. Consideremos, também, a correspondente partição da matriz modelo $\tilde{X} = (\tilde{X}_1, \tilde{X}_2)$.

A função escore total de β pode ser particionada da mesma forma que β como $U(\beta) = (U_1^\top(\beta), U_2^\top(\beta))^\top$, onde

$$U_1(\beta) = \frac{1}{\phi^2} \tilde{X}_1^\top W(Y - \mu) \quad \text{e} \quad U_2(\beta) = \frac{1}{\phi^2} \tilde{X}_2^\top W(Y - \mu);$$

a matriz W foi definida na Subseção (2.3). A matriz de informação de Fisher correspondente ao parâmetro β supondo ϕ conhecido é dada por

$$K(\beta) = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix},$$

que supomos ser positiva definida, com

$$K_{11}(\beta) = \frac{\delta_{(2,0,0,0,0)}}{\phi^2} \tilde{X}_1^\top \tilde{X}_1, \quad K_{22}(\beta) = \frac{\delta_{(2,0,0,0,0)}}{\phi^2} \tilde{X}_2^\top \tilde{X}_2,$$

$$K_{12}(\beta) = K_{21}(\beta) = \frac{\delta_{(2,0,0,0,0)}}{\phi^2} \tilde{X}_1^\top \tilde{X}_2.$$

Utilizando a fórmula de inversa de matrizes simétricas particionadas (Rao, 1973, p. 33) e após alguma álgebra, chega-se a

$$K^{-1} = \begin{bmatrix} K^{11} & K^{12} \\ K^{21} & K^{22} \end{bmatrix},$$

onde

$$\begin{aligned} K^{11} &= \frac{\phi^2}{\delta_{(2,0,0,0,0)}} (R^\top R)^{-1}, \\ K^{12} = K^{21} &= -\frac{\phi^2}{\delta_{(2,0,0,0,0)}} (R^\top R)^{-1} C^\top, \\ K^{22} &= \frac{\phi^2}{\delta_{(2,0,0,0,0)}} ((\tilde{X}_2^\top \tilde{X}_2)^{-1} + C(R^\top R)^{-1} C^\top), \end{aligned} \quad (2.7)$$

em que $R = \tilde{X}_1 - \tilde{X}_2 C$, com $C = (\tilde{X}_2^\top \tilde{X}_2)^{-1} \tilde{X}_2^\top \tilde{X}_1$ representando uma matriz $(p - q) \times q$ cujas colunas são vetores dos coeficientes de uma regressão linear das colunas de $\tilde{X}_1$ sobre a matriz modelo $\tilde{X}_2$.

Para testar a hipótese nula H_0 versus a hipótese alternativa H_1 , podemos utilizar a estatística da razão de verossimilhanças (RV) e a estatística score (SR), que, em problemas regulares e sob H_0 , têm distribuição assintótica χ_q^2 , onde q é a diferença entre as dimensões dos espaços paramétricos sob a hipótese alternativa e nula.

A estatística da razão de verossimilhanças para o teste de H_0 pode ser escrita da seguinte forma

$$RV = 2\{\ell(\hat{\beta}_1, \hat{\beta}_2) - \ell(\beta_1^{(0)}, \tilde{\beta}_2)\}, \quad (2.8)$$

onde $\ell(\cdot)$ é o logaritmo da função de verossimilhança definido em (2.3) e é dado por $\ell = \ell(\beta) = \sum_{l=1}^n \log \pi(y_l, \beta)$, $\hat{\beta}$ sendo o estimador de máxima verossimilhança irrestrito de β e $\tilde{\beta}_2$ denotando o estimador de máxima verossimilhança de β_2 sob H_0 .

Assintoticamente e sob a hipótese nula, temos que

$$RV \xrightarrow{D} \chi_q^2.$$

Rejeitamos a hipótese nula H_0 se $RV > q_{1-\alpha}$, onde $q_{1-\alpha}$ representa o quantil $1 - \alpha$ da distribuição qui-quadrado com q graus de liberdade e α é o nível de significância adotado, $0 < \alpha < 1$.

A estatística escore, originalmente sugerida por Rao (1947), é dada por

$$SR = U_1^\top(\tilde{\beta})K^{11}(\tilde{\beta})U_1(\tilde{\beta}), \quad (2.9)$$

onde $\tilde{\beta} = (\beta_1^{(0)\top}, \tilde{\beta}_2^\top)^\top$ e $\tilde{\beta}_2$ é o estimador de máxima verossimilhança de β_2 sob H_0 . Assintoticamente e sob a hipótese nula, temos que

$$SR \xrightarrow{D} \chi_q^2.$$

Assim, substituindo $U_1(\beta)$ e $K^{11}(\beta)$ pelas suas expressões dadas anteriormente, temos que a estatística escore para testar H_0 versus H_1 definida em (2.9) fica dada por

$$SR = \tilde{s}^\top \tilde{X}_1(R'R)^{-1} \tilde{X}_1^\top \tilde{s},$$

onde $s = (s_1, \dots, s_n)^\top$, com $s_l = \delta_{(2,0,0,0,0)}^{-1/2} z_l w_l$. Em grandes amostras, satisfeitas certas condições de regularidade, a estatística SR tem distribuição assintótica χ_q^2 sob H_0 .

Rejeitamos a hipótese nula H_0 se $SR > q_{1-\alpha}$, onde $q_{1-\alpha}$ representa o quantil $1 - \alpha$ da distribuição qui-quadrado com q graus de liberdade e α é o nível de significância adotado, $0 < \alpha < 1$.

O teste da razão de verossimilhanças envolve a estimação dos parâmetros sob as hipóteses nula e alternativa, enquanto o teste escore, requer estimação somente sob a hipótese nula. No caso, em que a hipótese nula for $H_0 : \beta = \beta^{(0)}$, o teste escore não requer estimação alguma. Nos casos em que a estimação sob a hipótese alternativa é complicada, o teste escore pode apresentar vantagem computacional em relação ao teste da razão de verossimilhanças.

2.4.2 Testes de hipóteses sobre o parâmetro β com ϕ desconhecido

Na seção anterior, apresentamos os testes de hipóteses para os parâmetros de regressão de modelos não-lineares simétricos supondo que o parâmetro escala ϕ é conhecido e igual para todas as observações. Nesta seção eliminamos a suposição de que ϕ é conhecido.

Desejamos testar a hipótese nula $H_0 : \beta_1 = \beta_1^{(0)}$ versus a hipótese alternativa $H_1 : \beta_1 \neq \beta_1^{(0)}$, com β_2 e ϕ representando os parâmetros de perturbação e $\beta_1^{(0)}$ sendo um vetor fixado de dimensão q ($q \leq p$). A estatística da razão de verossimilhanças para testar de H_0 pode ser escrita como

$$RV = 2\{l(\hat{\beta}_1, \hat{\beta}_2, \hat{\phi}) - l(\beta_1^{(0)}, \tilde{\beta}_2, \tilde{\phi})\}, \quad (2.10)$$

onde $\hat{\beta}_1, \hat{\beta}_2$ e $\hat{\phi}$ são os estimadores de máxima verossimilhança irrestritos de β_1, β_2 e ϕ , respectivamente, enquanto que $\tilde{\beta}_2$ e $\tilde{\phi}$ são os estimadores de máxima verossimilhança de β_2 e ϕ sob H_0 .

Por outro lado, a estatística escore, definida em (2.9), para testar H_0 é dada por

$$SR = \tilde{s}^\top \tilde{X}_1 (R^\top R)^{-1} \tilde{X}_1^\top \tilde{s},$$

onde $\tilde{s} = \delta_{(2,0,0,0)}^{-1/2} z_l w_l$, avaliado em $\theta^{(0)} = (\beta_1^{(0)\top}, \tilde{\beta}_2^\top, \tilde{\phi})^\top$.

Em problemas regulares, estas duas estatísticas têm distribuição assintótica χ^2 com q graus de liberdade, sob a hipótese nula.

Buse (1982) apresenta de forma bastante didática a interpretação geométrica dos testes da razão de verossimilhanças e escore para o caso de hipóteses simples.

2.5 Testes da Razão de Verossimilhanças e Escore Bootstrap

A técnica de bootstrap foi originalmente proposta por Efron (1979, 1981) num contexto de estimação em populações infinitas. Cysneiros e Cordeiro (2002) usaram esta técnica para obter uma aproximação bootstrap para a distribuição nula da estatística escore nos modelos não-lineares da família locação e escala. Lemonte (2006) usou esta técnica para obter a estatística da razão de verossimilhanças bootstrap no modelo de Birnbaum-Saunders. O objetivo encontrar a distribuição da estatística diretamente da amostra aleatória observada.

O esquema bootstrap pode ser realizado através dos seguintes passos:

- Retirar um número B de amostras aleatórias com reposição, supondo H_0 verdadeira, da amostra original $y = (y_1, \dots, y_n)^\top$.
- Calcular o valor da estatística desejada para as B novas amostras, tendo como vetor de valores da variável y , RV^{*b} ou SR^{*b} .

O percentil $1 - \alpha$ de RV^{*b} ou SR^{*b} é estimado pelo valor z_α , tal que

$$\# \frac{\{RV^{*b} \leq z_\alpha\}}{B} = 1 - \alpha$$

ou

$$\# \frac{\{SR^{*b} \leq z_\alpha\}}{B} = 1 - \alpha$$

c. Tomar o percentil, z_α , como segue: ordena-se as B réplicas bootstrap da estatística, a réplica $B \times (1 - \alpha)$ é o percentil estimado, z_α , assumindo que $B \times (1 - \alpha)$ é inteiro. Se $B \times (1 - \alpha)$ não for inteiro, podemos utilizar o seguinte procedimento: seja $k = \lfloor (B + 1) \times \alpha \rfloor$ o maior inteiro $\leq (B + 1) \times \alpha$; então, a quantidade de z_α é dada pelo $(B + 1 - k)$ -ésimo elemento ordenado de RV^{*b} . Uma boa discursão sobre o uso de bootstrap em testes de hipóteses pode ser encontrada em Efron e Tibshirani (1993, Cap.16).

O teste desejado consiste em rejeitar a hipótese nula H_0 quando a estatística é maior que o percentil estimado. O teste da razão de verossimilhanças bootstrap consiste em rejeitar a hipótese nula H_0 se

$$RV > z_\alpha;$$

no caso do teste escore, rejeitamos H_0 se

$$SR > z_\alpha.$$

O teste escore bootstrap foi considerado por Godfrey e Orme (1999), que usaram $B = 400$, Cribari-Neto e Zarkos (1999) que tomaram $B = 500$, e Cysneiros e Cordeiro (2002), tomando $B = 300$. Já Lemonte (2006) tomou $B = 600$ para fazer o teste da razão de verossimilhança por bootstrap.

Capítulo 3

Ajuste para o Teste da Razão de Verossimilhança e para o Teste Escore

3.1 Introdução

Em problemas de testes de hipóteses envolvendo grandes amostras, é muito comum o uso dos testes da razão de verossimilhanças (RV) e escore (SR). Em geral, há grande dificuldade em se determinar as distribuições nulas exatas das estatísticas destes testes, razão pela qual os testes têm sido construídos com base em resultados assintóticos. Em problemas regulares, as estatísticas da razão de verossimilhanças (RV) e escore (SR) têm, sob a hipótese nula, distribuição χ^2_q aproximadamente, em amostras grandes, ou seja, χ^2 com q graus de liberdade, onde q é a diferença entre as dimensões dos espaços paramétricos sob as hipóteses alternativa e nula. Os testes são comumente baseados na comparação das estatísticas com valores críticos obtidos da distribuição χ^2 de referência para níveis de significância nominal fixado. Em pequenas amostras ou mesmo em amostras de tamanho moderado, a aproximação pode não ser satisfatória, podendo conduzir a taxas de rejeição sob a hipótese nula bastante distorcidas. Para melhorar a qualidade da aproximação das distribuições das estatísticas RV e SR pela distribuição qui-quadrado são utilizadas as chamadas correções de Bartlett e tipo-Bartlett.

A primeira proposta de melhoria de testes estatísticos foi feita por Bartlett (1937). O autor modificou a estatística da razão de verossimilhanças por um fator de correção, visando a produzir uma estatística modificada com o primeiro momento igual ao da distribuição χ^2 de referência. Sob H_0 , o valor esperado desta estatística pode ser expandido, em problemas regulares, como $E(RV) = q + d_\beta + O(n^{-2})$, onde d_β é uma constante de ordem n^{-1} que pode ser consistentemente estimada sob H_0 e n é o tamanho da amostra. Assim, o valor esperado da estatística modificada $RV^* = RV/(1 + d_\beta)$ ou, equivalentemente, $RV^{**} = RV(1 - d_\beta)$, é dada por $q + O(n^{-2})$, estando ‘mais próximo’ daquele da distribuição χ^2_q do que o valor esperado de RV . Os fatores $1/(1 + d_\beta)$ e $(1 - d_\beta)$ ficaram conhecidos como correções de Bartlett. Vale ressaltar que tais fatores não dependem do valor da estatística da razão de verossimilhanças, mas podem

dependem de parâmetros desconhecidos; neste caso, estes devem ser substituídos por suas respectivas estimativas de máxima verossimilhança, sob H_0 , o que não afeta a ordem da aproximação resultante.

Hayakawa (1977) desenvolveu, sob a hipótese nula, uma expansão assintótica para a distribuição de RV e mostrou que, se a hipótese nula for simples, a estatística de RV^* tem distribuição χ^2 até ordem n^{-1} . Os resultados de Hayakawa (1977) apresentam um erro que foi, posteriormente, corrigido por Chesher e Smith (1995) (ver também Harris, 1986; Cordeiro, 1987; e Hayakawa, 1987).

Diversos trabalhos foram desenvolvidos nos últimos anos apresentando correções de Bartlett para a estatística da razão de verossimilhanças em problemas específicos e em modelos amplos. Em particular, correções de Bartlett para os modelos lineares generalizados quando o termo de escala é conhecido e desconhecido, respectivamente, foram obtidas por Cordeiro (1983, 1987), que mostrou que o uso do fator de correção representa um grande aperfeiçoamento nos testes de adequação dos modelos lineares generalizados. Para os modelos não-lineares da família exponencial com parâmetro de dispersão conhecido correções de Bartlett foram obtidas por Cordeiro e Paula (1989). DiCiccio (1986) estudou um aperfeiçoamento da estatística da razão de verossimilhanças para os modelos de locação baseado em inferência condicional. Correções de Bartlett na família exponencial uniparamétrica foram obtidas por Cordeiro, Cribari-Neto, Aubin e Ferrari (1995) e em modelos de regressão com erros t -Student foram obtidos por Ferrari e Arellano-Valle (1996). Correções similares para modelos lineares normais heteroscedásticos e em alguns modelos de regressão multivariada foram obtidas por Cribari-Neto e Ferrari (1995a) e Cribari-Neto e Zarkos (1995). Ferrari e Uribe-Opazo (2001) obtiveram a correção de Bartlett para os modelos lineares simétricos. Correções de Bartlett para os modelos não-lineares de locação e escala supondo que o parâmetro de escala é conhecido foram obtidos por Montenegro e Cordeiro (2002). Os resultados de Ferrari e Uribe-Opazo (2001) foram generalizados por Cordeiro (2004) que obteve uma correção de Bartlett para a estatística da razão de verossimilhanças nos modelos não-lineares simétricos. Ferrari, Cysneiros e Cribari-Neto (2004) obtiveram correção de Bartlett para a estatística da razão de verossimilhança perfilada modificada no modelo de regressão normal linear heteroscedástico, generalizando assim o artigo de Ferrari e Cribari-Neto (2002). A correção de Bartlett para a estatística da razão de verossimilhanças perfilada modificada em modelos de regressão não-lineares

da família exponencial foi obtida por Cysneiros e Ferrari (2006).

Posteriormente à obtenção do fator de correção de Bartlett para a estatística RV , foram obtidas correções similares para outras estatísticas, a saber, a estatística SR que converge para uma distribuição qui-quadrado. Tais correções são chamadas de correções tipo-Bartlett.

Harris (1985) obteve uma expressão assintótica até ordem n^{-1} para a distribuição da estatística escore SR , aperfeiçoando o teste baseado na própria estatística escore e em uma expansão, até ordem n^{-1} , para os seus quantis.

Cordeiro e Ferrari (1991) definiram uma estatística escore modificada (SR^*) que tem distribuição qui-quadrado até ordem n^{-1} . A estatística modificada é dada por SR multiplicada por um fator de correção, denominado fator de correção tipo-Bartlett, que envolve um polinômio de segundo grau na própria estatística e cujos coeficientes são, em geral, funções de parâmetros desconhecidos, que devem ser substituídos por suas respectivas estimativas de máxima verossimilhanças sob H_0 . Cordeiro, Ferrari e Paula (1993) obtiveram correções tipo-Bartlett para estatísticas escore em modelos lineares generalizados para o caso em que o parâmetro de dispersão é conhecido. Estes resultados foram generalizados por Cribari-Neto e Ferrari (1995b), que obtiveram para o ajuste no caso de dispersão desconhecida. Correções para testes escore em modelos lineares heteroscedásticos foram obtidos por Cribari-Neto e Ferrari (1995c). Correções tipo-Bartlett para a estatística escore em modelos de regressão com erros t -Student foram desenvolvidas por Ferrari e Arellano-Valle (1996). Correções tipo-Bartlett para os modelos lineares simétricos foram desenvolvidas por Uribe-Opazo (1997) para o caso em que o parâmetro de dispersão é desconhecido. Correções tipo-Bartlett para os modelos não-lineares de locação e escala foram obtidas por Cysneiros e Cordeiro (2002) considerando que o parâmetro de escala é conhecido.

Boas revisões sobre correções de Bartlett e tipo-Bartlett para as estatísticas da razão de verossimilhanças e escore, respectivamente, podem ser encontradas em Cribari-Neto e Cordeiro (1997) e em Cordeiro (1999).

Neste capítulo, apresentamos as correções de Bartlett e tipo-Bartlett para as estatísticas da razão de verossimilhanças e escore, respectivamente, no modelo de regressão não-linear simétrico.

3.2 Correção de Bartlett para a estatística da razão de verossimilhanças

No Capítulo 2, definimos a classe de modelos simétricos assumindo que o parâmetro de locação está relacionado com um vetor de parâmetros desconhecidos através da relação dada por (2.2) e considerando os casos em que o parâmetro de escala é conhecido e desconhecido. Também apresentamos os testes da razão de verossimilhanças e escore da hipótese nula $H_0 : \beta_1 = \beta_1^{(0)}$ versus a hipótese alternativa $H_1 : \beta_1 \neq \beta_1^{(0)}$. Assumimos, inicialmente, que ϕ é conhecido.

Lawley (1956) obteve, sob condições de regularidade, uma expansão de $L(\hat{\beta}_1, \hat{\beta}_2)$ em série de Taylor sob a hipótese nula até termos de ordem n^{-1} envolvendo derivadas até a quarta ordem do logaritmo da função de verossimilhança. Ele mostrou, após longos cálculos, que

$$2E[\ell(\hat{\beta}_1, \hat{\beta}_2) - \ell(\beta_1, \beta_2)] = p + \epsilon_{p,\beta} + O(n^{-2}),$$

onde o termo $\epsilon_{p,\beta}$ é de ordem n^{-1} e é escrito na forma

$$\epsilon_{p,\beta} = \sum' (\lambda_{rstu} - \lambda_{rstuvw}), \quad (3.1)$$

com os índices r, s, t, u, v, w referindo-se aos componentes do vetor β , e onde

$$\begin{aligned} \lambda_{rstu} &= \kappa^{rs} \kappa^{tu} \left(\frac{\kappa_{rstu}}{4} - \kappa_{rst}^{(u)} + \kappa_{rt}^{(su)} \right), \\ \lambda_{rstuvw} &= \kappa^{rs} \kappa^{tu} \kappa^{vw} \left\{ \kappa_{rtv} \left(\frac{\kappa_{suw}}{6} - \kappa_{sw}^{(u)} \right) \right. \\ &\quad \left. + \kappa_{rtu} \left(\frac{\kappa_{svw}}{4} - \kappa_{sw}^{(v)} \right) + \kappa_{rt}^{(v)} \kappa_{sw}^{(u)} + \kappa_{rt}^{(u)} \kappa_{sw}^{(v)} \right\}, \end{aligned} \quad (3.2)$$

em que $\sum'$ denota o somatório sobre todos os componentes do vetor β , isto é, sobre os p parâmetros.

Observe que este termo depende de cumulantes de derivadas do logaritmo da função de verossimilhança e suas derivadas, sendo avaliado no parâmetro verdadeiro.

Lawley demonstrou, também, que

$$2E[\ell(\beta_1^{(0)}, \tilde{\beta}_2) - \ell(\beta_1, \beta_2)] = p - q + \epsilon_{p-q,\beta} + O(n^{-2}),$$

onde

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

em que $\sum''$ denota o somatório apenas sobre os componentes do vetor β_2 , isto é, sobre $p - q$ parâmetros de perturbação, uma vez que β_1 está fixado em $\beta_1^{(0)}$.

A estatística da razão de verossimilhanças definida em (2.8) pode ser reescrita como

$$RV = 2\{\ell(\hat{\beta}_1, \hat{\beta}_2) - \ell(\beta_1, \beta_2) - \ell(\beta_1^{(0)}, \tilde{\beta}_2) + \ell(\beta_1, \beta_2)\},$$

$\ell(\beta_1^{(0)}, \tilde{\beta}_2)$ é o logaritmo da função de verossimilhança avaliada no parâmetro verdadeiro e o $\tilde{\beta}_2$ é o estimador de máxima verossimilhança de β_2 sob H_0 . Desta forma, sob H_0 o valor esperado da estatística da razão de verossimilhanças é dado por

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

Podemos melhorar a aproximação da média da estatística da razão de verossimilhanças pela média da distribuição χ_q^2 substituindo RV pela estatística modificada RV^* dada por

$$RV^* = \frac{RV}{1 + d_\beta},$$

onde o fator de correção de Bartlett é determinado através de

$$c = 1 + d_\beta,$$

em que

$$d_\beta = \frac{\epsilon_p - \epsilon_{p-q}}{q}. \quad (3.3)$$

Hayakawa (1977) (ver correção de Chesher e Smith, 1995) mostra que, sob certas condições de regularidade e sob a hipótese nula, a estatística modificada RV^* tem distribuição χ_q^2 até ordem n^{-1} .

Um teste da razão de verossimilhanças aperfeiçoado compara, então, a estatística modificada RV^* com a distribuição χ_q^2 de referência. No caso do teste de hipótese nula simples $H_0 : \beta = \beta^{(0)}$, a quantidade d_β dada em (3.3) que determina o fator de correção de Bartlett se reduz a $d_\beta = \frac{\epsilon_{p,\beta}}{p}$.

3.2.1 Correção de Bartlett em modelos não-lineares simétricos com ϕ conhecido

Os fatores de correção de Bartlett dependem da quantidade $\epsilon_{p,\beta}$, que é uma função aparentemente complicada dos cumulantes conjuntos de derivadas do logaritmo da função de verossimilhança (que são valores esperados de produtos de derivadas do logaritmo da função de verossimilhança) e suas derivadas. Nesta Subseção, apresentaremos $\epsilon_{p,\beta}$ em forma simples, para os testes da razão de verossimilhanças em modelos de regressão não-lineares simétricos. Uma propriedade fundamental dos modelos simétricos é que os cumulantes são invariantes sob a permutação de parâmetros de regressão. Isto facilita os cálculos na obtenção de d_β e, conseqüentemente, na estatística da razão de verossimilhanças corrigida.

A obtenção do aperfeiçoamento do teste da razão de verossimilhanças para o teste da hipótese $H_0 : \beta_1 = \beta_1^{(0)}$ versus $H_1 : \beta_1 \neq \beta_1^{(0)}$, considerando ϕ conhecido, é baseada no desenvolvimento das fórmulas definidas em (3.1) e (3.2) para a obtenção da constante d_β , definida em (3.3).

Usamos a notação $f_l^r = \partial\mu_l/\partial\beta_r$, $f_l^{rs} = \partial^2\mu_l/\partial\beta_r\partial\beta_s$, $f_l^{r,s} = \partial\mu_l/\partial\beta_r \times \partial\mu_l/\partial\beta_s$ e assumimos que o logaritmo da função de verossimilhança é regular com respeito às derivadas até a quarta ordem sob H_0 .

Temos os seguintes cumulantes para modelos não-lineares simétricos (vide Apêndice A):

$$\begin{aligned}\kappa_{rs} &= \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n f_l^{r,s}, \\ \kappa_{r,s} &= \frac{\delta_{(2,0,0,0,0)}}{\phi^2} \sum_{l=1}^n f_l^{r,s}, \\ \kappa_{rs}^{(t)} &= \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n (f_l^{rt,s} + f_l^{r,ts}), \\ \kappa_{rs}^{(tu)} &= \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n (f_l^{rtu,s} + f_l^{rs,tu} + f_l^{tsu,r} + f_l^{ts,ru}), \\ \kappa_{rst} &= \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n (f_l^{rt,s} + f_l^{r,st} + f_l^{rs,t}),\end{aligned}$$

$$\begin{aligned}\kappa_{rst}^{(u)} &= \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n (f_l^{rtu,s} + f_l^{ru,st} + f_l^{rs,tu} + f_l^{rt,su} + f_l^{r,stu} + f_l^{rsu,t}). \\ \kappa_{rstu} &= \frac{\delta_{(0,0,0,1,0)}}{\phi^4} \sum_{l=1}^n f_l^{r,s,t,u} + \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n (f_l^{rtu,s} + f_l^{rt,su} + f_l^{ru,st} \\ &\quad + f_l^{r,stu} + f_l^{rst,u} + f_l^{rsu,t} + f_l^{rs,tu}),\end{aligned}$$

onde $\delta_{(a,b,c,d,e)}$, é definido em (2.6). Observe que todos os cumulantes dependem somente de ϕ e dos parâmetros considerados conhecidos de cada distribuição e não dependem de β .

Definimos as matrizes de dimensão $n \times n$ positivas semi-definidas

$$Z = \tilde{X}(\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top, \quad Z_2 = \tilde{X}_2(\tilde{X}_2^\top \tilde{X}_2)^{-1} \tilde{X}_2^\top \quad (3.4)$$

de postos p e $p - q$, respectivamente. Se $p = q$, define-se $Z_2 = 0$, que denota uma matriz nula $n \times n$. É importante observar que as matrizes $Z\phi^2\delta_{(2,0,0,0,0)}^{-1}$ e $Z_2\phi^2\delta_{(2,0,0,0,0)}^{-1}$ são estruturas de covariância assintóticas de $\hat{\mu}$ e $\tilde{\mu}$. Definimos, ainda, $Z_d = \text{diag}\{z_{11}, \dots, z_{nn}\}$ e $Z_{2d} = \text{diag}\{z_{211}, \dots, z_{2nn}\}$, representando as matrizes com elementos diagonais de Z e Z_2 , respectivamente, $\tilde{X}_m$ uma matriz $p \times p$ cujo (r, s) -ésimo elemento é $f_m^{(r,s)}$, para $m = 1, \dots, n$, as matrizes B e C de dimensão $n \times n$ com os (l, m) -ésimos elementos dados por $b_{lm} = \text{tr}((\tilde{X}^\top \tilde{X})^{-1} \tilde{X}_l (\tilde{X}^\top \tilde{X})^{-1} \tilde{X}_m)$ e $c_{lm} = \tilde{x}_l^\top (\tilde{X}^\top \tilde{X})^{-1} \tilde{X}_m (\tilde{X}^\top \tilde{X})^{-1} \tilde{x}_l$, respectivamente, $\tilde{x}_l^\top$ sendo a l -ésima linha da matriz $\tilde{X}$, e D uma matriz diagonal de dimensão $n \times n$, onde o (l, l) -ésimo elemento definido como $d_l = \text{tr}(\tilde{X}_l (\tilde{X}^\top \tilde{X})^{-1})$, em que tr denota traço. Todas as matrizes acima escritas podem ser obtidas da componente sistemática μ_l , definida em (2.2), e requerem somente as duas primeiras derivadas do modelo com respeito às componentes de β .

A obtenção da constante d_β definida em (3.3) é feita substituindo os valores de κ 's na expressão dada em (3.2) e efetuando as somas sobre a amostra depois de avaliar as somas sobre os parâmetros. Desta maneira, aparecerão termos da forma $\sum' \kappa^{rs} f_l^r f_m^s$, $\sum' \kappa^{rs} f_l^{rv} \kappa^{vw} f_m^{ws}$, $\sum' f_l^r \kappa^{rs} f_m^{sw} \kappa^{vw} f_l^v$ e $\sum' \kappa^{rs} f_l^{rs}$, onde $-\kappa^{ij}$ é o (i, j) -ésimo elemento da matriz $\delta_{(2,0,0,0,0)}^{-1} \phi^2 K^{-1}$, $i, j = 1, \dots, p$. É fácil observar que estes somatórios representam $\delta_{(2,0,0,0,0)}^{-1} \phi^2 Z$, $\delta_{(2,0,0,0,0)}^{-1} \phi^2 B$, $\delta_{(2,0,0,0,0)}^{-1} \phi^2 C$ e $\delta_{(2,0,0,0,0)}^{-1} \phi^2 K D$, respectivamente. Para obter $\epsilon_{p,\beta}$, substituímos κ_{rstu} , κ_{rst}^u e κ_{rs}^u na expressão de λ_{rstu} dada em (3.2) e obtemos a parcela

$$\begin{aligned}\sum' \lambda_{rstu} = & \frac{1}{\phi^2} \sum' \frac{\delta_{(0,0,0,1,0)}}{4\phi^2} \sum_{l=1}^n f_l^{r,s,tu} + \frac{\delta_{(0,1,0,0,0)}}{4} \sum_{l=1}^n f_l^{rs,tu} \\ & - \frac{\delta_{(0,1,0,0,0)}}{2} \sum_{l=1}^n f_l^{us,rt}.\end{aligned}$$

Invertendo a ordem das somas e rearrumando os termos, temos

$$\begin{aligned}\sum' \lambda_{rstu} = & \frac{\delta_{(0,0,0,1,0)}}{4\delta_{(2,0,0,0,0)}^2} \sum_{l=1}^n z_{ll}^2 + \frac{\phi^2 \delta_{(0,1,0,0,0)}}{4\delta_{(2,0,0,0,0)}^2} \sum_{l=1}^n d_{ll}^2 \\ & - \frac{\phi^2 \delta_{(0,1,0,0,0)}}{2\delta_{(2,0,0,0,0)}^2} \sum_{l=1}^n b_{ll}.\end{aligned}$$

Finalmente, chegamos à expressão

$$\sum' \lambda_{rstu} = \frac{\delta_{(0,0,0,1,0)}}{4\delta_{(2,0,0,0,0)}^2} \text{tr}(Z_d Z_d) - \frac{\phi^2 \delta_{(0,1,0,0,0)}}{4\delta_{(2,0,0,0,0)}^2} \text{tr}(2B - D^2).$$

Da mesma forma, obtemos a segunda parcela $\sum' \lambda_{rstuvw}$. (ver Apêndice B)

$$\sum' \lambda_{rstuvw} = \frac{\delta_{(0,1,0,0,0)}^2}{2\delta_{(2,0,0,0,0)}^3} \left\{ \frac{1}{2} \iota^\top D Z D \iota - \text{tr}(ZB) \right\}.$$

De forma análoga, temos

$$\begin{aligned}\sum'' \lambda_{rstu} = & \frac{\delta_{(0,0,0,1,0)}}{4\delta_{(2,0,0,0,0)}^2} \text{tr}(Z_{2d} Z_{2d}) - \frac{\phi^2 \delta_{(0,1,0,0,0)}}{4\delta_{(2,0,0,0,0)}^2} \text{tr}(2B_2 - D_2^2) \\ \sum'' \lambda_{rstuvw} = & \frac{\delta_{(0,1,0,0,0)}^2}{2\delta_{(2,0,0,0,0)}^3} \left\{ \frac{1}{2} \iota^\top D_2 Z_2 D_2 \iota - \text{tr}(Z_2 B_2) \right\},\end{aligned}$$

onde ι corresponde um vetor $n \times 1$ de uns. A expressão resultante de $\epsilon_{p,\beta}$ é

$$\epsilon_{p,\beta} = L_p + N L_p,$$

onde L_p e $N L_p$ são dadas por

$$\begin{aligned}L_p = & \left[\frac{\delta_{(0,0,0,1,0)}}{4\delta_{(2,0,0,0,0)}^2} \text{tr}(Z_d Z_d) \right] = \frac{d_0}{n} \rho_{ZZ}, \\ N L_p = & \phi^2 \left[d_3 \text{tr}(2B_d - D^2) + d_4 \left\{ \frac{1}{2} \iota^\top D Z D \iota - \text{tr}(ZB) \right\} \right],\end{aligned}$$

onde ι é um vetor $n \times 1$ de uns, $Z^{(3)} = \{z_{lm}^3\}$, $\rho_{ZZ} = ntr(Z_d Z_d)$ e

$$d_0 = \frac{\delta_{(0,0,0,1,0)}}{4\delta_{(2,0,0,0,0)}^2}, \quad d_3 = \frac{\delta_{(0,1,0,0,0)}}{4\delta_{(2,0,0,0,0)}^2}, \quad d_4 = \frac{\delta_{(0,1,0,0,0)}^2}{2\delta_{(2,0,0,0,0)}^3}.$$

A parcela L_p reflete a linearidade e a parcela NL_p reflete a não-linearidade do modelo. Se a função $f(\cdot; \cdot)$ for linear, as matrizes B e D desaparecem e, por consequência $NL_p = 0$.

O fator de correção de Bartlett, d_β , para o teste da razão de verossimilhanças é obtido de (3.3) com as quantidades de ϵ_p e ϵ_{p-q} . Para o cálculo do termo ϵ_{p-q} a matriz $\tilde{X}$ deve ser substituída por $\tilde{X}_2$, ou seja, pela parte que corresponde aos parâmetros não especificados em H_0 .

Observa-se que a expressão de d_β é muito simples, não dependendo de estimação de parâmetros desconhecidos, considerando somente os parâmetros conhecidos de cada distribuição, e é decomposta em dois fatores, um que só depende da distribuição assumida para os dados e outra que só depende da matriz $\tilde{X}$.

Abaixo apresentamos os valores de d_0, d_3, d_4 para algumas distribuições pertencentes aos modelos não-lineares simétricos (no Apêndice D estão os δ 's para algumas distribuições)

- (i) Normal: $d_0 = 0, d_3 = -\frac{1}{4}, d_4 = \frac{1}{2}$.
- (ii) Cauchy: $d_0 = \frac{3}{4}, d_3 = -\frac{1}{2}, d_4 = 1$.
- (iii) t -Student com $\nu > 0$ conhecido: $d_0 = \frac{3(\nu+2)(\nu+3)^2}{2\nu(\nu+1)(\nu+5)(\nu+7)}, d_3 = -\frac{\nu+3}{4(\nu+1)}, d_4 = \frac{\nu+3}{2(\nu+1)}$.
- (iv) t -Student generalizada com $s, r > 0$ conhecidos: $d_0 = \frac{3(r+2)(r+3)^2}{2r(r+1)(r+5)(r+7)}, d_3 = -\frac{s(r+3)}{4r(r+1)}, d_4 = \frac{s(r+3)}{2r(r+1)}$.
- (v) Logística I: $d_0 \approx 0.4879224, d_3 \approx -0.16923449, d_4 = 0.33846899$.
- (vi) Logística II: $d_0 = \frac{3}{2}, d_3 = -\frac{3}{4}, d_4 = \frac{3}{2}$.

3.2.2 Correção de Bartlett em modelos não-lineares simétricos com ϕ desconhecido

Na Subseção anterior foi apresentada a correção de Bartlett para a estatística da razão de verossimilhanças nos modelos simétricos considerando o parâmetro escala ϕ é conhecido. Nesta seção eliminamos a suposição de que ϕ é conhecido e obtemos a correção de Bartlett para a estatística da razão de verossimilhanças associada ao teste da hipótese $H_0 : \beta_1 = \beta_1^{(0)}$ versus $H_1 : \beta_1 \neq \beta_1^{(0)}$, onde $\beta_1^{(0)}$ é um vetor especificado de dimensão $q \times 1$, $\beta_2 = (\beta_{q+1}, \dots, \beta_p)^\top$ e ϕ são parâmetros de perturbação.

Introduzimos a seguinte notação para definir os cumulantes que envolvem o parâmetro escala ϕ : $\kappa_{\phi\phi} = E[\partial^2 \ell(\theta)/\partial \phi^2]$, $\kappa_{\phi r} = E[\partial^2 \ell(\theta)/\partial \phi \partial \beta_r]$, $\kappa_{\phi rs} = E[\partial^3 \ell(\theta)/\partial \phi \partial \beta_r \partial \beta_s]$, $\kappa_{\phi\phi\phi} = E[\partial^3 \ell(\theta)/\partial \phi^3]$, etc. Então, além, dos cumulantes definidos na Subseção (3.2.1), temos os seguintes cumulantes para os modelos não-lineares simétricos:

$$\begin{aligned}
\kappa_{\phi\phi} &= -\frac{n}{\phi^2}(\delta_{(2,0,0,0,2)} - 1), \\
\kappa_{\phi,\phi} &= \frac{n}{\phi^2}(\delta_{(2,0,0,0,2)} - 1), \\
\kappa_{\phi r} &= 0, \\
\kappa_{\phi\phi}^{(\phi)} &= -\frac{2n}{\phi^3}(\delta_{(0,1,0,0,2)} - 1), \\
\kappa_{ij}^{(\phi)} &= \frac{2}{\phi^3}\delta_{(2,0,0,0,0)} \sum_{l=1}^n f_l^{i,j}, \\
\kappa_{ij\phi} &= -\frac{1}{\phi^3}(\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}) \sum_{l=1}^n f_l^{i,j}, \\
\kappa_{ij,\phi} &= \frac{1}{\phi^3}(\delta_{(0,0,1,0,1)} + 2\delta_{(2,0,0,0,0)} + 2\delta_{(0,1,0,0,0)}) \sum_{l=1}^n f_l^{i,j}, \\
\kappa_{i\phi,j} &= \frac{1}{\phi^3}(\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}) \sum_{l=1}^n f_l^{i,j}, \\
\kappa_{i,\phi,j} &= -\frac{2}{\phi^3}(\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)} + \delta_{(2,0,0,0,0)}) \sum_{l=1}^n f_l^{i,j}, \\
\kappa_{\phi\phi\phi} &= -\frac{n}{\phi^3}(6\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)} - 4), \\
\kappa_{ij\phi}^{(\phi)} &= \frac{3}{\phi^4}(\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}) \sum_{l=1}^n f_l^{i,j},
\end{aligned}$$

$$\kappa_{ij\phi\phi} = \frac{1}{\phi^4}(\delta_{(0,0,0,1,2)} - 6\delta_{(1,1,0,0,1)}) \sum_{l=1}^n f_l^{i,j}.$$

Observamos que todos os cumulantes são independentes de β e que os cumulantes relativos somente a ϕ não dependem da matriz $\tilde{X}$. Notamos também que os cumulantes $\kappa_{r\phi}$ são nulos para $r = 1, \dots, p$. Isto mostra que os parâmetros β e ϕ são globalmente ortogonais, segundo a definição de Cox e Reid (1987), nos modelos simétricos.

A estatística de teste da razão de verossimilhanças, RV , dada em (2.10), pode ser escrita como

$$RV = 2\{\{l(\hat{\beta}_1, \hat{\beta}_2, \hat{\phi}) - l(\beta_1, \beta_2, \phi)\} - \{l(\beta_1^{(0)}, \tilde{\beta}_2, \tilde{\phi}) + l(\beta_1, \beta_2, \phi)\}\},$$

onde ℓ é o logaritmo da função de verossimilhança.

Segundo Lawley (1956), sob H_0 , temos que

$$2E(l(\hat{\beta}_1, \hat{\beta}_2, \hat{\phi}) - l(\beta_1, \beta_2, \phi)) = p + 1 + \epsilon_{p+1} + O(n^{-2})$$

e

$$2E(l(\beta_1^{(0)}, \tilde{\beta}_2, \tilde{\phi}) - l(\beta_1, \beta_2, \phi)) = p - q + 1 + \epsilon_{p-q+1} + O(n^{-2}),$$

onde ϵ_{p+1} e ϵ_{p-q+1} , que tem ordem n^{-1} , são dados por

$$\begin{aligned} \epsilon_{p+1} &= \sum' (\lambda_{rstu} - \lambda_{rstuvw}) + \sum'_{\beta, \phi} (\lambda_{rstu} - \lambda_{rstuvw}) + (\ell_{\phi\phi\phi\phi} - \lambda_{\phi\phi\phi\phi\phi\phi}) \\ &= \epsilon_p + \epsilon_{p+1, \beta\phi} + \epsilon_\phi, \end{aligned}$$

$$\begin{aligned} \epsilon_{p-q+1} &= \sum'' (\lambda_{rstu} - \lambda_{rstuvw}) + \sum''_{\beta, \phi} (\lambda_{rstu} - \lambda_{rstuvw}) + (\ell_{\phi\phi\phi\phi} - \lambda_{\phi\phi\phi\phi\phi\phi}) \\ &= \epsilon_{p-q} + \epsilon_{p-q+1, \beta\phi} + \epsilon_\phi, \end{aligned}$$

sendo que $\sum'$ representa o somatório sobre todas as combinações dos parâmetros β , $\sum''$ representa o somatório sobre todas as combinações dos parâmetros β_2 , $\sum'_{\beta, \phi}$ representa o somatório sobre todas as combinações dos parâmetros de β e ϕ ($p+1$ parâmetros), mas com pelo menos um índice igual a ϕ , e $\sum''_{\beta, \phi}$ representa o somatório sobre todas as combinações dos parâmetros β_2 com pelo menos um índice igual a ϕ ($p-q+1$). Das equações acima segue que, sob H_0 , o valor esperado da estatística da razão de verossimilhanças tem a forma

$$\begin{aligned} E(RV) &= q + \epsilon_p + \epsilon_{p+1,\beta\phi} - (\epsilon_{p-q} + \epsilon_{p-q+1,\beta\phi}) + O(n^{-2}) \\ &= q(1 + d^*) + O(n^{-2}), \end{aligned}$$

onde d^* é um termo de ordem n^{-1} . Assim, o fator de correção de Bartlett é da forma $1 - d^*$, onde d^* pode ser escrito como

$$d^* = \frac{\epsilon_p - \epsilon_{p-q}}{q} + \frac{\epsilon_{p+1,\beta\phi} - \epsilon_{p-q+1,\beta\phi}}{q},$$

ou, alternativamente, como

$$d^* = d_\beta + d_{\beta\phi},$$

onde d_β é a mesma expressão do caso em que o parâmetro ϕ é conhecido e que foi apresentada na Subseção anterior e $d_{\beta\phi}$ representa a contribuição adicional devido ao fato de ϕ ser desconhecido, sendo da forma

$$d_{\beta\phi} = \frac{1}{q} \left\{ \sum_{\beta,\phi}' (\lambda_{rstu} - \lambda_{rstuvw}) - \sum_{\beta,\phi}'' (\lambda_{rstu} - \lambda_{rstuvw}) \right\}.$$

Depois de alguns cálculos, $d_{\beta\phi}$ se reduz a (ver Apêndice B)

$$d_{\beta\phi} = \frac{1}{n} \left\{ d_1 + \frac{(2p-q)}{2} d_2 \right\}, \quad (3.5)$$

sendo

$$\begin{aligned} d_1 &= -\frac{m_3}{2m_1^2} (m_2 + 2m_1 + m_3m_1) - \frac{m_4}{2m_1}, \\ d_2 &= \frac{m_3^2}{2m_1}, \end{aligned}$$

com

$$\begin{aligned} m_1 &= \delta_{(0,1,0,0,2)} - 1, \\ m_2 &= 4 - \delta_{(0,0,1,0,3)} - 6\delta_{(0,1,0,0,2)}, \\ m_3 &= \frac{(\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)})}{\delta_{(2,0,0,0,0)}}, \\ m_4 &= \frac{(\delta_{(0,0,0,1,2)} - 6\delta_{(1,1,0,0,1)})}{\delta_{(2,0,0,0,0)}}. \end{aligned}$$

Observamos que o coeficiente $d_{\beta\phi}$ dado em (3.5), além de não depender dos parâmetros β e ϕ , independe da estrutura da matriz modelo.

Assim, o fator de correção de Bartlett, $c = 1 + d^*$, para a família de modelos não lineares simétricos, é dado por

$$d^* = L_p + NL_p,$$

onde L_p e NL_p são

$$\begin{aligned} L_P &= d_0 h_0(\tilde{X}_1, \tilde{X}_2) + \frac{1}{n} \left\{ d_1 + \frac{(2p-q)}{2} d_2 \right\}, \\ NL_p &= \phi^2 \left[d_3 \{ \text{tr}(2B_d - D^2) - \text{tr}(2B_{2d} - D_2^2) \} \right. \\ &\quad \left. + d_4 \left\{ \left(\frac{1}{2} \iota^\top D Z D \iota - \text{tr}(ZB) - \frac{1}{2} \iota^\top D_2 Z_2 D_2 \iota + \text{tr}(Z_2 B_2) \right) \right\} \right] \quad (3.6), \end{aligned}$$

sendo ι um vetor $n \times 1$ de uns e $h_0(\tilde{X}_1, \tilde{X}_2) = \text{tr}(Z_d Z_d) - \text{tr}(Z_{2d} Z_{2d})$; as expressões de d_0, d_3 e d_4 foram dadas na Seção anterior.

Observe que expressão (3.6) é diferente daquela publicada em Cordeiro (2004, expressão (7) p 614). Em comunicação com o autor, ele nos afirmou que realmente alguns cumulantes tais como κ_{rstu} , $\kappa_{\phi\phi rs}$, $\kappa_{r,s,tu}$, $\kappa_{rs\phi\phi}$, $\kappa_{rs,\phi}$, $\kappa_{r,s,\phi\phi}$, continham erro e que concordava com os nossos resultados.

Obtivemos o termo ϵ_{p-q} substituindo a matriz $\tilde{X}$ pela matriz $\tilde{X}_2$. Note que L_p é uma combinação linear dos d 's, que depende somente dos δ 's de cada distribuição, com o coeficiente $h_0(\tilde{X}_1, \tilde{X}_2)$, o número de β 's, o número de parâmetros de interesse e o tamanho da amostra. O termo NL_p depende somente dos δ 's de cada distribuição e das matrizes Z, B, C e D , os quais dependem da matriz $\tilde{X}$ ou da matriz $\tilde{X}_2$, no caso do termo ϵ_{p-q} .

Para os modelos lineares simétricos ($\mu = X\beta$) a parcela NL_p se anula e o termo d^* é idêntico ao principal resultado obtido por Ferrari e Uribe-Opazo (2001).

Fornecemos, agora, os valores de d_1 e d_2 para algumas distribuições simétricas (no Apêndice D estão os δ 's para as distribuições). Cabe ressaltar que os valores de d_0, d_3, d_4 já foram fornecidos anteriormente na Subseção 3.2.1.

(i) Normal: $d_1 = 1, d_2 = 1$.

(ii) Cauchy: $d_1 = \frac{1}{2}, d_2 = 1$.

- (iii) t -Student com $\nu > 0$ conhecido: $d_1 = \frac{(\nu+2)(\nu+3)(\nu^2+9\nu+2)}{\nu(\nu+5)^2(\nu+7)}$, $d_2 = \frac{(\nu+3)(\nu+2)^2}{\nu(\nu+5)^2}$.
- (iv) t -Student generalizada com $s, r > 0$ conhecidos: $d_1 = \frac{(r+2)(r+3)(r^2+9r+2)}{r(r+5)^2(r+7)}$,
 $d_2 = \frac{(r+3)(r+2)^2}{r(r+5)^2}$.
- (v) Logística I: $d_1 \approx 1,470555027$, $d_2 \approx 1,362652631$.
- (vi) Logística II: $d_1 \approx 0,985182$, $d_2 \approx 0,786756$.

3.3 Correção tipo-Bartlett para a estatística escore

Na Seção 3.2, apresentamos os testes da razão de verossimilhanças modificado via correção de Bartlett considerando que o parâmetro de escala ϕ é conhecido e também abordamos o caso em que tal parâmetro é desconhecido, para a hipótese nula $H_0 : \beta_1 = \beta_1^{(0)}$ a ser testada versus a alternativa $H_1 : \beta_1 \neq \beta_1^{(0)}$. Assumimos, aqui, por enquanto, que ϕ seja conhecido.

Uma expansão assintótica sob a hipótese nula da distribuição da estatística escore SR para o teste discutido na Seção 2.3 foi obtida por Harris (1985).

Cordeiro e Ferrari (1991) demonstraram que qualquer estatística, S , cuja distribuição assintótica é qui-quadrado pode ser aperfeiçoada por um fator de correção multiplicativo expresso por um polinômio de grau k , de modo que os momentos da estatística são iguais aos correspondentes da distribuição qui-quadrada de referência, exceto por termos de ordem n^{-2} . A estatística modificada tem a forma:

$$S^* = S \left\{ 1 - \sum_{i=1}^k c_i S^{i-1} \right\},$$

onde os c_i 's, que são de ordem n^{-1} , são determinados de tal maneira que a distribuição de S^* sob a hipótese nula seja qui-quadrado (até esta ordem). Sob condições gerais de regularidade, a estatística S pode ser modificada por um fator de correção que envolve constantes de ordem n^{-1} e um polinômio de grau k na própria estatística, produzindo uma estatística ajustada com distribuição χ^2 até ordem n^{-1} a hipótese nula. Podemos definir a estatística escore modificada da seguinte forma:

$$SR^* = SR \left\{ 1 - (c + bSR + aSR^2) \right\}, \quad (3.7)$$

onde a, b, c , que são de ordem n^{-1} , são dados por

$$\begin{aligned} a &= \frac{A_3}{12q(q+2)(q+4)}, \\ b &= \frac{A_2 - 2A_3}{12q(q+2)}, \\ c &= \frac{A_1 - A_2 + A_3}{12q}, \end{aligned} \quad (3.8)$$

com

$$\begin{aligned} A_1 &= 3 \sum (\kappa_{ijk} + 2\kappa_{i,jk})(\kappa_{rst} + 2\kappa_{rs,t})a_{ij}a_{st}m_{kr} \\ &\quad - 6 \sum (\kappa_{ijk} + 2\kappa_{i,jk})\kappa_{r,s,t}a_{ij}a_{kr}m_{st} \\ &\quad + 6 \sum (\kappa_{ijk} - \kappa_{i,j,k})(\kappa_{rst} + 2\kappa_{rs,t})a_{ij}a_{kt}m_{ir} \\ &\quad - 6 \sum (\kappa_{i,j,k,r} + 2\kappa_{i,j,kr})a_{kr}m_{ij}, \end{aligned} \quad (3.9)$$

$$\begin{aligned} A_2 &= -3 \sum \kappa_{i,j,k}\kappa_{r,s,t}a_{kr}m_{ij}m_{st} + 6 \sum (\kappa_{ijk} + 2\kappa_{i,jk})\kappa_{r,s,t}a_{ij}m_{kr}m_{st} \\ &\quad - 6 \sum \kappa_{i,j,k}\kappa_{r,s,t}a_{kt}m_{ir}m_{jr} + 3 \sum \kappa_{i,j,k,r}m_{ij}m_{kr} \end{aligned} \quad (3.10)$$

e

$$A_3 = 3 \sum \kappa_{i,j,k}\kappa_{r,s,t}m_{ij}m_{kr}m_{st} + 2 \sum \kappa_{i,j,k}\kappa_{r,s,t}m_{ir}m_{js}m_{kt}, \quad (3.11)$$

onde as somas são tomadas em relação a todos os parâmetros do vetor β e a_{ij} e m_{ij} são os elementos das matrizes

$$A = \begin{pmatrix} 0 & 0 \\ 0 & I_{22}^{-1} \end{pmatrix}$$

e $M = I^{-1} - A$, respectivamente, onde I e I_{22}^{-1} são as matrizes de informação de Fisher e a matriz de covariância assintótica de $\tilde{\beta}_2$, respectivamente.

Em geral, SR^* é uma função de parâmetros desconhecidos. Neste caso, podemos usar a estatística $\tilde{SR}^*$ com parâmetros substituídos pelos seus respectivos estimadores de máxima verossimilhança sob H_0 . Sob H_0 e até ordem n^{-1} , as distribuições de SR^* e $\tilde{SR}^*$ coincidem (Cordeiro e Ferrari, 1991). Assim, até esta ordem, $\tilde{SR}^*$ tem também distribuição χ_q^2 . A forma da estatística escore modificada SR^* está fortemente relacionada com os quantis modificados de SR

obtidos por Harris (1985). Um teste escore aperfeiçoado pode ser reduzido, então, das seguintes maneiras: obtém-se a estatística escore aperfeiçoada dada em (3.7) e utiliza-se a distribuição χ_q^2 como referência, ou obtém-se a estatística escore usual e utilizam-se os quantis modificados como referência. Existem definições alternativas para a correção tipo-Bartlett que são equivalentes, tais como

$$SR_1^* = \frac{SR}{(1+B)} \quad \text{e} \quad SR_2^* = SR \times \exp(-B),$$

quando termos de ordem menor do que n^{-1} são ignorados, onde $B = (c + bSR + aSR^2)$. As estatísticas modificadas SR^* , SR_1^* e SR_2^* podem ser função não-monótonas de SR . Cordeiro, Ferrari e Cysneiros (1998) obtiveram uma estatística escore modificada, SR_3^* , que é assintoticamente equivalente a SR^* e tem a propriedade de ser monótona em SR . Esta estatística é dada por

$$SR_3^* = \begin{cases} \sqrt{\frac{\pi}{3a}} \exp\left(\frac{b^2}{3a} - c\right) \times \\ \left\{ \Phi\left(\sqrt{6a}SR + \sqrt{\frac{2}{3a}}b\right) - \Phi\left(\sqrt{\frac{2}{3a}}b\right) \right\} & \text{se } a > 0, \\ \frac{1}{2b} \exp(-c) \{1 - \exp(-2bSR)\} & \text{se } a = 0 \text{ e } b \neq 0. \end{cases}$$

Observe que se $a = 0$ e $b = 0$, $SR^* = SR(1 - c)$ e não há necessidade em definir uma estatística escore alternativa.

3.3.1 Correção tipo-Bartlett sobre o parâmetro β com ϕ conhecido

Os testes escore aperfeiçoados pela correção tipo Bartlett desenvolvida por Cordeiro e Ferrari (1991) dependem de quantidades A_1 , A_2 e A_3 que são funções aparentemente complicadas dos cumulantes conjuntos de derivadas do logaritmo da função de verossimilhança. O objetivo desta subseção é apresentar o cálculo para testes escores aperfeiçoados em modelos não-lineares simétricos, apresentando expressões simples para as quantidades A_1 , A_2 e A_3 .

Além dos cumulantes definidos nas Subseções anteriores, definiremos aqui mais alguns que iremos precisar para a correção tipo Bartlett para a estatística escore. São eles:

$$\begin{aligned}
\kappa_{i,jk} &= -\frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n f_l^{i,jk}, \\
\kappa_{i,j,k} &= -2\frac{\delta_{(0,0,1,0,0)}}{\phi^3} \sum_{l=1}^n f_l^{i,j,k} = 0, \\
\kappa_{i,j,k,r} &= \frac{1}{\phi^4} (\delta_{(4,0,0,0,0)} - 3\delta_{(0,1,0,0,0)}^2) \sum_{l=1}^n f_l^{i,j,k,r}, \\
\kappa_{i,j,kr} &= \frac{1}{\phi^4} (\delta_{(2,1,0,0,0)} + \delta_{(0,1,0,0,0)}^2) \sum_{l=1}^n f_l^{i,j,k,r},
\end{aligned}$$

onde $\sum$ é a soma sobre todas as observações.

Define-se também, $Q_2 = \text{diag}\{q_1, \dots, q_n\}$ sendo $q_l = \text{tr}\{(\tilde{X}_2^\top \tilde{X}_2)^{-1} D_{22}^{(l)}\}$, onde a matriz $D_{22}^{(l)}$ é obtida particionando a matriz D de dimensão $p \times p$ da mesma forma que β :

$$D^{(l)} = \frac{\partial^2 \mu_l}{\partial \beta_r \partial \beta_s} = \begin{pmatrix} D_{11}^{(l)} & D_{12}^{(l)} \\ D_{22}^{(l)} & D_{22}^{(l)} \end{pmatrix}, \quad l = 1, \dots, n.$$

Nós definimos a matriz de dimensão $n \times n$ $J_2 = \{j_{lm}\}$, onde

$$j_{lm} = \text{tr}\{D_{22}^{(l)}(\tilde{X}_2^\top \tilde{X}_2)^{-1} D_{22}^{(m)}(\tilde{X}_2^\top \tilde{X}_2)^{-1}\}.$$

Sejam Z e Z_2 as matrizes $n \times n$ definidas em (3.4). Expressões simples para os A_i 's podem ser obtidas substituindo os κ 's nas expressões (3.9) a (3.11) e efetuando as somas sobre a amostra depois de avaliar as somas sobre os parâmetros. Desta maneira, aparecerão termos da forma $\sum' a_{ij} f^{i,j}$ e $\sum' m_{ij} f^{i,j}$, sendo a_{ij} e m_{ij} os elementos (i, j) das matrizes A e M . É fácil observar que estes somatórios representam os elementos de $\phi^2/\delta_{(2,0,0,0,0)} Z_2$ e $\phi^2/\delta_{(2,0,0,0,0)} (Z - Z_2)$, respectivamente. Desta forma, obtemos as expressões dos A_i 's para os modelos não-lineares simétricos:

$$\begin{aligned}
A_{1,\beta} &= \phi^2 \left[\frac{3\delta_{(0,1,0,0,0)}^2}{\delta_{(2,0,0,0,0)}^3} \{\iota^\top Q_2 (Z - Z_2) Q_2 \iota - 2\iota^\top (Z - Z_2) \odot J_2 \iota\} \right. \\
&\quad \left. - \frac{6i}{\delta_{(2,0,0,0,0)}^2} \{\iota^\top (Z - Z_2)_d Z_{2d} \iota\} \right],
\end{aligned}$$

$$A_{2,\beta} = \phi^2 \frac{3e}{\delta_{(2,0,0,0,0)}^2} \{\iota^\top (Z - Z_2)_d^{(2)} \iota\},$$

$$A_{3,\beta} = 0, \tag{3.12}$$

onde

$$\begin{aligned} e &= \phi^{-2} \{\delta_{(4,0,0,0,0)} - 3\delta_{(0,1,0,0,0)}^2\}, \\ i &= \phi^{-2} \{\delta_{(4,0,0,0,0)} + 2\delta_{(0,1,0,0,0)}^2 + \delta_{(2,1,0,0,0)}\}. \end{aligned}$$

Substituindo na expressão dada em (3.8) os valores dos A_i 's e, conseqüentemente na expressão (3.7), obtemos a estatística escore aperfeiçoada para testar $H_0 : \beta_1 = \beta_1^{(0)}$ versus $H_1 : \beta_1 \neq \beta_1^{(0)}$.

Note que as expressões dos A_i 's são funções das matrizes $Z_2, Z - Z_2, Q_2, C_2, J_2$, do parâmetro conhecido ϕ e dos parâmetros μ_l desconhecidos, e somente envolvem operações simples com matrizes e vetores. Os A_i 's são invariantes à simetria do modelo aqui trabalhado.

Para os modelos lineares simétricos ($\mu = X\beta$) temos que as matrizes $J_2 = Q_2 = 0$ e as quantidades $A_{1,\beta}$, $A_{2,\beta}$ e $A_{3,\beta}$ coincidem com os resultados obtidos por Uribe-Opazo (1997).

A seguir são apresentados os valores de b_0 para algumas distribuições pertencentes a família de modelos simétricos não-linear (no Apêndice D estão os δ 's para cada distribuição):

- (i) Normal: $b_0 = 0$.
- (ii) Cauchy: $b_0 = \frac{1}{2}$.
- (iii) t -Student com $\nu > 0$ conhecido: $b_0 = \frac{6((\nu+2)^2-5)}{\nu(\nu+5)(\nu+7)}$.
- (iv) t -Student com $s, r > 0$ conhecidos: $b_0 = \frac{6((r+2)^2-5)}{r(r+5)(r+7)}$.
- (v) Logística I: $b_0 \approx -0,903461273$.
- (vi) Logística II: $b_0 = \frac{2}{5}$.
- (vii) Logística generalizada, com $m > 0$: $b_0 = \frac{2}{2m+3}$.
- (viii) Exponencial potência, com $-1 < k < -1/2$: $b_0 = 1 - \frac{(1-k)\Gamma(\frac{k+1}{2})\Gamma(\frac{3}{2}(1-k))}{2\Gamma^2(\frac{3-k}{2})}$.

3.3.2 Correção tipo-Bartlett sobre o parâmetro β com ϕ desconhecido

Nesta subseção, eliminamos a suposição de que o parâmetro escala, ϕ , é conhecido, sendo este agora um parâmetro de perturbação, e apresentaremos um fator de correção tipo-Bartlett para a estatística escore para testar a hipótese $H_0 : \beta_1 = \beta_1^{(0)}$ versus $H_1 : \beta_1 \neq \beta_1^{(0)}$. Lembremos que, devido à ortogonalidade global entre os parâmetros β e ϕ , a matriz de informação total de Fisher K e sua inversa K^{-1} são bloco-diagonais, ou seja,

$$K = \begin{bmatrix} K_{11} & K_{12} & 0 \\ K_{21} & K_{22} & 0 \\ 0 & 0 & \kappa_{\phi,\phi} \end{bmatrix}, \quad K^{-1} = \begin{bmatrix} K^{11} & K^{12} & 0 \\ K^{21} & K^{22} & 0 \\ 0 & 0 & \kappa_{\phi,\phi}^{-1} \end{bmatrix},$$

com

$$K_{11} = \frac{\delta_{(2,0,0,0,0)}}{\phi^2} \tilde{X}_1^\top \tilde{X}_1, \quad K_{22} = \frac{\delta_{(2,0,0,0,0)}}{\phi^2} \tilde{X}_2^\top \tilde{X}_2, \\ K_{12} = K_{21} = \frac{\delta_{(2,0,0,0,0)}}{\phi^2} \tilde{X}_1^\top \tilde{X}_2.$$

Assim, podemos definir as matrizes

$$A_\beta = \begin{bmatrix} 0 & 0 \\ 0 & K_{22}^{-1} \end{bmatrix}, \quad K_{\beta,\beta}^{-1} = \begin{bmatrix} K^{11} & K^{12} \\ K^{21} & K^{22} \end{bmatrix} \\ A = \begin{bmatrix} A_\beta & 0 \\ 0 & 0 \end{bmatrix}, \quad M = \begin{bmatrix} M_\beta & 0 \\ 0 & 0 \end{bmatrix},$$

onde $M_\beta = K_{\beta,\beta}^{-1} - A_\beta$.

Sejam $m_{r\phi}$ e $a_{r\phi}$ os $(r, p+1)$ -ésimos elementos das matrizes M e A , respectivamente, e sejam $m_{\phi\phi}$ e $a_{\phi\phi}$ os $(p+1, p+1)$ -ésimos elementos das matrizes M e A , respectivamente.

Quando o parâmetro de dispersão, ϕ , é desconhecido, Cribari-Neto e Ferrari (1995b) mostraram que, se β e ϕ são globalmente ortogonais, os A_i 's podem ser escritos como

$$A_1 = A_{1,\beta} + A_{1,\beta\phi}, \quad A_2 = A_{2,\beta} + A_{2,\beta\phi} \quad \text{e} \quad A_3 = A_{3,\beta} + A_{3,\beta\phi}, \quad (3.13)$$

onde $A_{1,\beta\phi}$, $A_{2,\beta\phi}$ e $A_{3,\beta\phi}$ são as quantidades que representam contribuições adicionais decorrentes da dispersão desconhecida, e $A_{1,\beta}$, $A_{2,\beta}$ e $A_{3,\beta}$ são os

mesmos definidos na Subseção 3.3.1, que correspondem às contribuições quando ϕ é conhecido.

Cribari-Neto e Ferrari (1995b) mostraram que as expressões dos $A_{i,\beta\phi}$ são

$$\begin{aligned}
A_{1,\beta\phi} = & 3 \sum' (\kappa_{\phi\phi k} + 2\kappa_{\phi,\phi k})(\kappa_{rst} + 2\kappa_{rs,t})a_{\phi\phi}a_{st}m_{kr} \\
& + 3 \sum' (\kappa_{ijk} + 2\kappa_{i,jk})(\kappa_{r\phi\phi} + 2\kappa_{r\phi,\phi})a_{ij}a_{\phi\phi}m_{kr} \\
& + 3 \sum' (\kappa_{\phi\phi k} + 2\kappa_{\phi,\phi k})(\kappa_{r\phi\phi} + 2\kappa_{r\phi,\phi})a_{\phi\phi}^2m_{kr} \\
& - 6 \sum' (\kappa_{\phi\phi k} + 2\kappa_{\phi,\phi k})\kappa_{r,s,t}a_{\phi\phi}a_{kr}m_{st} \\
& - 6 \sum' (\kappa_{ij\phi} + 2\kappa_{i,j\phi})\kappa_{\phi,s,t}a_{ij}a_{\phi\phi}m_{st} \\
& - 6 \sum' (\kappa_{\phi\phi\phi} + 2\kappa_{\phi,\phi\phi})\kappa_{\phi,s,t}a_{\phi\phi}^2m_{st} \\
& + 6 \sum' (\kappa_{i,\phi k} - \kappa_{i,\phi,k})(\kappa_{r\phi t} + 2\kappa_{r\phi,t})a_{\phi\phi}a_{kt}m_{ir} \\
& + 6 \sum' (\kappa_{i,j\phi} - \kappa_{i,j,\phi})(\kappa_{rs\phi} + 2\kappa_{rs,\phi})a_{\phi\phi}a_{js}m_{ir} \\
& + 6 \sum' (\kappa_{i,\phi\phi} - 2\kappa_{i,\phi,\phi})(\kappa_{r\phi\phi} + 2\kappa_{r\phi,\phi})a_{\phi\phi}^2m_{ir} \\
& - 6 \sum' (\kappa_{i,j,\phi,\phi} - \kappa_{i,j,\phi\phi})a_{\phi\phi}m_{ij},
\end{aligned}$$

$$\begin{aligned}
A_{2,\beta\phi} = & -3 \sum' \kappa_{i,j,\phi}\kappa_{\phi,s,t}a_{\phi\phi}m_{ij}m_{st} \\
& + 6 \sum' (\kappa_{\phi\phi k} + 2\kappa_{\phi,\phi k})\kappa_{r,s,t}m_{kr}a_{\phi\phi}m_{st} \\
& - 6 \sum' \kappa_{i,j,\phi}\kappa_{r,s,\phi}a_{\phi\phi}m_{ir}m_{js},
\end{aligned}$$

$$\begin{aligned}
A_{3,\beta\phi} = & 3 \sum' (\kappa_{\phi,j,k}\kappa_{r,s,t})m_{\phi j}m_{kr}m_{st} \\
& + 3 \sum' (\kappa_{i,\phi,k}\kappa_{r,s,t})m_{i\phi}m_{st}m_{kr} \\
& + 3 \sum' (\kappa_{i,j,\phi}\kappa_{r,s,t})m_{ij}m_{\phi r}m_{st} \\
& + 3 \sum' (\kappa_{i,j,k}\kappa_{\phi,s,t})m_{ij}m_{k\phi}m_{st} \\
& + 3 \sum' (\kappa_{i,j,k}\kappa_{r,\phi,t})m_{ij}m_{kr}m_{\phi t} \\
& + 3 \sum' (\kappa_{i,j,k}\kappa_{r,s,\phi})m_{ij}m_{kr}m_{s\phi} \\
& + 2 \sum' (\kappa_{\phi,j,k}\kappa_{r,s,t})m_{\phi r}m_{js}m_{kt} \\
& + 2 \sum' (\kappa_{i,\phi,k}\kappa_{r,s,t})m_{ir}m_{\phi s}m_{kt} \\
& + 2 \sum' (\kappa_{i,j,\phi}\kappa_{r,s,t})m_{ir}m_{js}m_{\phi t} \\
& + 2 \sum' (\kappa_{i,j,k}\kappa_{\phi,s,t})m_{i\phi}m_{js}m_{kt}
\end{aligned}$$

$$\begin{aligned}
& +2 \sum' (\kappa_{i,j,k} \kappa_{r,\phi,t}) m_{ir} m_{j\phi} m_{kt} \\
& +2 \sum' (\kappa_{i,j,k} \kappa_{r,s,\phi}) m_{ir} m_{js} m_{k\phi},
\end{aligned} \tag{3.14}$$

onde $\sum'$ denota soma sobre todos os parâmetros β . Como $m_{\phi r} = 0$ para todo r , então $A_{3,\beta\phi} = 0$.

Devido à simetria do modelo considerado, as expressões (3.14) se reduzem a

$$\begin{aligned}
A_{1,\beta\phi} = & -6 \sum' (\kappa_{ij\phi} + 2\kappa_{i,j\phi}) \kappa_{\phi,s,t} a_{ij} a_{\phi\phi} m_{st} \\
& -6 \sum' (\kappa_{\phi\phi\phi} + 2\kappa_{\phi,\phi\phi}) \kappa_{\phi,s,t} a_{\phi\phi}^2 m_{st} \\
& +6 \sum' (\kappa_{i,\phi k} - \kappa_{i,\phi,k}) ((\kappa_{r\phi t} + 2\kappa_{r\phi,t})) a_{\phi\phi} a_{kt} m_{ir} \\
& +6 \sum' (\kappa_{i,j\phi} - \kappa_{i,j,\phi}) ((\kappa_{rs\phi} + 2\kappa_{rs,\phi})) a_{\phi\phi} a_{js} m_{ir} \\
& -6 \sum' (\kappa_{i,j,\phi,\phi} - \kappa_{i,j,\phi\phi}) a_{\phi\phi} m_{ij},
\end{aligned}$$

$$\begin{aligned}
A_{2,\beta\phi} = & -3 \sum' \kappa_{i,j,\phi} \kappa_{\phi,s,t} a_{\phi\phi} m_{ij} m_{st} \\
& -6 \sum' \kappa_{i,j,\phi} \kappa_{r,s,\phi} a_{\phi\phi} m_{ir} m_{js},
\end{aligned}$$

$$A_{3,\beta\phi} = 0.$$

Temos, além dos cumulantes definidos nas subseções anteriores, os seguintes para os modelos simétricos não-lineares:

$$\begin{aligned}
\kappa_{i,j\phi} &= \frac{1}{\phi^3} [\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}] \sum_{l=1}^n f_l^{i,j}, \\
\kappa_{\phi,\phi\phi} &= \frac{4n}{\phi^3} \delta_{(0,0,1,0,2)} + \frac{n}{\phi^3} \delta_{(0,0,1,0,3)} + \frac{2n}{\phi^3}, \\
\kappa_{i,\phi,k} &= -\frac{2}{\phi^3} [\delta_{(0,0,1,0,1)} + 2\delta_{(2,0,0,0,0)} + 2\delta_{(0,1,0,0,0)}] \sum_{l=1}^n f_l^{i,k}, \\
\kappa_{i,\phi\phi} &= \frac{1}{\phi^3} [4\delta_{(0,1,0,0,1)} + \delta_{(0,0,1,0,2)}] \sum_{l=1}^n f_l^i, \\
\kappa_{i,\phi,\phi} &= -\frac{2}{\phi^3} [\delta_{(0,0,1,0,2)} + 2\delta_{(0,1,0,0,1)}] \sum_{l=1}^n f_l^i, \\
\kappa_{i\phi\phi} &= -\frac{1}{\phi^3} \delta_{(0,0,1,0,2)} \sum_{l=1}^n f_l^i,
\end{aligned}$$

$$\begin{aligned}
\kappa_{i,j,\phi\phi} &= \frac{1}{\phi^4} [2\delta_{(3,0,0,0,1)} - 2\delta_{(0,1,0,0,0)} + \delta_{(2,1,0,0,2)} \\
&\quad + \delta_{(0,1,0,0,0)}\delta_{(0,1,0,0,2)}] \sum_{l=1}^n f_l^{i,j}, \\
\kappa_{i,j,\phi,\phi} &= \frac{1}{\phi^4} [2\delta_{(3,0,0,0,1)} + \delta_{(4,0,0,0,2)} \\
&\quad - \delta_{(0,1,0,0,0)}\delta_{(0,1,0,0,2)}] \sum_{l=1}^n f_l^{i,j}.
\end{aligned}$$

Utilizando estes cumulantes obtemos as expressões dos $A_{i,\beta\phi}$'s na forma

$$A_{1,\beta\phi} = \frac{12b_1}{n}q(p-q) - \frac{6b_2}{n}q, \quad A_{2,\beta\phi} = -\frac{12b_3}{n}q^2 + 2q, \quad (3.15)$$

onde

$$b_1 = \frac{\delta_{(1,1,0,0,1)}(\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)})}{\delta_{(2,0,0,0,0)}^2(\delta_{(2,0,0,0,2)} - 1)},$$

$$b_2 = \{2\delta_{(1,1,0,0,1)}(2\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)})$$

$$+ (\delta_{(2,0,0,0,2)} - 1)[4\delta_{(3,0,0,0,1)} + \delta_{(4,0,0,0,2)} + \delta_{(2,1,0,0,2)} - 2\delta_{(0,1,0,0,0)}]\}$$

$$\times \{\delta_{(2,0,0,0,0)}(\delta_{(2,0,0,0,2)} - 1)^2\}^{-1},$$

$$b_3 = \frac{\delta_{(1,1,0,0,1)}^2}{\delta_{(2,0,0,0,0)}^2(\delta_{(2,0,0,0,2)} - 1)}.$$

Observa-se que os $A_{i,\beta\phi}$'s são funções do número total de parâmetros, do número de parâmetros não especificados em H_0 , dos δ 's de cada distribuição e do tamanho da amostra. Podemos escrever os coeficientes A_i 's usados na correção tipo-Bartlett do teste escore nos modelos simétricos, considerando o parâmetro de dispersão, ϕ , desconhecido, substituindo as expressões (3.15) e (3.12) em (3.13).

Para os modelos lineares simétricos ($\mu = X\beta$) temos que as quantidades $A_{1,\beta\phi}$ e $A_{2,\beta\phi}$ coincidem com os resultados obtidos por Uribe-Opazo (1997).

Abaixo estão apresentados os valores dos b_i 's para algumas distribuições pertencentes à família de modelos simétricos não-lineares (no Apêndice D estão os δ 's para cada distribuição):

- (i) Normal: $b_1 = 1, b_2 = 0, b_3 = \frac{1}{2}$.
- (ii) Cauchy: $b_1 = 0, b_2 = -1, b_3 = 0$.
- (iii) t -Student com $\nu > 0$ conhecido:

$$b_1 = \frac{(\nu - 1)(\nu + 2)(\nu + 3)}{\nu(\nu + 5)^2}, b_2 = -\frac{12(\nu + 1)(\nu + 2)(\nu + 3)}{\nu(\nu + 5)^2(\nu + 7)},$$

$$b_3 = \frac{(\nu - 1)^2(\nu + 3)}{2\nu(\nu + 5)^2}.$$

- (iv) t -Student com $s, r > 0$ conhecidos:

$$b_1 = \frac{(r - 1)(r + 2)(r + 3)}{r(r + 5)^2}, b_2 = -\frac{12(r + 1)(r + 2)(r + 3)}{r(r + 5)^2(r + 7)},$$

$$b_3 = \frac{(r - 1)^2(r + 3)}{2r(r + 5)^2}.$$

- (v) Logística I: $b_1 \approx 1,77437106, b_2 \approx 0,569031708, b_3 \approx 1,155242856$.
- (vi) Logística II: $b_1 \approx 0,52446, b_2 \approx -0,5835, b_3 \approx 0,174816$.
- (vii) Exponencial potência, com $-1 < k < -1/2$:

$$b_1 = \frac{1 - k}{1 + k}, b_2 = \frac{2(k - 1)k}{1 + k}, b_3 = \frac{(1 - k)^2}{2(1 + k)}.$$

Resultados numéricos

Neste capítulo, nosso objetivo é comparar, através de simulações de Monte Carlo, os desempenhos de testes baseados nas estatísticas da razão de verossimilhanças e escore usuais e em suas respectivas versões modificadas. Avaliaremos a eficácia das correções de Bartlett e tipo-Bartlett para os testes da razão de verossimilhanças e escore, respectivamente, em modelos de regressão não-lineares simétricos. Os testes a serem avaliados são: razão de verossimilhanças usual (RV), razão de verossimilhanças corrigida via correção de Bartlett (RV^*), sua versão assintoticamente equivalente (RV^{**}), razão de verossimilhanças bootstrap (RV_B), escore usual (SR), escore corrigida via correção tipo-Bartlett (SR^*), sua versão alternativa (SR_3^*) e escore bootstrap (SR_B).

Inicialmente, temos como objetivo analisar a influência do número de parâmetros de perturbação e do tamanho de amostra nos desempenhos dos testes e compará-los entre si. Os desempenhos são avaliados em função da proximidade das probabilidades efetivas de rejeição da hipótese nula, (probabilidade do erro tipo I) aos respectivos valores nominais. Comparamos a média, a variância e quantis amostrais das estatísticas de testes com as quantidades correspondentes da distribuição de referência e avaliamos os tamanhos empíricos dos testes para valores bem pequenos do nível nominal.

Em todas as simulações a componente sistemática é dada por

$$\mu_l = \beta_1 x_{l1} + \cdots + \beta_{p-2} x_{lp-2} + \beta_{p-1} \exp(\beta_p x_{lp}), \quad l = 1, \dots, n.$$

A hipótese nula considerada é $H_0 : \beta_p = 0$ e a hipótese alternativa é $H_1 : \beta_p \neq 0$. A variável resposta foi gerada assumindo que os β_l 's são iguais a um; o valor de ϕ foi fixado em 1,0 para todas as distribuições.

Os valores das covariadas $x_1, \dots, x_k$ foram geradas como realizações independentes da distribuição $\mathcal{U}(0, 1)$. O número de réplicas foi fixado em 10,000 réplicas de Monte Carlo e 500 réplicas de bootstrap para diversos tamanhos de amostras ($n = 10, 20, 30, 40, 50$) e foram considerados os seguintes níveis nominais: $\alpha = 10\%, 5\%, 1\%, 0, 5\%$. As simulações foram realizadas utilizando o linguagem de programação matricial **0x** (Doornik, 2001).

Para cada tamanho de amostra e cada nível nominal considerado, calculamos as taxas de rejeição de cada teste, ou seja, estimamos, via simulação

$P(RV > x_\alpha)$, $P(RV^* > x_\alpha)$, $P(RV^{**} > x_\alpha)$, $P(SR > x_\alpha)$, $P(SR^* > x_\alpha)$, $P(SR_3^* > x_\alpha)$, $P(RV > z_\alpha)$ e $P(SR > z_\alpha)$, onde x_α é o valor crítico obtido da distribuição χ_q^2 e z_α é encontrado da forma definida na Seção 2.6. Vale recordar que $RV^* = RV/(1+d)$, $RV^{**} = RV(1-d)$, $SR^* = SR\{1 - (c + bSR + aSR^2)\}$ e SR_3^* é a estatística modificada monótona em SR dada na Subseção 3.3.1.

Todas as entradas dos quadros apresentadas correspondem a porcentagens, com exceção daqueles dos Quadros 4.11-4.19. O estudo de simulação visa a analisar a influência do número de parâmetros de perturbação sobre os desempenhos dos testes.

Os resultados numéricos baseiam-se no seguinte modelo:

$$y_l = \mu_l + \xi_l, \quad l = 1, \dots, n,$$

em que $\xi_l \sim S(0, \phi^2)$ são erros independentemente distribuídos, onde S representa uma distribuição do modelo simétrico. Foram considerados 4 componentes sistemáticas para cada distribuição, a saber:

- (1) $\mu_l = \beta_4 \exp(\beta_5 x_5) \quad (p - q = 0)$,
- (2) $\mu_l = \beta_3 x_3 + \beta_4 \exp(\beta_5 x_5) \quad (p - q = 1)$,
- (3) $\mu_l = \beta_2 x_2 + \beta_3 x_3 + \beta_4 \exp(\beta_5 x_5) \quad (p - q = 2)$,
- (3) $\mu_l = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 \exp(\beta_5 x_5) \quad (p - q = 3)$.

A hipótese nula considerada é $H_0 : \beta_5 = 0$, indicando assim que o modelo sob a hipótese nula é linear. Dez mil amostras foram geradas para tamanho de amostra $n = 30$. Os resultados destas simulações são apresentados nos Quadros 4.6 a 4.10.

Nos Quadros 4.1 a 4.10 foram omitidos os resultados de simulação relativo ao teste baseado na estatística RV^* pois apresentou um desempenho similar ao teste RV^{**} .

Nos Quadros 4.1 a 4.5 apresentamos os tamanhos observados dos testes para a situação em que $p = 2$ e diferentes valores de n foram considerados ($n = 20, 30, 40, 50$). A análise destes quadros mostra que a estatística da razão de verossimilhanças (RV) não tem boa aproximação pela distribuição χ^2 e apresenta as taxas de rejeição bem maiores do que os níveis nominais, mesmo quando o número total de observações é razoavelmente grande ($n = 50$). Além disto, existe forte evidência de que a aproximação por χ^2 para os testes modificados é melhor do que para os testes da razão de verossimilhanças e escore usuais. Claramente, podemos observar que os desempenhos dos testes baseados na estatística da razão de verossimilhanças modificada via correção de Bartlett (RV^*)

e em sua estatística equivalente (RV^{**}) são melhores do que o teste baseado na estatística da razão de verossimilhanças usual (RV), sendo que, entre as estatísticas RV^* e RV^{**} aquela que produz melhor teste é a estatística RV^{**} . Já entre as estatísticas SR^* e SR_3^* , uma ordenação não é evidente. Os testes baseados na estatística escore modificada via correção de Bartlett (SR^*) e em sua versão equivalente (SR_3^*), na maioria das vezes, são conservativos a taxas de 1% e 0,5% pois exibem taxas de rejeição menores que os níveis nominais. As versões bootstrap RV_B e SR_B são os testes com melhores desempenhos, pois são os que mais atenuam a tendência dos testes a rejeitar demais a hipótese nula. Notamos que, conforme cresce o tamanho da amostra, as taxas de rejeição de todos os testes aproximam-se dos respectivos níveis nominais e, como era de se esperar, as correções vão sendo cada vez menos necessárias.

Nos Quadros 4.6 a 4.10 apresentamos os tamanhos observados dos testes para a situação em que fixamos o tamanho amostral em $n = 30$ para diferentes valores de p ($p = 2, 3, 4, 5$). Os desempenhos dos testes foram bastante similares aos obtidos anteriormente. Porém, notamos que o número de parâmetros de perturbação tem impacto considerável sobre a qualidade da aproximação χ^2 para as estatísticas RV , RV^* e sua versão equivalente RV^{**} . Isto também pode ser notado nas figuras 4.6 a 4.10, onde mostramos os gráficos dos tamanhos dos testes versus o número de elementos do vetor β . Por exemplo no Quadro 4.10 referente à distribuição logística tipo II, para $p = 5$ e $\alpha = 10\%$, estes testes são liberais, apresentando taxas de rejeição 17,24% e 14,00%, respectivamente, bem maiores do que o nível nominal. Este fato é bem mais acentuado no caso do teste RV , pois a correção de Bartlett atenua a tendência de rejeição demasiada da hipótese nula, porém não o bastante. É importante notar que o impacto do número de parâmetros de perturbação é bem menos marcante na versão bootstrap e nos outros testes baseados na estatística escore e sua correções. Para estes testes, quando o número de parâmetros de perturbação aumenta, as taxas de rejeição permanecem mais estáveis em relação aos respectivos valores nominais do que quando os testes são baseados na estatística da razão de verossimilhanças (RV), na sua versão corrigida via correção de Bartlett (RV^*) e sua estatística equivalente (RV^{**}). Como exemplo, podemos citar o Quadro 4.7, referente à distribuição Cauchy, onde as taxas de rejeição são 14,36%, 11,75% para os testes RV , RV^{**} , respectivamente, e são 5,60%, 5,64%, 5,61%, 5,90%, 5,65% para os testes RV_B , SR , SR^* , SR_3^* , SR_B , respectivamente, considerando $\alpha = 5\%$ e

$p = 4$.

Nos Quadros 4.11 a 4.14 apresentamos comparações das médias e variâncias das estatísticas RV , RV^* , RV^{**} , SR , SR^* , SR_3^* e também da distribuição χ_1^2 , para algumas distribuições e alguns valores de n e p . Os resultados contidos nessas Quadros mostram que as estatísticas corrigidas e a estatística escore são as que apresentam médias mais próximas da média da distribuição χ_1^2 . No caso da comparação das variâncias, notamos que as estatísticas cujas variâncias mais se aproximam da variância de uma varável aleatória χ_1^2 são a estatística escore (SR) e suas versões corrigidas, SR^* e SR_3^* . Por outro lado, as variâncias da estatística RV e de suas versões corrigidas, RV^* e RV^{**} , excedem a variância da distribuição de referência. Este mesmo comportamento foi observado nos Quadros 4.15 a 4.19 ao comparar os quantis amostrais de RV , RV^* , RV^{**} , SR , SR^* e SR_3^* com os da distribuição χ_1^2 . Por exemplo, no Quadro 4.17, referente à distribuição t -Student com 2 graus de liberdade, em todos os quantis, temos que as estatísticas SR , SR^* e SR_3^* apresentam quantis mais próximos ao da distribuição χ_1^2 , enquanto os quantis amostrais das estatísticas RV , RV^* , RV^{**} excederam o da distribuição χ_1^2 .

As Figuras 4.1 a 4.5 apresentam os gráficos dos quantis das estatísticas RV , RV^* , RV^{**} , SR , SR^* e SR_3^* versus os quantis assintóticos (qui-quadrado com um grau de liberdade) para tamanho amostral $n = 30$. Observe que as estatísticas dos testes da razão de verossimilhanças, da corrigida via Bartlett e a assintoticamente equivalente, apresentam quantis bem superiores aos quantis da distribuição qui-quadrado com um grau de liberdade, enquanto os quantis das estatísticas dos outros testes (RV_B , SR , SR^* , SR_B e SR_3^*) estão bem próximas dos quantis da distribuição qui-quadrado. Porém as estatísticas dos testes RV^* , RV^{**} , apresentam quantis mais próximos do que a estatística da razão de verossimilhanças, mostrando que a distribuição nula destas estatísticas são mais bem aproximadas pela distribuição assintótica utilizada do que a razão de verossimilhanças.

As Figuras 4.11 a 4.15 apresentam as distorções de tamanho dos testes RV , RV^{**} , RV_B , SR , SR^* , SR_B para diferentes níveis de significância α . Observe que o teste da razão de verossimilhanças apresenta tamanho estimado mais distorcido que os outros testes, principalmente em tamanhos de amostras pequenos. Note que, os testes escore, escore corrigido e escore bootstrap apresentam distorções de tamanhos próximas de zero em todos os tamanhos de amostras e níveis

de significância considerados.

Figura 4.1 Gráfico dos quantis - Distribuição Normal.

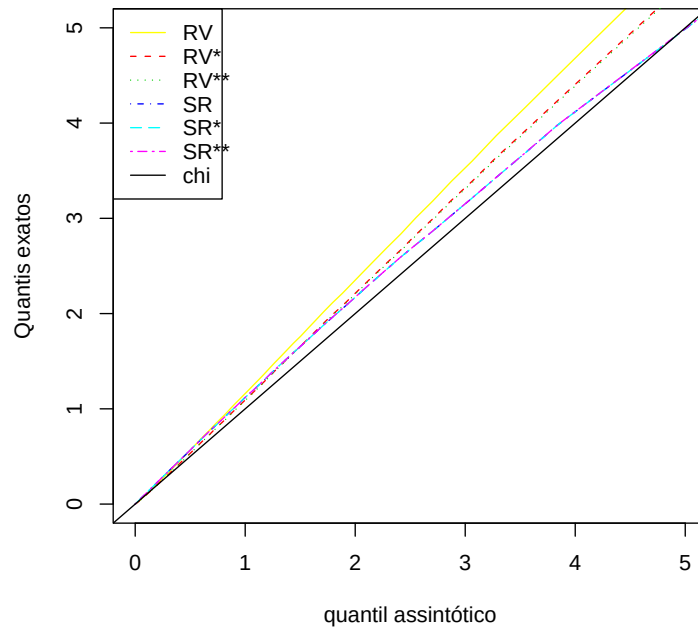


Figura 4.2 Gráfico dos quantis - Distribuição Cauchy.

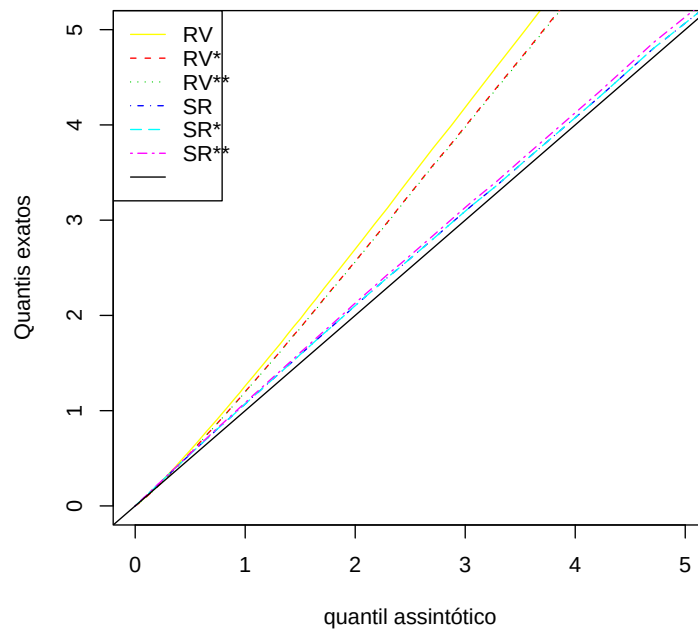


Figura 4.3 Gráfico dos quantis - Distribuição t -Student 2 grau de liberdade.

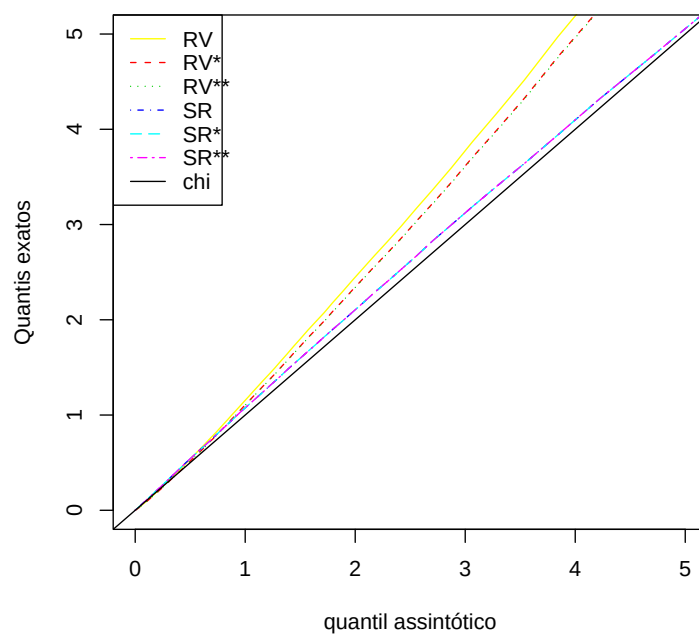


Figura 4.4 Gráfico dos quantis - Distribuição t -Student 4 grau de liberdade.

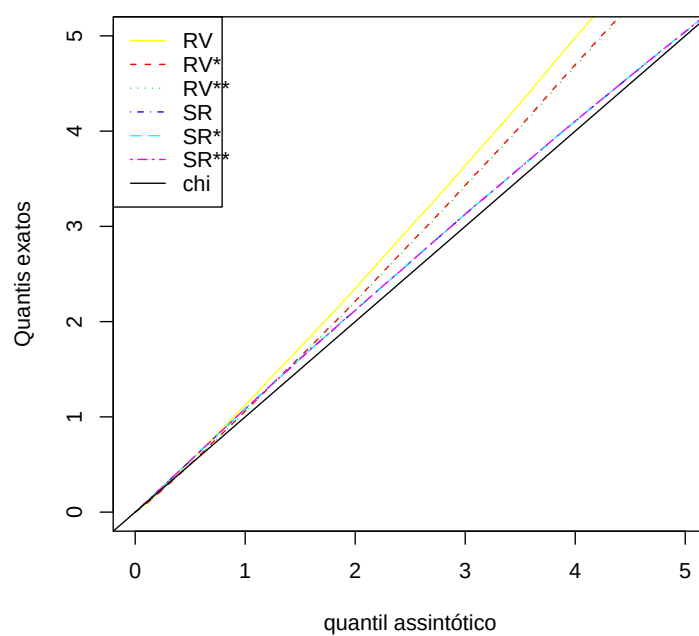
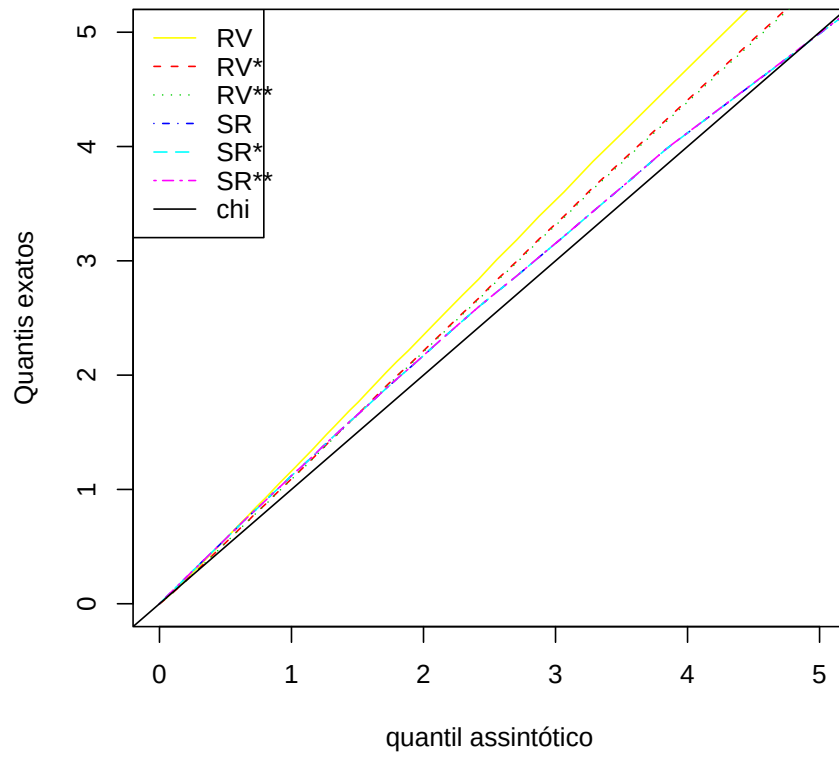


Figura 4.5 Gráfico dos quantis - Distribuição Logística.



Quadro 4.1. Taxas de rejeição (%), distribuição normal, $p = 2$.

teste	α	n			
		20	30	40	50
RV	10	14,08	13,27	12,86	12,01
	5	7,53	7,06	6,68	6,42
	1	2,12	1,55	1,42	1,35
	0,5	1,17	0,97	0,73	0,70
RV^{**}	10	11,36	10,94	10,68	10,91
	5	5,76	5,83	5,67	5,63
	1	1,32	1,22	1,04	1,01
	0,5	0,78	0,76	0,57	0,55
RV_B	10	10,35	10,30	10,27	10,15
	5	5,27	4,94	5,23	5,03
	1	1,25	1,11	1,09	1,03
	0,5	0,69	0,59	0,58	0,55
SR	10	11,30	10,61	10,30	10,27
	5	5,30	4,84	5,20	5,17
	1	0,84	0,86	1,00	1,00
	0,5	0,27	0,36	0,42	0,46
SR^*	10	10,33	10,07	9,86	10,05
	5	4,88	4,89	4,99	5,05
	1	0,99	0,99	1,09	1,03
	0,5	0,46	0,47	0,50	0,50
SR_B	10	10,24	10,25	10,16	10,07
	5	5,24	5,12	5,09	5,07
	1	1,20	1,15	1,12	1,11
	0,5	0,67	0,63	0,57	0,53
SR_3^*	10	10,37	10,09	9,88	10,05
	5	4,89	4,90	5,00	5,00
	1	1,00	0,94	1,09	1,05
	0,5	0,46	0,47	0,51	0,50

Quadro 4.2. Taxas de rejeição (%), distribuição Cauchy, $p = 2$.

teste	α	n			
		20	30	40	50
RV	10	19,87	17,11	14,63	13,93
	5	11,23	10,79	8,70	8,30
	1	4,74	3,78	2,91	2,71
	0,5	2,43	2,22	1,75	1,71
RV^{**}	10	16,20	15,70	13,75	13,34
	5	10,70	9,52	7,88	7,82
	1	4,85	3,09	2,63	2,43
	0,5	1,78	1,76	1,48	1,43
RV_B	10	10,83	10,34	10,23	10,11
	5	5,91	5,77	5,38	5,35
	1	1,44	1,42	1,20	1,16
	0,5	0,83	0,82	0,75	0,74
SR	10	12,83	11,30	10,62	10,09
	5	7,46	5,74	5,33	4,84
	1	1,82	0,96	1,04	1,02
	0,5	0,37	0,43	0,48	0,53
SR^*	10	11,34	10,99	10,44	10,34
	5	6,16	5,53	5,22	5,16
	1	0,69	0,93	1,01	1,11
	0,5	0,35	0,38	0,47	0,55
SR_B	10	11,20	10,28	10,05	9,97
	5	5,62	5,27	5,12	4,99
	1	1,46	1,39	1,15	1,14
	0,5	0,82	0,75	0,63	0,63
SR_3^*	10	11,83	11,30	10,62	10,09
	5	6,61	5,74	5,33	4,84
	1	0,83	0,96	1,04	1,06
	0,5	0,37	0,43	0,48	0,58

Quadro 4.3. Taxas de rejeição (%), distribuição t -Student
com $\nu = 2$ graus de liberdade, $p = 2$.

teste	α	n			
		20	30	40	50
RV	10	14,73	14,67	13,19	12,64
	5	8,87	8,80	7,52	7,15
	1	2,77	2,74	2,39	1,86
	0,5	1,66	1,63	1,34	0,97
RV^{**}	10	13,07	13,05	12,46	12,02
	5	7,64	7,63	7,02	6,55
	1	2,09	2,06	2,03	1,69
	0,5	1,30	1,29	1,19	0,86
RV_B	10	9,74	10,34	10,44	9,61
	5	4,97	5,77	5,20	4,97
	1	1,10	1,42	1,16	1,12
	0,5	0,71	0,82	0,62	0,64
SR	10	11,89	11,45	11,37	10,65
	5	6,44	5,65	5,66	5,56
	1	1,24	1,07	1,02	1,05
	0,5	0,60	0,48	0,55	0,56
SR^*	10	11,34	11,01	11,09	10,41
	5	5,86	5,38	5,22	5,51
	1	1,04	0,97	0,97	0,97
	0,5	0,52	0,43	0,49	0,44
SR_B	10	10,83	10,28	10,86	10,59
	5	5,49	5,27	5,63	5,40
	1	1,46	1,39	1,09	1,28
	0,5	0,67	0,65	0,60	0,57
SR_3^*	10	11,35	11,02	11,09	10,41
	5	5,87	5,38	5,53	5,51
	1	1,28	0,97	0,97	0,97
	0,5	0,73	0,43	0,49	0,44

Quadro 4.4. Taxas de rejeição (%), distribuição t -Student,
com $\nu = 4$ graus de liberdade, $p = 2$.

teste	α	n			
		20	30	40	50
RV	10	15,30	14,17	13,56	11,51
	5	9,04	8,49	7,69	6,39
	1	2,78	2,48	2,12	1,59
	0,5	1,51	1,35	1,25	0,82
RV^{**}	10	12,51	12,01	11,72	10,93
	5	7,14	6,90	7,07	5,92
	1	1,86	1,83	1,88	1,42
	0,5	1,01	1,08	0,99	0,72
RV_B	10	9,93	10,03	9,70	10,04
	5	4,99	5,16	4,86	5,33
	1	1,17	1,13	1,21	1,08
	0,5	0,75	0,70	0,76	0,64
SR	10	11,82	11,20	11,23	10,67
	5	6,08	5,56	5,49	5,27
	1	1,14	1,07	0,98	1,01
	0,5	0,67	0,47	0,53	0,50
SR^*	10	10,94	10,76	10,85	10,45
	5	5,63	5,26	5,23	5,08
	1	0,98	0,87	0,89	0,94
	0,5	0,41	0,37	0,50	0,43
SR_B	10	10,52	10,45	10,38	10,21
	5	5,62	5,53	5,51	5,24
	1	1,38	1,35	1,33	1,17
	0,5	0,82	0,81	0,74	0,65
SR_3^*	10	10,96	10,85	10,76	10,45
	5	5,63	5,26	5,23	5,08
	1	0,98	0,87	0,90	0,94
	0,5	0,43	0,37	0,50	0,43

Quadro 4.5. Taxas de rejeição (%), distribuição logística, $p = 2$.

teste	α	n			
		20	30	40	50
RV	10	13,01	12,34	12,20	11,28
	5	7,43	6,61	6,39	5,85
	1	1,90	1,68	1,44	1,39
	0,5	1,08	0,90	0,83	0,76
RV^{**}	10	11,91	11,61	10,40	10,34
	5	6,53	6,06	5,33	5,17
	1	1,57	1,53	1,18	1,01
	0,5	0,84	0,82	0,78	0,51
RV_B	10	10,32	10,28	10,09	9,64
	5	5,17	5,13	5,10	5,03
	1	1,27	1,21	1,18	1,17
	0,5	0,73	0,70	0,58	0,52
SR	10	11,53	11,08	10,87	10,52
	5	5,46	5,43	5,36	5,27
	1	0,77	1,01	1,06	0,98
	0,5	0,32	0,44	0,61	0,51
SR^*	10	10,97	10,83	10,53	10,43
	5	5,26	5,25	4,75	5,03
	1	0,50	0,82	0,83	0,83
	0,5	0,20	0,29	0,33	0,34
SR_B	10	11,47	10,34	10,24	10,21
	5	6,15	5,30	5,24	5,25
	1	1,52	1,33	1,14	1,15
	0,5	0,80	0,67	0,60	0,57
SR_3^*	10	10,84	10,54	10,48	10,06
	5	4,76	5,05	5,21	5,00
	1	0,50	0,83	1,00	0,88
	0,5	0,23	0,30	0,51	0,44

Quadro 4.6. Taxas de rejeição (%), distribuição normal, $n = 30$.

teste	α	p			
		2	3	4	5
RV	10	13,27	14,80	15,52	16,65
	5	7,06	8,36	8,90	9,80
	1	1,55	2,23	2,40	2,51
	0,5	0,97	1,22	1,29	1,47
RV^{**}	10	10,94	12,37	12,46	12,50
	5	5,83	6,26	6,34	6,41
	1	1,22	1,36	1,37	1,37
	0,5	0,76	0,80	0,80	0,85
RV_B	10	10,30	10,64	10,74	10,83
	5	4,94	5,25	5,74	5,69
	1	1,11	1,09	1,25	1,36
	0,5	0,59	0,78	0,67	0,69
SR	10	10,61	11,46	12,17	12,89
	5	4,84	5,63	5,96	6,68
	1	0,86	1,02	1,23	1,33
	0,5	0,36	0,45	0,52	0,59
SR^*	10	10,07	10,18	10,19	10,20
	5	4,89	4,95	4,91	5,24
	1	0,99	1,01	1,05	1,07
	0,5	0,47	0,47	0,48	0,51
SR_B	10	10,25	10,50	10,56	10,88
	5	5,12	5,18	5,37	5,60
	1	1,15	1,06	1,07	1,14
	0,5	0,63	0,60	0,63	0,67
SR_3^*	10	10,09	10,21	10,25	10,35
	5	4,90	5,04	4,98	5,34
	1	0,94	1,01	1,07	1,08
	0,5	0,47	0,49	0,50	0,51

Quadro 4.7. Taxas de rejeição (%), distribuição Cauchy, $n = 30$.

teste	α	p			
		2	3	4	5
RV	10	17,11	19,91	21,76	25,46
	5	10,79	12,75	14,36	17,05
	1	3,78	4,43	5,64	6,85
	0,5	2,22	2,64	3,52	4,64
RV^{**}	10	15,70	17,61	18,53	20,69
	5	9,52	10,85	11,75	13,33
	1	3,09	3,22	3,82	4,51
	0,5	1,76	1,89	2,19	2,72
RV_B	10	10,34	10,42	11,01	11,54
	5	5,77	5,36	5,60	5,66
	1	1,42	1,21	1,38	1,45
	0,5	0,82	0,75	0,85	0,89
SR	10	10,79	11,48	11,49	11,52
	5	5,49	5,36	5,64	5,90
	1	0,83	0,92	1,04	1,07
	0,5	0,36	0,38	0,42	0,48
SR^*	10	10,79	11,12	11,14	11,19
	5	4,91	5,21	5,61	5,38
	1	0,69	0,85	1,00	1,00
	0,5	0,24	0,35	0,44	0,37
SR_B	10	10,28	10,04	10,57	10,63
	5	5,27	5,09	5,65	5,72
	1	1,39	1,22	1,40	1,42
	0,5	0,75	0,76	0,81	0,86
SR_3^*	10	10,79	11,49	11,48	11,52
	5	5,49	5,36	5,90	5,64
	1	0,82	0,92	1,04	1,07
	0,5	0,36	0,38	0,48	0,42

Quadro 4.8. Taxas de rejeição (%), distribuição t -Student $\nu = 2$, $n = 30$.

teste	α	p			
		2	3	4	5
RV	10	14,67	17,46	18,02	19,84
	5	8,80	10,66	10,76	12,79
	1	2,74	3,35	3,31	4,16
	0,5	1,63	2,05	2,05	2,50
RV^{**}	10	13,05	12,21	13,29	14,91
	5	7,63	6,40	7,85	8,39
	1	2,06	1,49	2,19	2,27
	0,5	1,29	0,80	1,35	1,34
RV_B	10	9,74	10,34	10,79	11,18
	5	4,97	5,37	5,34	5,64
	1	1,10	1,23	1,18	1,32
	0,5	0,71	0,66	0,75	0,75
SR	10	11,45	11,84	12,28	12,51
	5	5,65	5,80	6,20	6,29
	1	1,07	1,11	1,13	1,15
	0,5	0,48	0,48	0,53	0,65
SR^*	10	11,01	11,33	11,53	11,87
	5	5,38	5,30	5,57	5,88
	1	0,97	0,95	0,99	0,99
	0,5	0,43	0,41	0,51	0,41
SR_B	10	10,28	11,07	10,34	11,40
	5	5,27	5,39	5,25	5,86
	1	1,39	1,24	1,23	1,23
	0,5	0,75	0,70	0,80	0,75
SR_3^*	10	11,02	11,33	11,53	11,87
	5	5,38	5,30	5,58	5,89
	1	0,97	0,95	0,99	0,99
	0,5	0,43	0,41	0,51	0,41

Quadro 4.9. Taxas de rejeição (%), distribuição t -Student $\nu = 4$, $n = 30$.

teste	α	p			
		2	3	4	5
RV	10	14,17	15,30	15,90	17,31
	5	8,49	9,04	9,78	10,57
	1	2,48	2,78	3,02	3,45
	0,5	1,35	1,51	1,73	1,81
RV^{**}	10	12,01	13,45	13,78	14,00
	5	6,90	7,63	7,90	8,23
	1	1,83	2,10	2,05	2,06
	0,5	1,08	1,10	1,07	1,15
RV_B	10	10,03	10,06	10,28	10,82
	5	5,16	5,02	5,12	5,51
	1	1,13	1,13	1,17	1,27
	0,5	0,70	0,68	0,72	0,78
SR	10	11,21	11,79	12,30	13,02
	5	5,56	6,22	6,44	6,52
	1	1,00	1,17	1,22	1,28
	0,5	0,47	0,52	0,59	0,60
SR^*	10	10,76	11,05	11,37	11,79
	5	5,26	5,52	5,68	5,70
	1	0,87	0,97	0,92	0,95
	0,5	0,37	0,39	0,39	0,40
SR_B	10	10,52	10,58	10,67	10,83
	5	5,53	5,55	5,40	5,68
	1	1,35	1,20	1,23	1,24
	0,5	0,67	0,72	0,75	0,87
SR_3^*	10	10,85	11,06	11,39	11,84
	5	5,26	5,53	5,69	5,70
	1	0,87	0,97	0,93	0,95
	0,5	0,37	0,39	0,39	0,40

Quadro 4.10. Taxas de rejeição (%), distribuição Logística, $n = 30$.

teste	α	p			
		2	3	4	5
RV	10	12,34	15,15	16,06	17,24
	5	6,61	8,88	9,26	10,34
	1	1,60	2,40	2,62	3,01
	0,5	0,90	1,45	1,58	1,81
RV^{**}	10	11,91	13,32	13,41	14,00
	5	6,01	7,44	7,24	7,88
	1	1,53	1,78	1,80	1,91
	0,5	0,82	0,97	0,98	1,03
RV_B	10	10,28	10,30	10,82	10,83
	5	5,13	5,27	5,34	5,73
	1	1,24	1,25	1,28	1,29
	0,5	0,70	0,69	0,72	0,78
SR	10	11,08	11,74	12,60	13,11
	5	5,43	5,94	6,19	6,66
	1	1,01	1,13	1,27	1,31
	0,5	0,44	0,57	0,59	0,63
SR^*	10	10,83	10,92	11,43	11,38
	5	5,25	5,25	5,38	5,48
	1	0,82	0,83	0,83	0,87
	0,5	0,29	0,33	0,34	0,35
SR_B	10	10,34	10,47	10,56	10,97
	5	5,30	5,37	5,40	5,47
	1	1,33	1,32	1,32	1,41
	0,5	0,67	0,77	0,77	0,78
SR_3^*	10	10,54	10,94	11,47	11,44
	5	5,05	5,21	5,26	5,43
	1	0,83	0,87	0,88	0,90
	0,5	0,30	0,33	0,34	0,35

Figura 4.6 Tamanho dos testes, $n = 30$, distribuição normal, $\alpha = 10\%$.

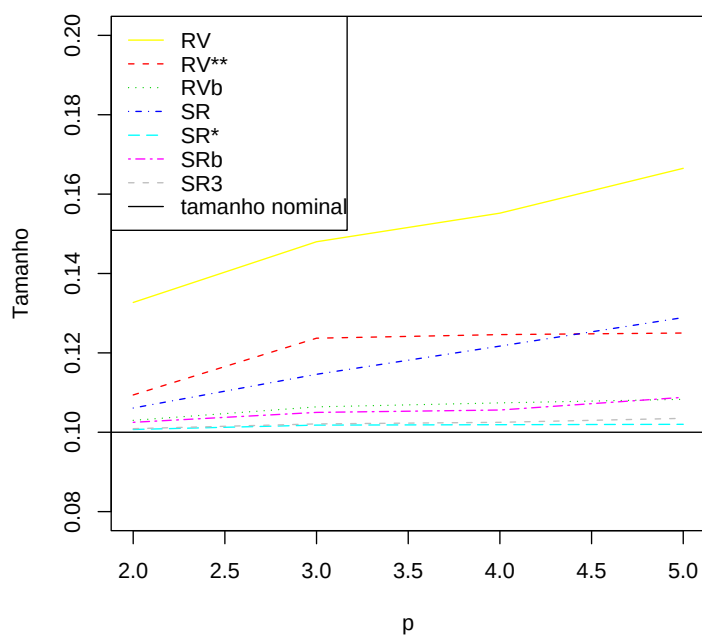


Figura 4.7 Tamanho dos testes, $n = 30$, Distribuição Cauchy, $\alpha = 5\%$.

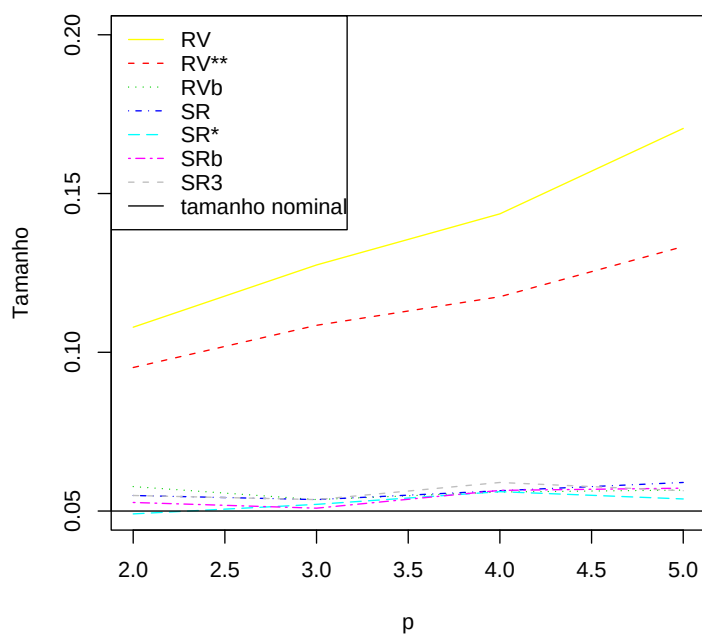


Figura 4.8 Tamanho dos testes, $n = 30$, Distribuição t -Student 2 graus de liberdade, $\alpha = 1\%$.

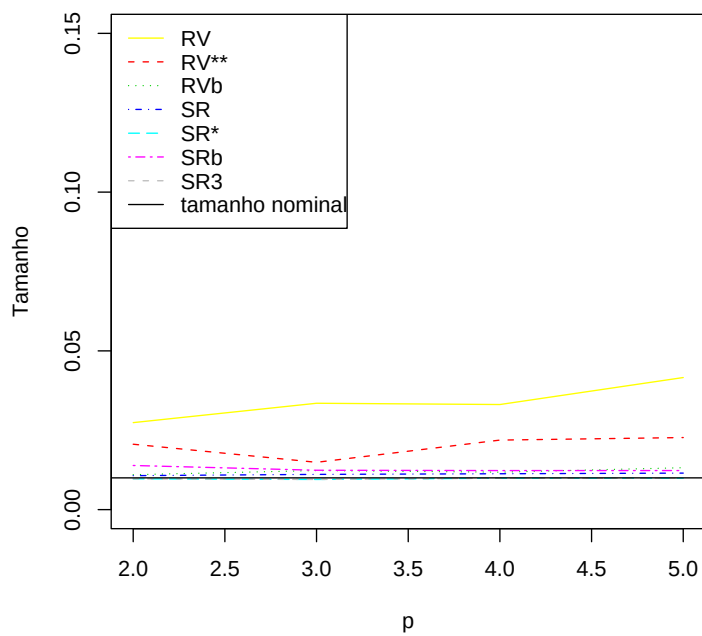


Figura 4.9 Tamanho dos testes, $n = 30$, Distribuição t -Student 4 graus de liberdade, $\alpha = 0.5\%$.

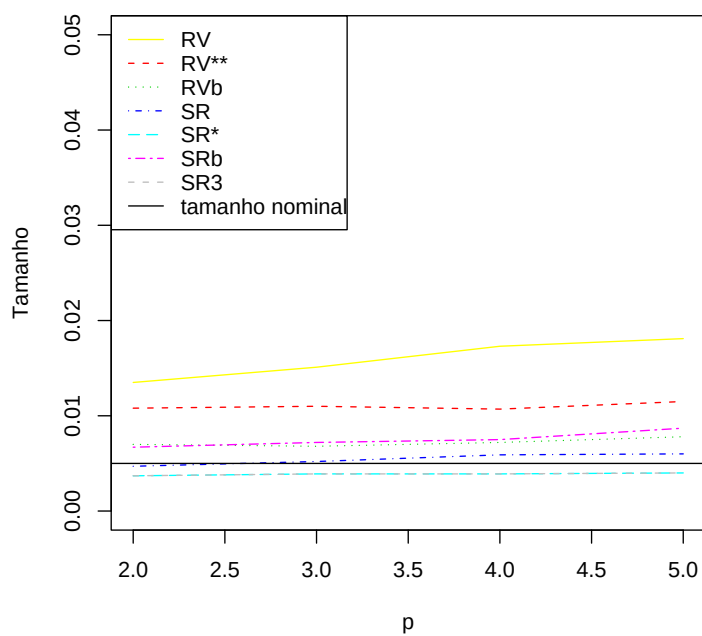


Figura 4.10 Tamanho dos testes, $n = 30$, Distribuição Logística II, $\alpha = 10\%$.

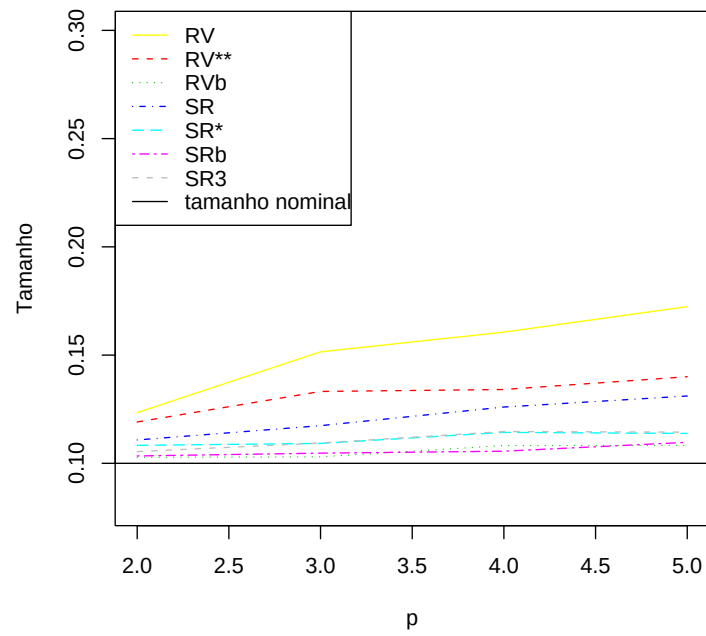


Figura 4.11 Distorção de tamanho dos testes, $p = 2$, Distribuição Normal, $\alpha = 0.5\%$.

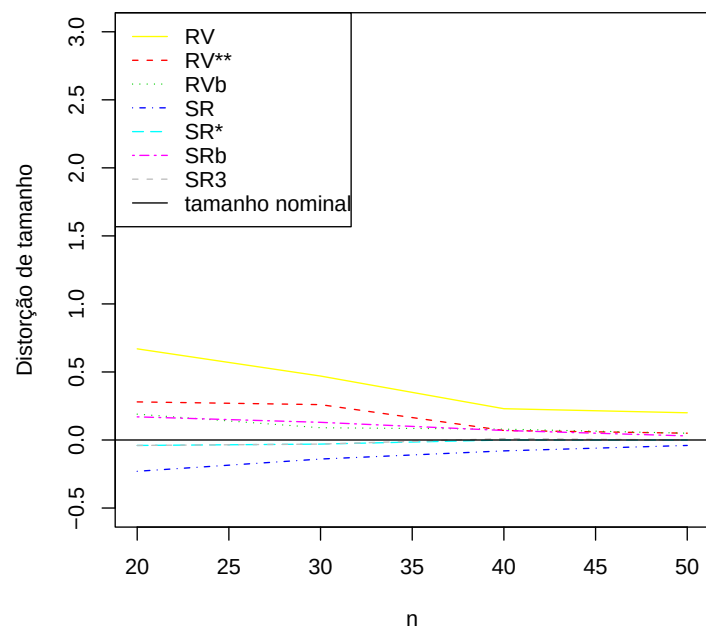


Figura 4.12 Distorção de tamanho dos testes, $p = 2$, Distribuição Cauchy, $\alpha = 1\%$.

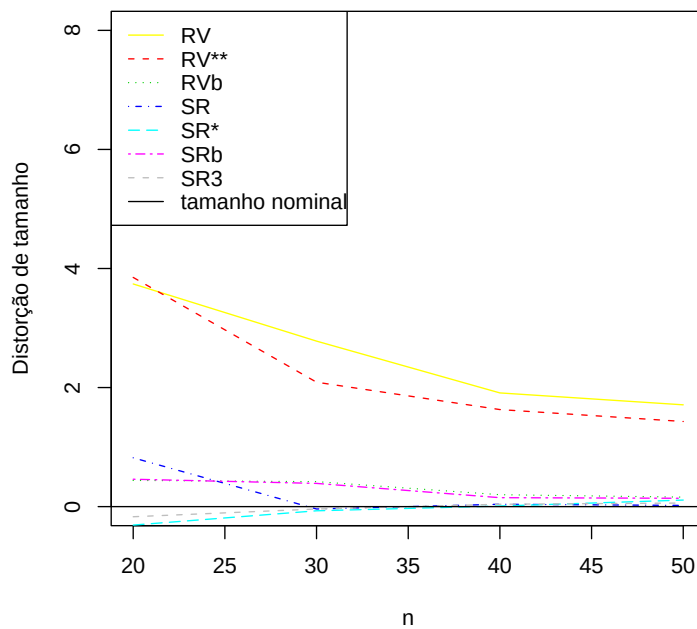


Figura 4.13 Distorção de tamanho dos testes, $p = 2$, Distribuição t -Student 2 graus de liberdade, $\alpha = 5\%$.

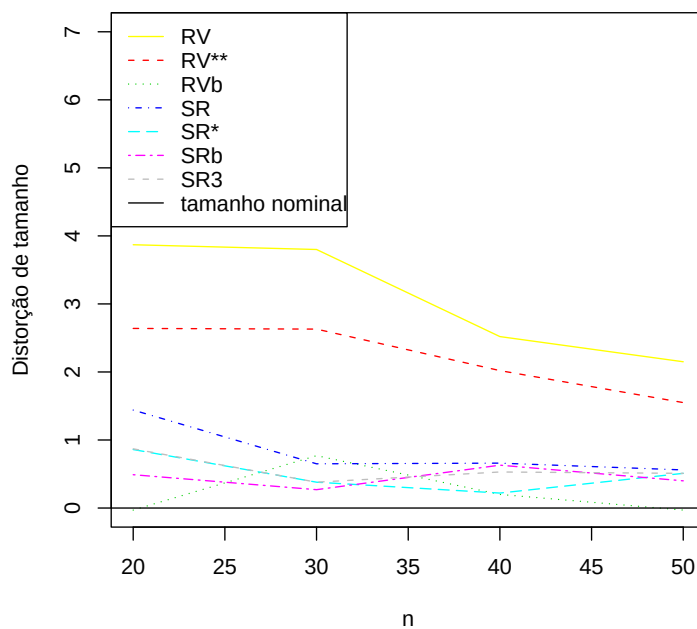


Figura 4.14 Distorção de tamanho dos testes, $p = 2$, Distribuição t -Student 4 graus de liberdade, $\alpha = 10\%$.

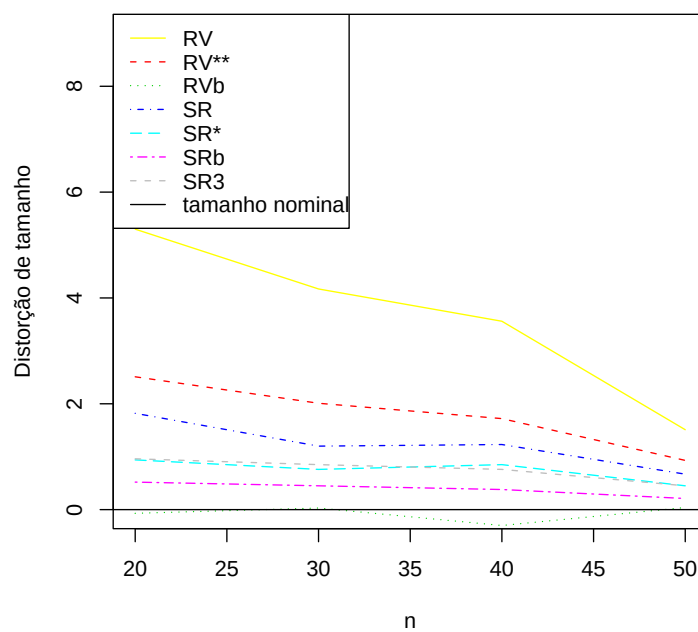
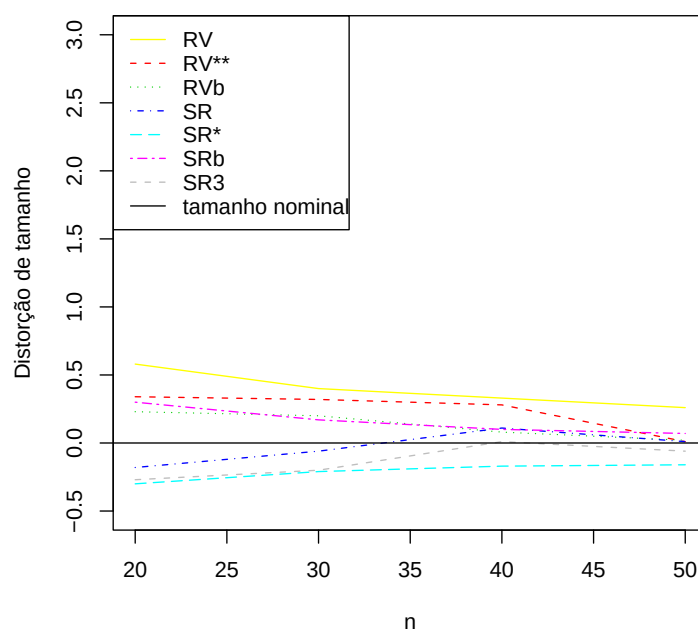


Figura 4.15 Distorção de tamanho dos testes, $p = 2$, Distribuição Logística II, $\alpha = 0.5\%$.



Quadro 4.11. Médias e variâncias das estatísticas, distribuição normal, $p = 2$ e $n = 20$.

momentos	Estatísticas						
	χ^2	RV	RV^*	RV^{**}	SR	SR^*	SR_3^*
média	1,0	1,2	1,1	1,1	1,1	1,1	1,1
variância	2,0	3,1	2,4	2,4	1,9	1,8	1,8

Quadro 4.12 . Médias e variâncias das estatísticas, distribuição logística I, $p = 3$ e $n = 30$.

momentos	Estatísticas						
	χ^2	RV	RV^*	RV^{**}	SR	SR^*	SR_3^*
média	1,0	1,3	1,2	1,2	1,1	1,0	1,0
variância	2,0	3,5	2,9	2,9	2,2	1,9	1,9

Quadro 4.13 . Médias e variâncias das estatísticas, distribuição t -Student, com $\nu = 2$ graus de liberdade, $p = 2$ e $n = 40$.

momentos	Estatísticas						
	χ^2	RV	RV^*	RV^{**}	SR	SR^*	SR_3^*
média	1,0	1,2	1,1	1,1	1,1	1,0	1,0
variância	2,0	3,3	3,1	3,1	2,1	2,0	2,0

Quadro 4.14 . Médias e variâncias das estatísticas, distribuição t -Student, com $\nu = 4$ graus de liberdade, $p = 5$ e $n = 30$.

momentos	estatísticas						
	χ^2	RV	RV^*	RV^{**}	SR	SR^*	SR_3^*
média	1,0	1,5	1,3	1,3	1,2	1,1	1,1
variância	2,0	4,1	3,2	3,1	2,3	2,0	2,0

Quadro 4.15 . Quantis das estatísticas, distribuição normal, $p = 2$ e $n = 40$

estatísticas	quantis (%)			
	90,0	95,0	99,0	99,5
χ^2	2,7	3,8	6,6	7,9
RV	3,1	4,4	7,6	9,2
SR	2,8	3,9	6,4	7,4
RV^*	2,9	4,2	7,2	8,6
RV^{**}	2,9	4,1	7,1	8,6
SR^*	2,9	3,9	6,2	7,1
SR_3^*	2,9	3,9	6,2	7,1

Quadro 4.16 . Quantis das estatísticas, distribuição Cauchy, $p = 2$ e $n = 40$.

estatísticas	quantis (%)			
	90,0	95,0	99,0	99,5
χ^2	2,7	3,8	6,6	7,9
RV	3,6	5,3	9,5	11,2
SR	2,8	4,0	6,7	7,8
RV^*	3,4	5,0	9,1	10,7
RV^{**}	3,4	5,0	9,1	10,6
SR^*	2,8	3,9	6,6	7,7
SR_3^*	2,8	4,0	6,7	7,8

Quadro 4.17 . Quantis das estatísticas, distribuição t -Student com $\nu = 2$ graus de liberdade, $p = 2$ e $n = 40$.

estatísticas	quantis (%)			
	90,0	95,0	99,0	99,5
χ^2	2,7	3,8	6,6	7,9
RV	3,3	4,8	8,7	10,4
SR	2,9	4,1	6,7	8,0
RV^*	3,2	4,6	8,4	10,0
RV^{**}	3,2	4,6	8,4	9,9
SR^*	2,9	4,0	6,6	7,8
SR_3^*	2,9	4,1	6,6	7,8

Quadro 4.18 . Quantis das estatísticas, distribuição t -Studentcom $\nu = 4$ graus de liberdade, $p = 2$ e $n = 40$.

estatísticas	quantis (%)			
	90,0	95,0	99,0	99,5
χ^2	2,7	3,8	6,6	7,9
RV	3,2	4,6	7,7	8,9
SR	2,8	3,9	6,6	7,6
RV^*	3,0	4,4	7,4	8,5
RV^{**}	3,0	4,4	7,4	8,5
SR^*	2,8	3,9	6,4	7,3
SR_3^*	2,8	3,9	6,4	7,3

Quadro 4.19 . Quantis das estatísticas, distribuição Logística II, $p = 2$ e $n = 40$.

estatísticas	quantis (%)			
	90,0	95,0	99,0	99,5
χ^2	2,7	3,8	6,6	7,9
RV	2,9	4,2	7,4	9,0
SR	2,8	3,9	6,6	7,9
RV^*	2,8	4,0	7,0	8,6
RV^{**}	2,8	4,0	7,0	8,6
SR^*	2,7	3,8	6,4	7,6
SR_3^*	2,7	3,8	6,4	7,6

Conclusão

Nesta dissertação, foram apresentadas características e propriedades de algumas distribuições pertencentes à classe de modelos simétricos, em especial modelos não-lineares. Revisamos alguns resultados teóricos acerca da obtenção das correções de Bartlett e tipo-Bartlett.

Tratamos, neste trabalho de refinamentos para testes de hipóteses em modelos de regressão não-lineares simétricos, assumindo que o parâmetro de escala é desconhecido. Apresentamos, em notação matricial, o fator de correção de Bartlett para melhorar a estatística da razão de verossimilhanças, corrigindo aqui a forma apresentada em Gauss (2004), e tipo-Bartlett para melhorar a estatística *escore*, nesta classe de modelos. Apresentamos, também, versões bootstrap dos testes da razão de verossimilhanças e *escore*.

Além disso, analisamos estudos de simulação via Monte Carlo para averiguar o desempenho dos testes ($RV, RV^*, RV^{**}, RV_B, SR, SR^*, SR_3, SR_B$).

As simulações mostraram que:

- (a) Para todas as distribuições, tamanhos de amostras considerados e para o número de componentes do vetor β , as versões corrigidas das estatísticas da razão de verossimilhanças e *escore* apresentaram uma melhora no desempenho em relação às estatísticas usuais.
- (b) A versão bootstrap da razão de verossimilhanças foi a que apresentou taxas de rejeição mais próximas aos níveis nominais considerados, para todas as distribuições, tamanhos de amostras e para o número de componentes do vetor β considerados, do que a estatística da razão de verossimilhanças usual.
- (c) Para as distribuições normal, cauchy, e *t*-Student (com 2 e 4 graus de liberdades), a versão bootstrap *escore* foi a que apresentou melhor desempenho em relação às estatísticas *escore* usual, e versões corrigidas, sendo que estas últimas apresentaram equivalentes. Porém para a distribuição logística II, a estatística SR_3^* apresentou melhores taxas de rejeição.

Assim, nos modelos não-lineares simétricos simulados, para os tamanhos de amostra considerados e para o número de componentes do vetor β , os resultados de simulação mostraram que, em geral, os testes *escore*, sua versão corrigida,

sua versão alternativa e a versão bootstrap tendem a apresentar desempenhos melhores do que os testes da razão de verossimilhanças e suas versões corrigidas. Este é um resultado importante, pois o teste escore apresenta cálculos mais simples, já que só envolve estimação sob o modelo restrito (sob H_0). E que as correções de Bartlett para a estatística da razão de verossimilhanças apresentam uma melhoria em relação ao teste da razão de verossimilhanças.

Para concluir, acrescentamos os estudos de simulação que poderão ser realizados no futuro:

- (i) comparação dos poderes dos testes corrigidos e não corrigidos;
- (ii) análise da influência dos valores das variáveis auxiliares no desempenho dos testes;
- (iii) comparação dos testes corrigidos e não corrigidos na presença de pontos aberrantes.

Apêndice A

Neste apêndice, apresentamos a obtenção de alguns cumulantes conjuntos de derivadas do logaritmo da função de verossimilhança do modelo não-linear simétrico necessários aos cálculos do termo d_β que define a correção de Bartlett para a estatística RV e das quantidades A_1, A_2 e A_3 que definem a correção tipo-Bartlett para as estatísticas SR .

Parâmetro de escala (ϕ) conhecido

O logaritmo da função de verossimilhança total do parâmetro β , dado o vetor de observações $(y_1, \dots, y_n)$, do modelo não-linear simétrico descrito na Seção 2.2, tem a forma

$$\ell(\beta) = -n \log \phi + \sum_{l=1}^n g(z_l),$$

onde $g(z_l) = \log h(z_l^2)$, com $z_l = (y_l - \mu_l)/\phi$ sendo $\mu_l = f(x_l; \beta)$ e $\sum$ denotando somatório sobre os dados. Em problemas regulares, por simples diferenciação, temos

$$\begin{aligned} \frac{\partial \ell(\beta)}{\partial \beta_r} &= \sum_{l=1}^n \frac{\partial g(z_l)}{\partial \beta} = \sum_{l=1}^n \frac{dg(z_l)}{dz_l} \frac{\partial z_l}{\partial \mu_l} \frac{\partial \mu_l}{\partial \beta} = -\frac{1}{\phi} \sum_{l=1}^n g_l^{(1)} f_l^r, \quad r = 1, \dots, p \\ \frac{\partial^2 \ell(\beta)}{\partial \beta_r \partial \beta_s} &= -\frac{1}{\phi} \sum_{l=1}^n \frac{\partial g_l^{(1)}}{\partial \beta_s} f_l^r = \frac{1}{\phi^2} \sum_{l=1}^n g_l^{(2)} f_l^{r,s} - \frac{1}{\phi} \sum_{l=1}^n g_l^{(1)} f_l^{rs}. \end{aligned}$$

De forma análoga, temos que

$$\begin{aligned} \frac{\partial^3 \ell(\beta)}{\partial \beta_r \partial \beta_s \partial \beta_t} &= -\frac{1}{\phi^3} \sum_{l=1}^n g_l^{(3)} f_l^{r,s,t} + \frac{1}{\phi^2} \sum_{l=1}^n g_l^{(2)} (f_l^{r,st} + f_l^{rs,t} + f_l^{rt,s}) - \frac{1}{\phi} \sum_{l=1}^n f_l^{rst} \\ \frac{\partial^4 \ell(\beta)}{\partial \beta_r \partial \beta_s \partial \beta_t \partial \beta_u} &= \frac{1}{\phi^4} \sum_{l=1}^n g_l^{(4)} f_l^{r,s,t,u} - \frac{1}{\phi^3} \sum_{l=1}^n g_l^{(3)} (f_l^{ur,s,t} + f_l^{us,r,t} + f_l^{ut,r,s} \\ &\quad + f_l^{u,tr,s} + f_l^{u,r,st} + f_l^{u,rs,t}) + \frac{1}{\phi^2} \sum_{l=1}^n g_l^{(2)} (f_l^{urt,s} + f_l^{us,rt} \\ &\quad + f_l^{ur,st} + f_l^{ust,r} + f_l^{urs,t} + f_l^{ut,rs} + f_l^{u,trs}) - \frac{1}{\phi} \sum_{l=1}^n f_l^{rstu} \end{aligned}$$

onde $g_l^{(1)}, g_l^{(2)}, g_l^{(3)}, g_l^{(4)}$ denotam as quatro primeiras derivadas de g , isto é $\partial g(z_l)/\partial z_l^r$, $f_l^r = \partial \mu_l / \beta_r$, $f_l^{r,s} = f_l^r f_l^s$, sendo que f_l^r é o elemento da (l, r) matriz $\tilde{X}$.

Considerando a notação $\delta_{(a,b,c,d,e)} = E\{g^{(1)a}g^{(2)b}g^{(3)c}g^{(4)d}z_l^e\}$ e tomando as esperanças encontram-se os cumulantes

$$\begin{aligned}\kappa_r &= -\frac{1}{\phi} \sum_{l=1}^n E(g_l^{(1)}) f_l^r = -\frac{\delta_{(1,0,0,0,0)}}{\phi} \sum_{l=1}^n f_l^r = 0, \\ \kappa_{rs} &= \frac{1}{\phi} \sum_{l=1}^n E(g_l^{(2)}) f_l^{r,s} = \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n f_l^{r,s}\end{aligned}$$

e

$$\kappa_{r,s} = \frac{1}{\phi^2} \left\{ \sum_{l=1}^n E(g_l^{(1)2}) f_l^{r,s} + \sum_{l \neq l'}^n E(g_l^{(1)} g_{l'}^{(1)}) f_l^r f_{l'}^s \right\} = \frac{\delta_{(2,0,0,0,0)}}{\phi^2} \sum_{l=1}^n f_l^{r,s},$$

mas $\kappa_{rs} = -\kappa_{r,s}$, então $\delta_{(0,1,0,0,0)} = -\delta_{(2,0,0,0,0)}$. Também temos que $E[g_l^{(3)}] = 0$ para todas as distribuições simétricas, assim

$$\begin{aligned}\kappa_{rst} &= \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n (f_l^{r,st} + f_l^{rs,t} + f_l^{rt,s}), \\ \kappa_{r,st} &= -\frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n f_l^{r,st}, \\ \kappa_{r,s,t} &= -\frac{2\delta_{(0,0,1,0,0)}}{\phi^3} \sum_{l=1}^n f_l^{r,s,t} = 0.\end{aligned}$$

Derivando os cumulantes acima, temos

$$\begin{aligned}\kappa_{rs}^{(u)} &= \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n (f_l^{r,su} + f_l^{ru,s}), \\ \kappa_{rs}^{(tu)} &= \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n (f_l^{rtu,s} + f_l^{rs,tu} + f_l^{tsu,r} + f_l^{ts,ru}), \\ \kappa_{rst}^{(u)} &= \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n (f_l^{rtu,s} + f_l^{ru,st} + f_l^{rs,tu} + f_l^{rt,su} + f_l^{r,stu} + f_l^{rsu,t}).\end{aligned}$$

Os demais cumulantes são obtidos de forma análoga. Assim,

$$\kappa_{rstu} = \frac{\delta_{(0,0,0,1,0)}}{\phi^4} \sum_{l=1}^n f_l^{r,s,t,u} + \frac{\delta_{(0,1,0,0,0)}}{\phi^2} \sum_{l=1}^n (f_l^{urt,s} + f_l^{us,rt} +$$

$$\begin{aligned}
& f_l^{ur,st} + f_l^{ust,r} + f_l^{urs,t} + f_l^{ut,rs} + f_l^{u,trs}), \\
\kappa_{r,s,t,u} &= \frac{\delta_{(4,0,0,0,0)} - 3\delta_{(2,0,0,0,0)}}{\phi^2} \sum_{l=1}^n f_l^{r,s,t,u}, \\
\kappa_{r,s,tu} &= \frac{\delta_{(2,1,0,0,0)}}{\phi^4} \sum_{l=1}^n f_l^{r,s,t,u} + \frac{\delta_{(0,1,0,0,0)}^2}{\phi^4} \sum_{l=1}^n f_l^{r,s,t,u}. \\
\kappa_{rst\phi} &= \frac{\delta_{(0,0,0,1,1)}}{\phi^4} \sum_{l=1}^n f_l^{r,s,t} - \frac{[2\delta_{(0,1,0,0,0)} + \delta_{(0,0,1,0,1)}]}{\phi^3} \sum_{l=1}^n (f_l^{rs,t} + f_l^{rt,s} + f_l^{r,st}).
\end{aligned}$$

Parâmetro de escala (ϕ) desconhecido

Seja $\ell = \ell(\theta)$ o logaritmo da função de verossimilhança que depende dos parâmetros $\beta = (\beta_1, \dots, \beta_p)$ e ϕ , ambos desconhecidos. Diferenciando $\ell(\theta)$ obtemos

$$\begin{aligned}
\frac{\partial \ell(\theta)}{\partial \phi} &= -\frac{1}{\phi} \sum_{l=1}^n (1 + g_l^{(1)} z_l), \\
\frac{\partial^2 \ell(\theta)}{\partial \phi^2} &= \frac{1}{\phi^2} \sum_{l=1}^n (1 + 2g_l^{(1)} z_l + g_l^{(2)} z_l^2), \\
\frac{\partial^2 \ell(\theta)}{\partial \beta_r \partial \phi} &= \frac{1}{\phi^2} \sum_{l=1}^n (g_l^{(1)} + g_l^{(2)} z_l) f_l^r, \\
\frac{\partial^3 \ell(\theta)}{\partial \phi^3} &= -\frac{1}{\phi^3} \sum_{l=1}^n (2 + 6g_l^{(1)} z_l + 6g_l^{(2)} z_l^2 + g_l^{(3)} z_l^3), \\
\frac{\partial^3 \ell(\theta)}{\partial \beta_r \partial \beta_s \partial \phi} &= -\frac{1}{\phi^3} \sum_{l=1}^n (g_l^{(3)} z_l + 2g_l^{(2)} f_l^{r,s} + \frac{1}{\phi^2} \sum_{l=1}^n (g_l^{(2)} z_l + g_l^{(1)} f_l^{rs}), \\
\frac{\partial^3 \ell(\theta)}{\partial \beta_r \partial \phi^2} &= -\frac{1}{\phi^3} \sum_{l=1}^n (2g_l^{(1)} + 4g_l^{(2)} z_l + g_l^{(3)} z_l^2) f_l^r, \\
\frac{\partial^4 \ell(\theta)}{\partial \beta_r \partial \beta_s \partial \phi^2} &= \frac{1}{\phi^4} \sum_{l=1}^n (6g_l^{(2)} + g_l^{(4)} z_l^2 + 6g_l^{(3)} z_l) f_l^{r,s} \\
&\quad - \frac{1}{\phi^3} \sum_{l=1}^n (g_l^{(3)} z_l^2 + 4g_l^{(2)} z_l + 2g_l^{(1)} f_l^{rs}).
\end{aligned}$$

Tomando esperança, encontramos os cumulantes

$$\begin{aligned}
\kappa_\phi &= 0, \\
\kappa_{\phi\phi} &= \frac{n}{\phi^2} [\delta_{(0,1,0,0,2)} - 1], \\
\kappa_{r\phi} &= 0, \\
\kappa_{\phi\phi\phi} &= -\frac{n}{\phi^3} [6\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)} - 4], \\
\kappa_{rs\phi} &= -\frac{1}{\phi^3} [\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}] \sum_{l=1}^n f_l^{r,s}, \\
\kappa_{r\phi\phi} &= -\frac{1}{\phi^3} [\delta_{(0,0,1,0,2)}] \sum_{l=1}^n f_l^r = 0,
\end{aligned}$$

$$\begin{aligned}
\kappa_{rs\phi\phi} &= \frac{1}{\phi^4} [6\delta_{(0,1,0,0,0)} + 6\delta_{(0,0,1,0,1)} + \delta_{(0,0,0,1,2)}] \sum_{l=1}^n f_l^{r,s} - \frac{1}{\phi^3} [\delta_{(0,0,1,0,2)}] \sum_{l=1}^n f_l^{rs}, \\
\kappa_{\phi\phi}^{(\phi)} &= -\frac{2n}{\phi^3} [\delta_{(0,1,0,0,2)} - 1].
\end{aligned}$$

Como $\kappa_{\phi\phi}^{(\phi)} = \kappa_{\phi\phi\phi} + \kappa_{\phi\phi,\phi}$, temos que

$$\kappa_{\phi\phi,\phi} = \frac{n}{\phi^3} [4\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)} - 2]$$

De $\kappa_{\phi,\phi,\phi} = 2\kappa_{\phi\phi\phi} - 3\kappa_{\phi\phi}^{(\phi)}$, segue que

$$\kappa_{\phi,\phi,\phi} = \frac{2n}{\phi^3} [1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}]$$

Temos também

$$\begin{aligned}
\kappa_{r,s\phi} &= \frac{1}{\phi^3} [\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}] \sum_{l=1}^n f_l^{r,s}, \\
\kappa_{r,\phi\phi} &= \frac{1}{\phi^3} [4\delta_{(0,1,0,0,1)} + \delta_{(0,0,1,0,2)}] \sum_{l=1}^n f_l^r = 0, \\
\kappa_{r\phi,\phi} &= \frac{1}{\phi^3} [2\delta_{(0,1,0,0,1)} + \delta_{(0,0,1,0,2)}] \sum_{l=1}^n f_l^r = 0, \\
\kappa_{rs,\phi} &= \frac{1}{\phi^3} \delta_{(0,0,1,0,1)} \sum_{l=1}^n f_l^{r,s}, \\
\kappa_{r,s,\phi} &= -\frac{2}{\phi^3} [\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)} + \delta_{(2,0,0,0,0)}] \sum_{l=1}^n f_l^{r,s},
\end{aligned}$$

$$\kappa_{rs}^{(\phi)} = -\frac{2}{\phi^3}[\delta_{(0,1,0,0,0)}] \sum_{l=1}^n f_l^{r,s},$$

$$\kappa_{\phi\phi}^{(r)} = 0.$$

Das relações dos cumulantes apresentadas no Capítulo 1, obtemos as seguintes relações entre os δ 's:

$$\begin{aligned}\delta_{(0,1,0,0,1)} &= \delta_{(1,0,0,0,0)} = 0, \\ \delta_{(0,0,0,1,0)} &= -\delta_{(1,0,1,0,0)}, \\ \delta_{(1,1,0,0,1)} + \delta_{(0,0,1,0,1)} &= -\delta_{(0,1,0,0,0)}, \\ \delta_{(0,1,0,0,2)} &= 2 - \delta_{(2,0,0,0,2)}, \\ \delta_{(4,0,0,0,0)} &= -\delta_{(2,1,0,0,0)}, \\ \delta_{(0,1,0,0,0)} &= -\delta_{(2,0,0,0,0)}.\end{aligned}$$

Apêndice B

Neste apêndice apresentamos a obtenção do termo d_β , referente ao teste $H_0 : \beta_1 = \beta_1^{(0)}$ versus $H_1 : \beta_1 \neq \beta_1^{(0)}$ no modelo não-linear simétrico definido no Capítulo 2 considerando o parâmetro de escala ϕ conhecido e desconhecido.

Obtenção de d_β em modelos não-lineares simétricos com ϕ conhecido

De acordo com a expressão (3.2), temos que

$$\lambda_{rstu} = \kappa^{rs} \kappa^{tu} \left(\frac{1}{4} \kappa_{rstu} - \kappa_{rst}^{(u)} + \kappa_{rt}^{(su)} \right)$$

Substituindo os valores de κ 's, encontrados para os modelos não-lineares simétricos, no apêndice A, temos

$$\begin{aligned} \lambda_{rstu} &= \frac{1}{\phi^2} \sum' \kappa^{rs} \kappa^{tu} \left\{ \frac{\delta_{(0,0,0,1,0)}}{4\phi^2} \sum_{l=1}^n f_l^{r,s,t,u} - \delta_{(0,1,0,0,0)} \sum_{l=1}^n (f_l^{rsu,t} \right. \\ &\quad \left. + f_l^{rt,su}) + \frac{\delta_{(0,1,0,0,0)}}{4} \left\{ 4 \sum_{l=1}^n f_l^{rts,u} + 2 \sum_{l=1}^n (f_l^{us,rt} + f_l^{rs,tu}) \right\} \right\} \\ &= \frac{1}{\phi^2} \sum' \kappa^{rs} \kappa^{tu} \left\{ \frac{\delta_{(0,0,0,1,0)}}{4\phi^2} \sum_{l=1}^n f_l^{r,s,tu} + \frac{\delta_{(0,1,0,0,0)}}{4} \times \right. \\ &\quad \left. \sum_{l=1}^n f_l^{rs,tu} - \frac{\delta_{(0,1,0,0,0)}}{2} \sum_{l=1}^n f_l^{us,rt} \right\}. \end{aligned}$$

Invertendo os somatórios e rearranjando os termos, temos

$$\begin{aligned} \lambda_{rstu} &= \frac{1}{\phi^2} \left\{ \frac{\delta_{(0,0,0,1,0)}}{4\phi^2} \sum_{l=1}^n \left(\sum' f_l^r \kappa^{rs} f_l^s \right) \left(\sum' f_l^t \kappa_{tu} f_l^u \right) \right. \\ &\quad \left. + \frac{\delta_{(0,1,0,0,0)}}{4} \sum_{l=1}^n \left(\sum' f_l^{rs} \kappa^{rs} \right) \left(\sum' f_l^{tu} \kappa_{tu} \right) - \frac{\delta_{(0,1,0,0,0)}}{2} \times \right. \\ &\quad \left. \sum_{l=1}^n \left(\sum' f_l^{rt} \kappa_{tu} f_l^{us} \kappa^{sr} \right) \right\}. \end{aligned}$$

Das definições das matrizes Z , B e D , definidas na seção 3.1, obtemos a expressão

$$\lambda_{rstu} = \frac{\delta_{(0,0,0,1,0)}}{4m^2} \sum_{l=1}^n z_{ll}^2 + \frac{\delta_{(0,1,0,0,0)}}{4m^2} \phi^2 \sum_{l=1}^n d_{ll}^2 - \frac{1}{2m^2} \phi^2 \sum_{l=1}^n b_{ll}.$$

que em notação matricial tem a forma

$$\sum' \lambda_{rstu} = \frac{\delta_{(0,0,0,1,0)}}{4m^2} \text{tr}(Z_d^2) - \frac{\phi^2}{4m^2} \delta_{(0,1,0,0,0)} \text{tr}(2B - D^2).$$

Mostraremos agora a expressão de λ_{rstuvw} . Assim temos a expressão

$$\begin{aligned} \sum' \lambda_{rstuvw} = & \kappa^{rs} \kappa^{tu} \kappa^{vw} \left\{ \kappa_{rtv} \left(\frac{1}{6} \kappa_{suw} - \kappa_{sw}^{(u)} \right) \right. \\ & \left. + \kappa_{rtu} \left(\frac{1}{4} \kappa_{svw} - \kappa_{sw}^{(v)} \right) + \kappa_{rt}^{(v)} \kappa_{sw}^{(u)} + \kappa_{rt}^{(u)} + \kappa_{sw}^{(v)} \right\}. \end{aligned}$$

Então, substituindo os κ 's temos

$$\begin{aligned} \sum' \lambda_{rstuvw} = & \sum' \kappa^{rs} \kappa^{tu} \kappa^{vw} \left\{ \frac{1}{6} \delta_{(0,1,0,0,0)}^2 \sum_{l=1}^n (f_l^{rv,t} + f_l^{r,tv} + f_l^{rt,v}) \right. \\ & \sum_{m=1}^n (f_m^{sw,u} + f_m^{s,uw} + f_m^{su,w}) - \delta_{(0,1,0,0,0)}^2 \sum_{l=1}^n (f_l^{rv,t} + f_l^{r,w} \\ & + f_l^{rt,v}) \sum_{m=1}^n (f_m^{su,w} + f_m^{s,wu}) + \frac{\delta_{(0,1,0,0,0)}^2}{4} \sum_{l=1}^n (f_l^{ru,t} + f_l^{r,tu} \\ & + f_l^{rt,u}) \sum_{m=1}^n (f_m^{sv,w} + f_m^{s,vw} + f_m^{sv,w}) - \delta_{(0,1,0,0,0)}^2 \sum_{l=1}^n (f_l^{ru,t} \\ & + f_l^{r,tu} + f_l^{rt,u}) \sum_{m=1}^n (f_m^{sv,w} + f_m^{s,wv}) - \delta_{(0,1,0,0,0)}^2 \sum_{l=1}^n (f_l^{rv,t} \\ & + f_l^{r,tv}) \sum_{m=1}^n (f_m^{su,w} + f_m^{s,wu}) + \delta_{(0,1,0,0,0)}^2 \\ & \left. \sum_{l=1}^n (f_l^{ru,t} + f_l^{r,tu}) \sum_{m=1}^n (f_l^{sv,w} + f_l^{s,wv}) \right\} \end{aligned}$$

$$\begin{aligned} \sum' \lambda_{rstuvw} = & \sum' \kappa^{rs} \kappa^{tu} \kappa^{vw} \delta_{(0,1,0,0,0)}^2 \sum_{l,m} \left\{ \frac{7}{6} f_l^{rv,t} f_m^{sw,u} - \frac{5}{6} f_l^{rv,t} f_m^{s,uw} \right. \\ & + \frac{1}{6} f_l^{rv,t} f_m^{su,w} + \frac{7}{6} f_l^{r,tv} f_m^{sw,u} - \frac{5}{6} f_l^{r,tv} f_m^{s,uw} + \frac{1}{6} f_l^{r,tv} f_m^{su,w} \\ & + \frac{1}{6} f_l^{rt,v} f_m^{sw,u} - \frac{5}{6} f_l^{rt,v} f_m^{s,uw} - \frac{5}{6} f_l^{rt,v} f_m^{su,w} + \frac{5}{4} f_l^{ru,t} f_m^{sv,w} \\ & \left. - \frac{3}{4} f_l^{ru,t} f_m^{s,vw} + \frac{1}{4} f_l^{ru,t} f_m^{sv,w} + \frac{5}{4} f_l^{r,tu} f_m^{sv,w} - \frac{3}{4} f_l^{r,tu} f_m^{s,vw} \right\}. \end{aligned}$$

Assim a expressão de λ_{rstuvw} trocando a ordem dos somatórios é

$$\begin{aligned}\lambda_{rstuvw} = & \frac{1}{4} \sum_{l,m} f_l^{rt,tu} f_m^{sv,w} + \frac{1}{4} \sum_{l,m} f_l^{rt,u} f_m^{sw,v} - \frac{3}{4} \sum_{l,m} f_l^{rt,u} f_m^{s,vw} \\ & - \frac{3}{4} \sum_{l,m} f_l^{rt,u} f_m^{sv,w} \} - \delta_{0,1,0,0,0}^2 \sum_{l,m} \left\{ \frac{1}{2} \left\{ \sum' \kappa^{tu} f_l^t f_m^u \right\} \right. \\ & \left\{ \sum' \kappa^{sr} f_l^{rv} \kappa^{vw} f_m^{ws} \right\} - \frac{1}{4} \left\{ \sum' \kappa^{tu} f_l^{tu} \right\} \left\{ \sum' \kappa^{wv} f_m^{wv} \right\} \\ & \left. \left\{ \sum' \kappa^{rs} f_l^r f_m^s \right\} \right\}\end{aligned}$$

Das definições das matrizes temos que

$$\begin{aligned}\sum' \lambda_{rstuvw} &= \sum_{l,m} \frac{\delta_{(0,1,0,0,0)}^2}{m^3} \left\{ \frac{1}{2} z_{lm} b_{lm} - \frac{1}{4} d_l d_m z_{lm} \right\} \\ &= - \frac{\delta_{(0,1,0,0,0)}^2}{2m^3} \left\{ \frac{1}{2} 1^\top D Z D 1 - \text{tr}(ZB) \right\}\end{aligned}$$

Obtenção de $d_{\beta\phi}$ em modelos não-lineares simétricos com ϕ desconhecido

A quantidade $d_{\beta\phi}$ definida na Seção (3.2) tem a forma

$$d_{\beta\phi} = q^{-1} \left(\sum'_{\beta,\phi} (\lambda_{rstu} - \lambda_{rstuvw}) - \sum''_{\beta,\phi} (\lambda_{rstu} - \lambda_{rstuvw}) \right) \quad (\text{B.1})$$

onde $\sum'_{\beta,\phi}$ representa o somatório sobre todas as combinações dos parâmetros β e ϕ , mas com pelo menos um índice igual a ϕ e $\sum''_{\beta,\phi}$ representa o somatório sobre todas as combinações de parâmetros $\{\beta_{q+1}, \dots, \beta_p; \phi\}$ com pelo menos um índice igual a ϕ . As expressões de λ_{rstu} e λ_{rstuvw} são definidos em (3.2). Devido à ortogonalidade global de β e ϕ e substituindo os cumulantes deste Apêndice na expressão (B.1), temos

$$\begin{aligned}\sum'_{\beta,\phi} \lambda_{rstu} &= \sum' \left\{ \kappa^{\phi\phi} \left[\frac{1}{4} \kappa^{tu} \kappa_{\phi\phi tu} + \kappa^{rs} \left(\frac{1}{4} \kappa_{rs\phi\phi} - \kappa_{rs\phi}^{(\phi)} \right) \right] \right\} \\ &+ \kappa^{\phi\phi^2} \left(\frac{1}{4} \kappa_{\phi\phi\phi\phi} - \kappa_{\phi\phi\phi}^{(\phi)} + \kappa_{\phi\phi}^{(\phi\phi)} \right),\end{aligned} \quad (\text{B.2})$$

$$\begin{aligned}
\sum_{\beta, \phi}' \lambda_{rstuvw} = & \sum' \left\{ \frac{1}{4} \kappa^{\phi\phi^2} \kappa^{vw} \kappa_{\phi\phi\phi} \kappa_{\phi vw} \right. \\
& + \kappa^{\phi\phi^2} \kappa^{tu} \left(\frac{1}{4} \kappa_{\phi tu} \kappa_{\phi\phi\phi} - \kappa_{\phi tu} \kappa_{\phi\phi}^{(\phi)} \right) \\
& + \kappa^{\phi\phi} \kappa^{tu} \kappa^{vw} \left(\frac{1}{6} \kappa_{\phi tv} \kappa_{\phi uw} + \frac{1}{4} \kappa_{\phi tu} \kappa_{\phi vw} \right) \\
& + \kappa^{\phi\phi} \kappa^{rs} \kappa^{vw} \left(\frac{1}{6} \kappa_{r\phi v} \kappa_{s\phi w} - \kappa_{r\phi v} \kappa_{sw}^{(\phi)} \right) \\
& + \frac{1}{6} \kappa^{\phi\phi} \kappa^{rs} \kappa^{tu} \kappa_{rt\phi} \kappa_{su\phi} \left. \right\} \\
& + \kappa^{\phi\phi^3} \left\{ \frac{1}{6} \kappa_{\phi\phi\phi}^2 + \frac{1}{4} \kappa_{\phi\phi\phi}^2 - 2 \kappa_{\phi\phi\phi} \kappa_{\phi\phi}^{(\phi)} + 2 \kappa_{\phi\phi}^{(\phi^2)} \right\} \quad (B.3)
\end{aligned}$$

onde $\sum'$ é o somatório sobre todos os componentes de β . De forma similar, $\sum''_{\beta, \phi} (\lambda_{rstu} - \lambda_{rstuvw})$ vem de (B.2) e (B.3), mas com $\sum''$ representando o somatório sobre $\{\beta_{q+1}, \dots, \beta_p\}$. Consideramos inicialmente a obtenção de $\sum' \lambda_{rstu}$. Assim, substituindo os cumulantes e rearranjando os somatórios, temos

$$\begin{aligned}
\sum' \lambda_{rstu} = & \ell_{\phi\phi\phi\phi} + \frac{\phi^2}{n(\delta_{(0,1,0,0,2)} - 1)} \sum_{l=1}^n \left\{ \frac{1}{4\phi^4} (\delta_{(0,0,0,1,2)} - 6\delta_{(1,1,0,0,1)}) \right. \\
& \sum' \kappa^{tu} f_l^{t,u} + \frac{1}{4\phi^2} (\delta_{(0,0,0,1,2)} - 6\delta_{(1,1,0,0,1)}) \sum' \kappa^{rs} f_l^{r,s} \\
& \left. - \frac{3}{\phi^4} (\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}) \sum' \kappa^{rs} f_l^{r,s} \right\}.
\end{aligned}$$

Como $-\sum' \kappa^{rs} f_l^{r,s}$ é o elemento (l, l) da matriz $Z\phi^2/\delta_{(2,0,0,0,0)}$, temos

$$\begin{aligned}
\sum' \lambda_{rstu} = & \ell_{\phi\phi\phi\phi} + \frac{(\delta_{(0,0,0,1,2)} - 6\delta_{(1,1,0,0,1)})}{2n\phi^2(\delta_{(0,1,0,0,2)} - 1)} \\
& \sum_{l=1}^n \left(-\frac{\phi^2}{\delta_{(2,0,0,0,0)}} z_{ll} \right) - \frac{3(\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)})}{n\phi^2(\delta_{(0,1,0,0,2)} - 1)} \times \\
& \sum_{l=1}^n \left(-\frac{\phi^2}{\delta_{(2,0,0,0,0)}} z_{ll} \right).
\end{aligned}$$

Escrevendo em notação matricial, vem

$$\sum' \lambda_{rstu} = \ell_{\phi\phi\phi\phi} - \frac{m_4}{2nm_1} \text{tr}(Z) + 3 \frac{m_3}{nm_1} \text{tr}(Z)$$

onde

$$\begin{aligned}
m_1 &= \delta_{(0,1,0,0,2)} - 1, \\
m_2 &= 4 - \delta_{(0,0,1,0,3)} - 6\delta_{(0,1,0,0,2)} \\
m_3 &= (\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)})/\delta_{(2,0,0,0,0)}, \\
\text{e } m_4 &= (\delta_{(0,0,0,1,2)} - 6\delta_{(1,1,0,0,1)})/\delta_{(2,0,0,0,0)}.
\end{aligned}$$

Como $\text{tr}(Z) = \text{tr}(\tilde{X}(\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top) = \text{tr}(I_p) = p$, sendo p o posto da matriz $\tilde{X}$ e I_p a matriz identidade $p \times p$. Assim, temos

$$\sum' \lambda_{rstu} = \ell_{\phi\phi\phi\phi} + \left(\frac{6m_3 - m_4}{2nm_1} \right) p.$$

De forma análoga,

$$\begin{aligned}
\sum'' \lambda_{rstu} &= \ell_{\phi\phi\phi\phi} + \left(\frac{6m_3 - m_4}{2nm_1} \right) \text{tr}(Z_d - Z_{2d}) \\
&= \ell_{\phi\phi\phi\phi} + \left(\frac{6m_3 - m_4}{2nm_1} \right) \text{tr}(p - q)
\end{aligned}$$

Para obter $\sum' \lambda_{rstuvw}$ temos, de (B.3), que

$$\begin{aligned}
\sum_{\beta, \phi}' \lambda_{rstuvw} &= \ell_{\phi\phi\phi\phi\phi\phi} \\
&+ \frac{1}{4} \kappa_{\phi\phi\phi} \kappa^{\phi\phi^2} \left(-\frac{1}{\phi^3} \right) (\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}) \sum_{l=1}^n (\sum' \kappa^{vw} f_l^{v,w}) \\
&+ \frac{1}{4} \kappa_{\phi\phi\phi} \kappa^{\phi\phi^2} \left(-\frac{1}{\phi^3} \right) (\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}) \sum_{l=1}^n (\sum' \kappa^{tu} f_l^{t,u}) \\
&- \kappa^{\phi\phi^2} \kappa_{\phi\phi}^{(\phi)} \left(-\frac{1}{\phi^3} \right) (\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}) \sum_{l=1}^n (\sum' \kappa^{tu} f_l^{t,u}) \\
&+ \frac{1}{6} \kappa^{\phi\phi} \left(-\frac{1}{\phi^3} \right)^2 (\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)})^2 \sum_{l,m} (\sum' \kappa^{tu} f_l^t f_m^u) \times \\
&(\sum' \kappa^{vw} f_l^v f_m^w)
\end{aligned}$$

$$\begin{aligned}
& + \frac{1}{4} \kappa^{\phi\phi} \left(-\frac{1}{\phi^3} \right)^2 (\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)})^2 \sum_{l,m} (\sum' \kappa^{tu} f_l^{t,u}) \times \\
& (\sum' \kappa^{vw} f_m^{v,w}) \\
& + \frac{1}{6} \kappa^{\phi\phi} \left(-\frac{1}{\phi^3} \right)^2 (\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)})^2 \sum_{l,m} (\sum' \kappa^{rs} f_l^r f_m^s) \times \\
& (\sum' \kappa^{vw} f_l^v f_m^w) \\
& - \kappa^{\phi\phi} \left(-\frac{1}{\phi^3} \right) (\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}) \times \\
& \left(\frac{2}{\phi^3} \right) \delta_{(2,0,0,0,0)} \sum_{l,m} (\sum' \kappa^{rs} f_l^r f_m^s) (\sum' \kappa^{vw} f_l^v f_m^w) \\
& + \frac{1}{6} \kappa^{\phi\phi} \left(-\frac{1}{\phi^3} \right)^2 (\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}) \sum_{l,m} (\sum' \kappa^{rs} f_l^r f_m^s) \times \\
& (\sum' \kappa^{tu} f_l^t f_m^u)
\end{aligned}$$

Como

$$\begin{aligned}
\kappa^{\phi\phi} &= \kappa_{\phi\phi}^{-1} = \frac{\phi^2}{n} (\delta_{(0,1,0,0,2)} - 1)^{-1}, \\
\kappa_{\phi\phi}^{(\phi)} &= -\frac{2}{\phi} \kappa_{\phi\phi}, \\
\kappa_{\phi\phi\phi} &= -\frac{n}{\phi^3} (6\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)} - 4)
\end{aligned}$$

e $-\sum \kappa^{rs}(r,s)_l$ é o elemento (l,l) da matriz $Z\phi^2/\delta_{(2,0,0,0,0)}$, temos

$$\begin{aligned}
\sum_{\beta,\phi} \lambda_{rstuvw} &= \ell_{\phi\phi\phi\phi\phi\phi} \\
& - \frac{1}{2} \frac{(6\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)} - 4)}{n(\delta_{(0,1,0,0,2)} - 1)^2} \left(\frac{\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}}{\delta_{(2,0,0,0,0)}} \right) \sum_{l=1}^n z_{ll} \\
& + \frac{2}{n} \left(\frac{1}{(\delta_{(0,1,0,0,2)} - 1)} \right) \left(\frac{\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}}{\delta_{(2,0,0,0,0)}} \right) \sum_{l=1}^n z_{ll} \\
& + \frac{1}{2n} \left(\frac{1}{(\delta_{(0,1,0,0,2)} - 1)} \right) \left(\frac{\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}}{\delta_{(2,0,0,0,0)}} \right)^2 \sum_{l,m} z_{lm} z_{ml} \\
& + \frac{1}{4n} \left(\frac{1}{(\delta_{(0,1,0,0,2)} - 1)} \right) \left(\frac{\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}}{\delta_{(2,0,0,0,0)}} \right)^2 \sum_{l=1}^n z_{ll}^2 \\
& + \frac{2}{n} \left(\frac{1}{(\delta_{(0,1,0,0,2)} - 1)} \right) \left(\frac{\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}}{\delta_{(2,0,0,0,0)}} \right)^2 \sum_{l,m} z_{lm} z_{ml}.
\end{aligned}$$

Escrevendo de forma matricial vem

$$\begin{aligned}\sum_{\beta,\phi}'\lambda_{rstuvw} = & \ell_{\phi\phi\phi\phi\phi\phi} + \frac{1}{2n} \frac{m_2 m_3}{m_1^2} \text{tr}(Z_d) + \frac{2}{n} \frac{m_3}{m_1} \text{tr}(Z_d) + \frac{1}{2n} \frac{m_3^2}{m_1} \text{tr}(ZZ) \\ & + \frac{1}{4n} \frac{m_3^2}{m_1} \text{tr}(Z)^2 + \frac{2}{n} \frac{m_3}{m_1} \text{tr}(ZZ).\end{aligned}$$

Mas

$$\begin{aligned}\text{tr}(Z) &= \text{tr}(\tilde{X}(\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top) = \text{tr}(I_p) = p \\ \text{e } \text{tr}(ZZ) &= \text{tr}(\tilde{X}(\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top \tilde{X}(\tilde{X}^\top \tilde{X})^{-1} \tilde{X}^\top) = \text{tr}(I_p) = p,\end{aligned}$$

sendo p o posto da matriz $\tilde{X}$. Assim, temos

$$\sum_{\beta,\phi}'\lambda_{rstuvw} = \ell_{\phi\phi\phi\phi\phi\phi} + \frac{1}{2n} \frac{m_2 m_3}{m_1^2} p + \frac{4}{n} \frac{m_3}{m_1} p + \frac{1}{2n} \frac{m_3^2}{m_1} p + \frac{1}{2} \frac{m_3^2}{m_1} \frac{p^2}{2n}. \quad (1)$$

De forma análoga,

$$\begin{aligned}\sum_{\beta,\phi}''\lambda_{rstuvw} = & \ell_{\phi\phi\phi\phi\phi\phi} + \frac{1}{2n} \frac{m_2 m_3}{m_1^2} (p - q) \\ & + \frac{4}{n} \frac{m_3}{m_1} (p - q) + \frac{1}{2n} \frac{m_3^2}{m_1} (p - q) + \frac{1}{2n} \frac{m_3^2}{m_1} \frac{(p - q)^2}{2}.\end{aligned} \quad (2)$$

Assim, substituindo (1) e (2) em (B.1), temos

$$d_{\beta\phi} = \frac{1}{n} \left\{ d_1 + \frac{(2p - q)}{2} d_2 \right\},$$

onde

$$\begin{aligned}d_1 &= -\frac{m_3}{2m_3^2} (m_2 + m_1(2 + m_3)) - \frac{m_4}{2m_1} \\ \text{e } d_2 &= -\frac{m_3^2}{2m_1}.\end{aligned}$$

Apêndice C

Neste apêndice apresentamos a obtenção das quantidades $A_{1,\beta}$, $A_{2,\beta}$ e $A_{3,\beta}$ definidas em (3.8) e (3.10) referente ao teste $H_0 : \beta_1 = \beta_1^{(0)}$ versus $H_1 : \beta_1 \neq \beta_1^{(0)}$ para modelos não lineares simétricos definido no Capítulo 2 considerando o parâmetro de escala ϕ conhecido e desconhecido.

Obtenção de $A_{1,\beta}$ e $A_{2,\beta}$ em modelos não-lineares com ϕ conhecido

Para a obtenção destas quantidades basta substituir os cumulantes correspondentes aos modelos não-lineares simétricos. Chamaremos, aqui, os termos A_1 e A_2 de $A_{1,\beta}$ e $A_{2,\beta}$, respectivamente. Consideremos, inicialmente, a obtenção de $A_{1,\beta}$, de acordo com a expressão (3.8).

$$\begin{aligned} A_{1,\beta} = & 3 \sum' \left\{ \frac{1}{\phi^4} \delta_{(0,1,0,0,0)}^2 \sum_{l=1}^n (f_l^{ik,j} + f_l^{ij,k} - f_l^{i,jk}) \sum_{m=1}^n (f_m^{ts,r} + f_m^{s,rt} - f_m^{t,rs}) \right\} a_{ij} a_{st} m_{kr} \\ & - 6 \sum' \left\{ \frac{\delta_{(0,1,0,0,0)}^2}{\phi^4} \sum_{l,m} f_l^{i,jk} (f_m^{ts,r} + f_m^{s,rt} - f_m^{t,rs}) \right\} a_{js} a_{kt} m_{ir} \\ & + 12 \sum' \left\{ \frac{\delta_{(2,1,0,0,0)} + \delta_{(2,0,0,0,0)}^2}{\phi^4} \sum_{l=1}^n f_l^{i,j,k,r} \right\} a_{kr} m_{ij}. \end{aligned}$$

Sabemos que os termos $\sum' f_l^{i,j} a_{ij}$, $\sum' f_m^{s,t} m_{st}$ e $\sum' f_l^{ij} a_{ij}$, são os (l, l) -ésimos elementos das matrizes $\phi^2 Z_2 / \delta_{(2,0,0,0,0)}$, $\phi^2 (Z - Z_2) / \delta_{(2,0,0,0,0)}$ e $\phi^2 Q_2 / \delta_{(2,0,0,0,0)}$, respectivamente, o termos $\sum' f_l^{jk} a_{js} f_l^{st} a_{tk}$ corresponde ao (l, m) -ésimos elementos da matriz $\phi^2 J_2 / \delta_{(2,0,0,0,0)}$. Assim, invertendo os somatórios e ajustando os elementos das matrizes acima, temos

$$\begin{aligned} A_{1,\beta} = & 3 \frac{\delta_{(0,1,0,0,0)}^2}{\phi^4} \sum_{l,m} \left\{ \left(\frac{\phi^2}{\delta_{(2,0,0,0,0)}} q_l \right) \left(\frac{\phi^2}{\delta_{(2,0,0,0,0)}} (z_{lm} - z_{2lm}) \right) \left(\frac{\phi^2}{\delta_{(2,0,0,0,0)}} q_m \right) \right\} \\ & - 6 \frac{\delta_{(0,1,0,0,0)}^2}{\phi^4} \sum_{l,m} \left\{ \left(\frac{\phi^2}{\delta_{(2,0,0,0,0)}} (z_{lm} - z_{2lm}) \right) \left(\frac{\phi^2}{\delta_{(2,0,0,0,0)}} j_{lm} \right) \right\} \\ & + 12 \frac{\delta_{(2,1,0,0,0)} + \delta_{(2,0,0,0,0)}^2}{\phi^4} \sum_{l,m} \left\{ \left(\frac{\phi^2}{\delta_{(2,0,0,0,0)}} (z_{lm} - z_{2lm}) \right) \left(\frac{\phi^2}{\delta_{(2,0,0,0,0)}} z_{2ll} \right) \right\}. \end{aligned}$$

Onde a forma matricial é

$$\begin{aligned}
A_{1,\beta} = & 3 \frac{\delta_{(0,1,0,0,0)}^2}{\delta_{(2,0,0,0,0)}^3} \phi^2 \left\{ \iota^\top Q_2 (Z - Z_2) Q_2 \iota \right\} \\
& - 6 \frac{\delta_{(0,1,0,0,0)}^2}{\delta_{(2,0,0,0,0)}^3} \phi^2 \left\{ \iota^\top (Z - Z_2) \otimes J_2 \iota \right\} \\
& + 12 \frac{\delta_{(2,1,0,0,0)} + \delta_{(2,0,0,0,0)}^2}{\delta_{(2,0,0,0,0)}^2} \text{tr}((Z - Z_2)_d Z_{2d}).
\end{aligned}$$

Considere agora o termo de $A_{2,\beta}$ definido em (3.9). Como $\kappa_{r,s,t} = 0$, temos

$$A_{2,\beta} = 3 \sum' \kappa_{i,j,k,r} m_{ij} m_{kr}.$$

Substituindo o valor do cumulante, temos

$$A_{2,\beta} = 3 \sum' \frac{[\delta_{(4,0,0,0,0)} - 3\delta_{(2,0,0,0,0)}^2]}{\phi^4} \sum_{l=1}^n f_l^{i,j,k,r} m_{ij} m_{kr}.$$

Invertendo os somatórios, e substituindo os termos $\sum' f^i m_{ij} f^j$ por sua respectiva matriz, temos

$$\begin{aligned}
A_{2,\beta} = & 3 \frac{[\delta_{(4,0,0,0,0)} - 3\delta_{(2,0,0,0,0)}^2]}{\phi^4} \sum_{l=1}^n (\sum' f_l^i m_{ij} f_{lj}) (\sum' f_l^k m_{kr} f^r) \\
= & 3 \frac{[\delta_{(4,0,0,0,0)} - 3\delta_{(2,0,0,0,0)}^2]}{\phi^4} \sum_{l=1}^n (\phi^2 \delta_{(2,0,0,0,0)}^{-1} (z_{ll} - z_{2l})) (\phi^2 \delta_{(2,0,0,0,0)}^{-1} (z_{ll} - z_{2l})).
\end{aligned}$$

Matricialmente, temos

$$A_{2,\beta} = 3 \frac{[\delta_{(4,0,0,0,0)} - 3\delta_{(2,0,0,0,0)}^2]}{\delta_{(2,0,0,0,0)}^2 \phi^4} \left\{ \iota^\top (Z - Z_2)_d^{(2)} \iota \right\}.$$

Finalmente, substituindo os cumulantes em (3.10) teremos $A_{3,\beta} = 0$.

Obtenção de $A_{1,\beta\phi}$ e $A_{2,\beta\phi}$ em modelos não-lineares com ϕ desconhecido

As quantidades $A_{1,\beta\phi}$ e $A_{2,\beta\phi}$ são agora obtidas substituindo os cumulantes nas expressões (3.11). Consideremos, inicialmente, a obtenção de $A_{1,\beta\phi}$.

$$\begin{aligned}
A_{1,\beta\phi} = & -6 \sum' \left\{ -\frac{1}{\phi^3} (\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)}) \left(\sum_{l=1}^n f_l^{i,j} a_{ij} \right) \right. \\
& \left. \left(\frac{2}{\phi^3} \right) \delta_{(1,1,0,0,1)} \sum_{l=1}^n f_l^{s,t} m_{st} \right\} a_{\phi\phi} \\
& -6 \sum' \left\{ \frac{n}{\phi^3} (2\delta_{(0,1,0,0,2)} + \delta_{(0,1,0,0,3)}) \frac{2}{\phi^3} \delta_{(1,1,0,0,1)} \sum_{l=1}^n f_l^{s,t} m_{st} \right\} a_{\phi\phi}^2 \\
& +6 \sum' \left\{ -\frac{1}{\phi^3} (3\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)}) \left(\sum_{l=1}^n f_l^{i,k} \right) \right. \\
& \left. \left\{ -\frac{1}{\phi^3} (\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)}) \sum_{m=1}^n f_m^{t,r} \right\} a_{\phi\phi} a_{kt} m_{rv} \right. \\
& +6 \sum' \left\{ -\frac{1}{\phi^3} (3\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)}) \left(\sum_{l=1}^n f_l^{i,j} \right) \right. \\
& \left. \left\{ -\frac{1}{\phi^3} (\delta_{(2,0,1,0,0,0)} - \delta_{(0,0,1,0,1)}) \sum_{m=1}^n f_m^{r,s} \right\} a_{\phi\phi} a_{js} m_{rv} \right. \\
& +6 \sum' \left\{ \frac{1}{\phi^4} (4\delta_{(3,0,0,0,1)} + \delta_{(4,0,0,0,2)} + \delta_{(2,1,0,0,2)} \right. \\
& \left. \left. - 2\delta_{(0,1,0,0,0)}) \left(\sum_{l=1}^n f_l^{i,j} m_{ij} \right) \right\} a_{\phi\phi}
\end{aligned}$$

onde $\sum'$ representa o somatório sobre todos os β 's e a_{ij} e m_{ij} são os (i,j) -ésimos elementos das matrizes A e M , respectivamente.

Invertendo a ordem dos somatórios e rearranjando os termos temos,

$$\begin{aligned}
A_{1,\beta\phi} = & \frac{12\phi^{-4}}{n(\delta_{(2,0,0,0,2)} - 1)} \{ (\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)}) \delta_{(1,1,0,0,1)} \} \\
& \sum_{l=1}^n \left(\sum' f_l^{i,j} a_{ij} \right) \sum_{m=1}^n \left(\sum' f_m^{s,t} m_{st} \right)
\end{aligned}$$

$$\begin{aligned}
& + \frac{12\phi^{-2}}{n(\delta_{(2,0,0,0,2)} - 1)^2} \{ \delta_{(1,1,0,0,1)}(2\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)}) \sum_{l=1}^n (\sum' f_l^{s,r} m_{sr}) \} a_{\phi\phi} \\
& + \frac{6\phi^{-4}}{n(\delta_{(2,0,0,0,2)} - 1)} \{ (3\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)})(\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)}) \} \\
& \sum_{l,m} (\sum' f_{lm}^{i,r} m_{ir}) (\sum' f_{lm}^{k,t} a_{kt}) \\
& \frac{6\phi^{-4}}{n(\delta_{(2,0,0,0,2)} - 1)} \{ (3\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)})(\delta_{(0,0,1,0,1)} - 2\delta_{(0,1,0,0,0)}) \} \\
& \sum_{l,m} (\sum' f_{lm}^{i,r} m_{ir}) (\sum' f_{lm}^{j,s} a_{js}) \\
& - \frac{6\phi^{-4}}{n(\delta_{(2,0,0,0,2)} - 1)} (4\delta_{(3,0,0,0,1)} + \delta_{(4,0,0,0,2)} + \delta_{(2,1,0,0,2)}) \\
& - 2\delta_{(0,1,0,0,0)} \sum_{l=1}^n (\sum' f_l^{i,j} m_{ij})
\end{aligned}$$

Agora, escrevendo os termos $\sum' f_l^{i,j} a_{ij}$ e $\sum' f_m^{s,t} m_{st}$ como os (l, l) -ésimos elementos das matrizes $\phi^2 Z_2 / \delta_{(2,0,0,0,0)}$ e $\phi^2 (Z - Z_2) / \delta_{(2,0,0,0,0)}$, respectivamente, temos

$$\begin{aligned}
A_{1,\beta\phi} &= \frac{12\phi^{-4}}{n(\delta_{(2,0,0,0,2)} - 1)\delta_{(2,0,0,0,0)}^2} \{ (\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)})\delta_{(1,1,0,0,1)} \} \\
& \sum_{l=1}^n (z_{2l}) \sum_{m=1}^n (z_{mm} - z_{2mm}) \\
& + \frac{12}{n(\delta_{(2,0,0,0,2)} - 1)^2 \delta_{(2,0,0,0,0)}} \{ \delta_{(1,1,0,0,1)}(2\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)}) \\
& \sum_{l=1}^n (z_{ll} - z_{2l}) \} \\
& + \frac{6}{n(\delta_{(2,0,0,0,2)} - 1)\delta_{(2,0,0,0,0)}^2} \{ (3\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)})(\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)}) \} \\
& \sum (z_{lm} - z_{2lm}) z_{lm} \\
& - \frac{6}{n(\delta_{(2,0,0,0,2)} - 1)\delta_{(2,0,0,0,0)}^2} \{ (3\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)})(\delta_{(0,0,1,0,1)} - 2\delta_{(0,1,0,0,0)}) \}
\end{aligned}$$

$$\begin{aligned}
& \sum (z_{lm} - z_{2lm}) z_{lm} \\
& - \frac{6}{n(\delta_{(2,0,0,0,2)} - 1)\delta_{(2,0,0,0,0)}} (4\delta_{(3,0,0,0,1)} + \delta_{(4,0,0,0,2)} + \delta_{(2,1,0,0,2)} - 2\delta_{(0,1,0,0,0)}) \\
& \sum_{l=1}^n (z_{ll} - z_{2ll})
\end{aligned}$$

Como $\sum_{l=1}^n z_{2ll} = \text{tr}(Z_2) = p - q$, $\sum_{l=1}^n (z_{ll} - z_{2ll}) = \text{tr}(Z_d - Z_{2d}) = q$ e $\sum (z_{lm} - z_{2lm}) z_{2lm} = \text{tr}(ZZ_2) - \text{tr}(Z_2) = 0$, temos

$$\begin{aligned}
A_{1,\beta\phi} &= \frac{12}{n} \frac{(\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)})\delta_{(1,1,0,0,1)}}{(\delta_{(2,0,0,0,2)} - 1)\delta_{(2,0,0,0,0)}} (p - q)q \\
& - \frac{12}{n} \frac{\delta_{(1,1,0,0,1)}(2\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)})}{(\delta_{(2,0,0,0,2)} - 1)^2\delta_{(2,0,0,0,0)}} q \\
& - \frac{6}{n} \frac{(4\delta_{(3,0,0,0,1)} + \delta_{(4,0,0,0,2)} + \delta_{(2,1,0,0,2)} - 2\delta_{(0,1,0,0,0)})}{(\delta_{(2,0,0,0,2)} - 1)\delta_{(2,0,0,0,0)}} q.
\end{aligned}$$

A obtenção de $A_{2,\beta\phi}$ é conduzida de maneira análoga. De (3.11), temos

$$\begin{aligned}
A_{2,\beta\phi} &= -12 \sum' \frac{1}{\phi^6} \delta_{(1,1,0,0,1)}^2 \sum_{l=1}^n (f_l^{i,j}) \sum_{m=1}^n (f_m^{s,t}) a_{\phi\phi} m_{st} m_{ij} \\
& - 24 \sum' \delta_{(1,1,0,0,1)}^2 \left(\sum_{m=1}^n f_m^{s,r} \right) \left(\sum_{l=1}^n f_l^{i,j} \right) m_{ir} m_{js} a_{\phi\phi} \\
& = - \frac{12\phi^{-4}}{n(\delta_{(2,0,0,0,2)} - 1)} \delta_{(1,1,0,0,1)}^2 \left\{ \sum_{l=1}^n \left(\sum' (f_l^{i,j}) m_{ij} \right) \sum_{m=1}^n \left(\sum' (f_m^{s,t}) m_{st} \right) \right. \\
& \quad \left. + 2 \sum_{l,m} \left(\sum' f_{ml}^{i,r} m_{ir} \right) \sum_{l,m} \left(\sum' f_{lm}^{j,s} m_{js} \right) \right\} \\
& = - \frac{12\delta_{(1,1,0,0,1)}^2}{n(\delta_{(2,0,0,0,2)} - 1)\delta_{(2,0,0,0,0)}^2} \left\{ \sum_{l=1}^n (z_{ll} - z_{2ll}) \sum_{m=1}^n (z_{mm} - z_{2mm}) \right. \\
& \quad \left. + 2 \sum (z_{lm} - z_{2lm})(z_{lm} - z_{2lm}) \right\}
\end{aligned}$$

que se reduz a

$$\begin{aligned}
A_{2,\beta\phi} &= - \frac{12\delta_{(1,1,0,0,1)}^2}{n(\delta_{(2,0,0,0,2)} - 1)\delta_{(2,0,0,0,0)}^2} \{ (\text{tr}(Z_d - Z_{2d}))^2 + 2\text{tr}(Z_d - Z_{2d}) \} \\
& - \frac{12\delta_{(1,1,0,0,1)}^2}{n(\delta_{(2,0,0,0,2)} - 1)\delta_{(2,0,0,0,0)}^2} (q^2 + 2q).
\end{aligned}$$

Apêndice D

Neste Apêndice, encontraremos os valores dos δ 's associados à correção de Bartlett e tipo Bartlett para cada distribuição na classe dos modelos simétricos. A notação g^r e $\delta(a, b, c, d, e)$ estão definidos na seção 3.2. Consideramos $z_l = (y_l - \mu_l)/\phi$ com $\mu_l = f(x_l; \beta)$.

1. Distribuição Normal

Seja $y_l \sim N(\mu_l, \phi^2)$ com função densidade escrita como

$$\pi(y; \mu_l, \phi) = \frac{1}{\sqrt{2\pi\phi}} \exp\left\{-\frac{1}{2} \left(\frac{y - \mu_l}{\phi}\right)^2\right\}, y, \mu_l \in R.$$

Temos que a função $g(z_l) = -z_l^2/2$. Assim, como $g^{(r)} = \partial^r g / \partial z^r$, podemos dizer que

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

Usando a notação $\delta(a, b, c, d, e)$ e fazendo as esperanças temos

$$\begin{array}{ll} \delta_{(2,0,0,0,0)} = 1, & \delta_{(0,0,0,1,0)} = 0, \\ \delta_{(2,1,0,0,0)} = -1, & \delta_{(0,0,0,1,2)} = 0, \\ \delta_{(0,0,1,0,1)} = 0, & \delta_{(0,1,0,0,0)} = -1, \\ \delta_{(0,0,1,0,3)} = 0, & \delta_{(1,1,0,0,1)} = 1, \\ \delta_{(4,0,0,0,2)} = 15, & \delta_{(2,0,0,0,2)} = 3, \\ \delta_{(1,0,1,0,0)} = 0, & \delta_{(0,1,0,0,2)} = -1, \\ \delta_{(2,1,0,0,2)} = -3, & \delta_{(3,0,0,0,1)} = -3, \end{array}$$

2. Distribuição de Cauchy

Seja $y_l \sim C(\mu_l, \phi^2)$ com função densidade escrita como

$$\pi(y; \mu_l, \phi) = \frac{1}{\pi\phi} \left[1 + \left(\frac{y - \mu_l}{\phi}\right)^2\right]^{-1}, y, \mu_l \in R.$$

Temos que a função $g(z_l) = \log\left(\frac{1}{\phi}[1 + z_l^2]^{-1}\right)$.

Assim podemos dizer que $g^{(1)} = 2zw$, onde $w = -1/(1 + z^2)$, $g^{(2)} = 2w + 4z^2w^2$, $g^{(3)} = 12zw^2 + 16z^3w^3$, $g^{(4)} = 12w^2 + 96z^2w^2 + 96z^4w^4$.

Usando a notação $\delta(a, b, c, d, e)$ e fazendo as esperanças temos

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

3. Distribuição t -Student

Seja $y_l \sim C(\mu_l, \phi^2, \nu)$ com função densidade escrita como

$$\pi(y; \mu_l, \phi, \nu) = \frac{1}{\phi} \frac{\nu^{\nu/2}}{B(1/2, \nu/2)} \left[\nu + \left(\frac{y - \mu_l}{\phi} \right)^2 \right]^{-\frac{\nu+1}{2}}, y, \mu_l \in R.$$

Temos que a função $g(z_l) = \log \left(\frac{\nu^{\nu/2}}{B(1/2, \nu/2)} [\nu + z_l^2]^{-\frac{\nu+1}{2}} \right)$.

Derivando $g(z_l)$ com respeito a z , temos:

$$\begin{aligned} g^{(1)} &= 2zw, \quad \text{onde } w = -(\nu+1)/[2(\nu+z^2)], \quad g^{(2)} = 2w + \frac{8z^2w^2}{(\nu+1)}, \quad g^{(3)} = \\ &= \frac{24}{(\nu+1)}zw^2 + \frac{64}{(\nu+1)^2}z^3w^3, \quad g^{(4)} = \frac{24}{(\nu+1)}w^2 + \frac{384}{(\nu+1)^2}z^2w^3 + \frac{768}{(\nu+1)^3}z^4w^4. \end{aligned}$$

Usando a notação $\delta(a, b, c, d, e)$ e fazendo as esperanças temos:

$$\begin{array}{ll} \delta_{(2,0,0,0,0)} = \frac{\nu+1}{\nu+3}, & \delta_{(0,0,0,1,0)} = \frac{6(\nu+1)(\nu+2)}{\nu(\nu+5)(\nu+7)}, \\ \delta_{(0,0,0,1,2)} = \frac{6}{(\nu+3)} \left(\frac{\nu-19}{\nu+5} + \frac{120}{(\nu+5)(\nu+7)} \right), & \delta_{(2,1,0,0,0)} = -\frac{(\nu+1)^3(\nu+2)}{\nu(\nu+3)(\nu+5)(\nu+7)}, \\ \delta_{(0,0,1,0,1)} = \frac{6(\nu+1)}{(\nu+3)(\nu+5)}, & \delta_{(2,0,0,0,2)} = \frac{3(\nu+1)}{(\nu+3)}, \\ \delta_{(0,1,0,0,2)} = \frac{3-\nu}{(\nu+3)}, & \delta_{(0,1,0,0,0)} = -\frac{(\nu+1)}{\nu+3}, \\ \delta_{(1,1,0,0,1)} = \frac{(\nu+1)(\nu-1)}{(\nu+3)(\nu+5)}, & \delta_{(0,0,1,0,3)} = \frac{6(3\nu-5)}{(\nu+3)(\nu+5)}, \\ \delta_{(4,0,0,0,2)} = \frac{15(\nu+1)^3}{(\nu+3)(\nu+5)(\nu+7)}, & \delta_{(3,0,0,0,1)} = -\frac{3(\nu+1)^2}{(\nu+3)(\nu+5)}, \\ \delta_{(2,1,0,0,2)} = \frac{3(\nu+1)^2(3-\nu)}{(\nu+3)(\nu+5)(\nu+7)}, & \delta_{(1,0,1,0,0)} = -\frac{6(\nu+1)(\nu+2)}{\nu(\nu+5)(\nu+7)}. \end{array}$$

4. Distribuição t -Student generalizada

Seja $y_l \sim C(\mu_l, \phi^2, r, s)$ com função densidade escrita como

$$\pi(y; \mu_l, \phi, r, s) = \frac{1}{\phi} \frac{s^{r/2}}{B(1/2, r/2)} \left[s + \left(\frac{y - \mu_l}{\phi} \right)^2 \right]^{-\frac{r+1}{2}}, y, \mu_l \in R.$$

Temos que a função $g(z_l) = \log \left(\frac{s^{r/2}}{B(1/2, r/2)} [s + z_l^2]^{-\frac{r+1}{2}} \right)$. Derivando $g(z_l)$ com respeito a z , temos:

$g^{(1)} = 2zw$, onde $w = -(r+1)/[2(s+z^2)]$, $g^{(2)} = 2w + \frac{8}{(r+1)}z^2w^2$, $g^{(3)} = \frac{24}{(r+1)}zw^2 + \frac{64}{(r+1)^2}z^3w^3$, $g^{(4)} = \frac{24}{(r+1)}w^2 + \frac{384}{(r+1)^2}z^2w^3 + \frac{768}{(r+1)^3}z^4w^4$. Usando a notação $\delta(a, b, c, d, e)$ e fazendo as esperanças temos:

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

5. Distribuição Logística I

Seja $y_l \sim LI(\mu_l, \phi^2)$ com função densidade escrita como

$$\pi(y; \beta, \phi) = \frac{c}{\phi} \frac{\exp \left\{ - \left(\frac{y - \mu_l}{\phi} \right)^2 \right\}}{\left(1 + \exp \left\{ - \left(\frac{y - \mu_l}{\phi} \right)^2 \right\} \right)^2}, \quad y, \mu_l \in \mathbb{R}.$$

Temos que a função $g(z_l) = \log \left(\frac{e^{-z_l^2}}{1+e^{-z_l^2}} \right)$. Derivando $g(z_l)$ com respeito a z , temos:

$$\begin{aligned} g^{(1)} &= 2zw, \text{ onde } w = (e^{-z^2} - 1)/(1 + e^{-z^2}), \\ g^{(2)} &= 2w + 2z^2(w^2 - 1), \quad g^{(3)} = 6z(w^2 - 1) + 4z^3w(w^2 - 1), \\ g^{(4)} &= 6(w^2 - 1) + 24wz^2(w^2 - 1) + 4z^4(w^2 - 1)(3w^2 - 1). \end{aligned}$$

Usando a notação $\delta(a, b, c, d, e)$ e fazendo as esperanças temos:

$$\begin{aligned}
\delta_{(2,0,0,0,0)} &\approx 1,477240176, & \delta_{(0,0,0,1,0)} &\approx 4,259052264, \\
\delta_{(2,1,0,0,2)} &\approx -10,92854975, & \delta_{(2,1,0,0,0)} &\approx 4,153806544, \\
\delta_{(0,0,0,1,2)} &\approx 2,65931983, & \delta_{(2,0,0,0,2)} &\approx 4,013783934, \\
\delta_{(0,0,1,0,1)} &\approx -1,27916363, & \delta_{(0,1,0,0,0)} &\approx -1,477240176, \\
\delta_{(0,1,0,0,2)} &\approx -2,013783934, & \delta_{(0,0,1,0,3)} &\approx -0,508877866, \\
\delta_{(1,1,0,0,1)} &\approx 2,756409976, & \delta_{(3,0,0,0,1)} &\approx 46,76577386, \\
\delta_{(4,0,0,0,2)} &\approx 46,76577386, & \delta_{(1,0,1,0,0)} &\approx -4,259052264.
\end{aligned}$$

6. Distribuição Logística II

Seja $y_l \sim LII(\mu_l, \phi^2)$ com função densidade escrita como

$$\pi(y; \beta, \phi) = \frac{1}{\phi} \frac{\exp\left\{\left(\frac{y-\mu_l}{\phi}\right)\right\}}{\left(1 + \exp\left\{\left(\frac{y-\mu_l}{\phi}\right)\right\}\right)^2}, \quad y, \mu_l \in \mathbb{R}.$$

Temos que a função $g(z_l) = \log\left(\frac{e^{z_l}}{(1+e^{z_l})^2}\right)$.

Derivando $g(z_l)$ com respeito a z , temos:

$$g^{(1)} = \frac{1-e^z}{1+e^z}, \quad g^{(2)} = \frac{-2e^z}{(1+e^z)^2}, \quad g^{(3)} = \frac{2(e^{2z}-e^z)}{(1+e^z)^3}, \quad g^{(4)} = \frac{-2(e^{3z}-4e^{2z}+e^z)}{(1+e^z)^4}.$$

Usando a notação $\delta(a, b, c, d, e)$ e fazendo as esperanças temos:

$$\begin{aligned}
\delta_{(2,0,0,0,0)} &= 1/3, & \delta_{(0,0,0,1,0)} &= 1/15, \\
\delta_{(1,0,1,0,0)} &= -1/15 & \delta_{(2,1,0,0,0)} &= -1/15, \\
\delta_{(0,0,0,1,2)} &\approx -0,11401, & \delta_{(2,1,0,0,2)} &\approx -0,21932 \\
\delta_{(0,0,1,0,1)} &= 1/6, & \delta_{(0,1,0,0,0)} &= -1/3, \\
\delta_{(0,1,0,0,2)} &\approx -0,42996, & \delta_{(0,0,1,0,3)} &\approx 0,64493, \\
\delta_{(1,1,0,0,1)} &= 1/6, & \delta_{(3,0,0,0,1)} &= -2/3, \\
\delta_{(4,0,0,0,2)} &\approx 1,99131,
\end{aligned}$$

Apêndice E

Este apêndice apresenta os programas escritos na linguagem de programação Ox para geração dos resultados de simulação obtidos nesta dissertação. No primeiro programa, encontra-se o programa de simulação de Monte Carlo e Bootstrap; no segundo, contém as funções das distribuições consideradas. No terceiro, encontram-se as diretrizes de ligação entre os outros dois arquivos.

Programa Principal

```

/*****
Autor: Katya Silene Porto Rodrigues Braga
Programa: Prog_tese.ox

Funcao: Calcular a taxa de rejeicao dos testes da razao de verossimilhanças
usual, teste Score, correcao
de Bartlett, correcao tipo-Bartlett e as versoes bootstrap.

A media ou o parametro \mu tem uma componente sistematica da forma:
\mu = \beta_0 + \beta_1 *x_2 + \beta_2 *x_3 + \beta_3 *exp{\beta_4 *x_5}
*****/

#include <oxstd.h>
#include <oxprob.h>
#import <maximize>
#import <solvenle>
#include <oxfloat.h>
#import <funcoes1>
// Definicao de variaveis globais:
static decl NBOOT= 500;
static decl INREP =10000;
decl s_mX;
decl s_vy,dphi_hat_beta,s_vyboot;
static decl N ;

decl fpout,fpout1,fpout2,fpout3,fpout4,fpout5,file; //arquivo de saida

//Corpo principal do programa
main()
{
decl vp,vpr,vpr_1,vprnull,vprnull_1,irr,irr1,irrnull,irrnull1,dfuncr,
dfuncr_1,dfuncrnull,dfuncrnull_1,
ib,iq, dphi, dphi2, vbeta,vbetaMLE,vbetaMLEnull,vbetaMLE1,
vbetaMLEnull1,dphiMLE1,dphiMLEnull1,
vbetahat_phi,i,l,ii,j,jj,t,m,q,time,time2,cprint,mu,mu_hat,mMoments,
vR,vR1,vR2,vR3;
decl cj,vp_Ho_B,vprB00T,vprnullB00T,irrB00T,irrnullB00T,dfuncrB00T,
dfuncrnullB00T,dMaxLogB00T,dMaxLogNullB00T,mrazaoverossimilhancaboot;

```



```

decl mEstQuant,vLR,vLRp,vLRB,vLR_1,W_s,W_tm,Wc,vLRc,vLRcc,vLRB_1,
    k_ML1,vcv10,vcv5,vcv1,vcv05,mRej,mRej_1,tn,B,k;
decl dMaxLog_1,dMaxLogNull_1,dMaxProf_1,dMaxProfNull_1;
decl Shat,J,U,vphiMLE,Ihat,Itil,U_1,Ihat_1,Itil_1,
    mB,mC,d_m,mBd,mCd,mB2,mC2,d2_m,mB2d,mC2d,max,gm;
cprint = floor(INREP/5); // print iteration flow control parameter
/* exit file */
file= fopen("saida22c1.txt", "w");
//fpout=fopen( "RVlog2.txt","w");
//fpout1=fopen("Wlog2.txt","w");
//fpout2=fopen("RVclog2.txt","w");
//fpout3=fopen("RVcclog2.txt","w");
//fpout4=fopen("SRclog2.txt","w");
//fpout5=fopen("SRcclog2.txt","w");
// Marcador de tempo
time = timer();
oxwarning(0);
// Usando o George Marsaglia
decl m1 = 2700;
decl m2 = 1000;
ranseed({m1, m2});

tn=<30;10;20;40;50>;
for(B=0;B<5;++B){
    N=tn[B];

max=1;
gm=0.0;

    // Critical values for the tests:
    mEstQuant = zeros(8, 4);
    mEstQuant[0][0]=vcv10=quanchi(0.90,1);
    mEstQuant[0][1]=vcv5=quanchi(0.95,1);
    mEstQuant[0][2]=vcv1=quanchi(0.99,1);
    mEstQuant[0][3]=vcv05=quanchi(0.995,1);

for (jj = 0; jj < max; ++jj) {

//Declaracoes para \phi
decl k_MLB00T=0; //contador de falha para \beta para Bootstrap
k_ML1=0; //contador de falha \beta1 para usual
ib=1; // numeros de parametros betas de interesse
k=1; //numeros de parametros betas de pertubacao
iq = 1; // numeros de parametros phi
s_mX = ranu(N,4); // nao linear
dphi =1;
dphi2=sqrt(dphi);
vbeta=zeros(k,1);
decl vbetap= constant(gm,ib,1);

```

```

vbetaMLE1 = zeros(ib+k, INREP);
vbetaMLEnull1= zeros(ib+k,INREP);
dphiMLE1= zeros(1,INREP);
dphiMLEnull1= zeros(1,INREP);
vLR_1 = zeros(1, INREP);
vLRc= zeros(1, INREP);
vLRcc= zeros(1, INREP);
mrazaoverossimilhancaboot= zeros(1,NBOOT);
W_s = zeros(1, INREP);
Wc = zeros(1, INREP);
W_tm = zeros(1, INREP);
mRej_1 = zeros(15, 4);
decl vcvBoot= zeros(4,INREP);
decl pcwBoot= zeros(4,INREP);
mMoments = zeros(7,2);

U_1=zeros(ib+iq+k,1);
mB=zeros(N,N);
mBd=zeros(N,N);
mC= zeros(N,N);
mCd=zeros(N,N);
d_m= zeros(1,N);
mB2=zeros(N,N);
mB2d=zeros(N,N);
mC2= zeros(N,N);
mC2d=zeros(N,N);
d2_m= zeros(1,N);

// Set maximum number of iterations:
MaxControlEps( 1e-4, 5e-3); // set the accuracy for BFGS estimation
MaxControl(300,-1,-1); // set maximum # of iterations =100

//declaracao da matriz de derivadas mu(theta,x)=beta0+beta1*exp(beta2*x)
decl x1=s_mX[] [0];
decl x2=s_mX[] [1];
decl x3=s_mX[] [2];
decl x4=exp(vbetap.*s_mX[] [3]);
decl x5=vbeta[0]*s_mX[] [3].*exp(vbetap.*s_mX[] [3]);
decl X= x5~x4;
decl X1=x5;
decl X2=x4;
//fim da declaracao da matriz de derivadas (X til)

decl vlpredp=vbeta[0]*(exp(vbetap .*s_mX[] [3])); // 2 beta
// decl vlpredp=vbeta[0]*(exp(vbetap .*s_mX[] [3]))+vbeta[1]
// .s_mX[] [2]; // 3 betas
// decl vlpredp=vbeta[0]*(exp(vbetap .*s_mX[] [3]))+vbeta[1]
// .s_mX[] [2]+ vbeta[2].s_mX[] [1]; //4 betas
// decl vlpredp=vbeta[0]*(exp(vbetap .*s_mX[] [3]))+vbeta[1]

```

```

        .*s_mX[] [2]+ vbeta[2].*s_mX[] [1]+ vbeta[3].*s_mX[] [0]; //5 betas
decl mup=vlpredp;

//Declaracao das esperancas
////Normal
//decl delta_10001= -1;
//decl delta_1= 0;
//decl delta_01001= 0;
//decl delta_2= 1;
//decl delta_0001= 0;
//decl delta_01= -1;
//decl delta_01002= -1;
//decl delta_00103= 0;
//decl delta_00101= 0;
//decl delta_00012= 0;
//decl delta_11001= 1;
//decl delta_101= 0;
//decl delta_21= -1;
//decl delta_20002= 3;
//decl delta_30001= -3;
//decl delta_40002= 15;
//decl delta_21002= -3;

////
////T-Student v=1, Cauchy
//decl delta_10001= -1;
//decl delta_1= 0;
//decl delta_01001= 0;
//decl delta_2= 1/2;
//decl delta_0001= 3/4;
//decl delta_01= -1/2;
//decl delta_01002= 1/2;
//decl delta_00103= -1/2;
//decl delta_00101= 1/2;
//decl delta_00012= -3/4;
//decl delta_11001= 0;
//decl delta_101= -3/4;
//decl delta_21= -1/8;
//decl delta_20002=3/2;
//decl delta_30001=-1/2;
//decl delta_40002=5/8;
//decl delta_21002=1/8;

//T-Student v=2
decl delta_10001= -1;
decl delta_1= 0;

```

```

decl delta_01001= 0;
decl delta_2= 3/5;
decl delta_0001= 4/7;
decl delta_01= -3/5;
decl delta_01002= 1/5;
decl delta_00103= 6/35;
decl delta_00101= 18/35;
decl delta_00012= -66/105;
decl delta_11001= 3/35;
decl delta_101= -5/7;
decl delta_21= -6/35;
decl delta_20002=9/5;
decl delta_30001=-27/35;
decl delta_40002=9/7;
decl delta_21002=3/35;

```

```
//
```

```

////T-Student v=4
//decl delta_10001= -1;
//decl delta_1= 0;
//decl delta_01001= 0;
//decl delta_2= 5/7;
//decl delta_0001= 5/11;
//decl delta_01= -5/7;
//decl delta_01002= -1/7;
//decl delta_00103= 2/3;
//decl delta_00101= 10/21;
//decl delta_00012= -30/77;
//decl delta_11001= 5/21;
//decl delta_101= -5/11;
//decl delta_21= -125/462;
//decl delta_20002=15/7;
//decl delta_30001=- 25/21;
//decl delta_40002=625/231;
//decl delta_21002=-25/231;

```

```
//
```

```

////Logistica
//decl delta_10001= -1;
//decl delta_1= 0;
//decl delta_01001= 0;
//decl delta_2= 1/3;
//decl delta_0001= 1/15;
//decl delta_01= -1/3;
//decl delta_01002= -0.42996;
//decl delta_00103= 0.64493;
//decl delta_00101= 1/6;
//decl delta_00012= -0.11401;

```

```

//decl delta_11001= 1/6;
//decl delta_101= -1/15;
//decl delta_21= -1/15;
//decl delta_20002= 2.42996;
//decl delta_30001= -2/3;
//decl delta_40002= 1.99131;
//decl delta_21002= -0.21932;

//Declaracao dos d's e m's
decl m1= delta_01002 -1;
decl m2= 4- delta_00103 -(6*delta_01002);
decl m3= (delta_00101+(2*delta_01))/delta_2;
decl m4= (delta_00012-(6*delta_11001))/delta_2;
decl d0= delta_0001/(4*(delta_2^2));
decl d1= -((m3/(2*(m1^2)))*(m2+(m1*(2+m3))))-(m4/(2*m1));
decl d2= -((m3^2)/(2*m1));
decl d3= -delta_01/(4*(delta_2^2));
decl d4= delta_01^2/(2*(delta_2^3));

/*Declaracao das matrizes Z, X_m, B, C e D para calculo do */
//fator de correcao de Bartlett
decl pXinv= invertgen(X'*X);
decl mX2inv= invertgen(X2'*X2);
decl mZ= X*pXinv*X';
decl mZd= diagonalize(mZ);
decl mZ2= X2*mX2inv*X2';
decl mZ2d= diagonalize(mZ2);

decl mX_m = new array [N];
for(l = 0; l < N; ++l)
{ mX_m[l]=((((s_mX[l][3]) .^2)*vbeta[0] .*exp(vbetap .*s_mX[l][3])~
(s_mX[l][3]*exp(vbetap .*s_mX[l][3])))|
(s_mX[l][3] .*exp(vbetap .*s_mX[l][3])~0));}

decl mX2_m = new array [N];
for(l = 0; l < N; ++l)
{ mX2_m[l]=((0));}

for(ii=0; ii < N; ++ii)
{
for(j=0; j < N; ++j)
{
mB[ii][j]= trace(pXinv*mX_m[ii]*pXinv*mX_m[j]);
}
}
mBd= diagonalize(mB);

```

```

for(ii=0; ii < N; ++ii)
{
  for(j=0; j < N; ++j)
  {
    mB2[ii][j]= trace(mX2inv*mX2_m[ii]*mX2inv*mX2_m[j]);
  }
}
mB2d= diagonalize(mB2);

for(t=0 ; t < N; ++t)
{
  for(m=0; m < N; ++m)
  {
    mC[t][m]= X[t][]*pXinv*mX_m[m]*pXinv*X[t][]';
  }
}
mCd=diagonalize(mC);

for(t=0 ; t < N; ++t)
{
  for(m=0; m < N; ++m)
  {
    mC2[t][m]= X2[t][]*mX2inv*mX2_m[m]*mX2inv*X2[t][]';
  }
}
mC2d=diagonalize(mC2);
for(q=0; q < N; ++q)
{
  d_m[q]= trace(mX_m[q]*pXinv);}
decl mD= diag(d_m);

for(q=0; q < N; ++q)
{
  d2_m[q]= trace(mX2_m[q]*mX2inv);}
decl mD2= diag(d2_m);

// Declaracao dos A_i's para o fator de correcao tipo Bartlett
decl b_0= (delta_21/(delta_2^2)) +1;
decl b_1= (delta_11001*(delta_11001-delta_01))/(delta_2^2*(delta_20002-1));
decl b_21= (2*delta_11001*((2*delta_01002)+delta_00103));
decl b_22= (delta_20002-1)*((4*delta_30001)+delta_40002+delta_21002-
  (2*delta_01));
decl b_23= (delta_2*((delta_20002-1)^2))^( -1);
decl b_2= b_23*(b_21+b_22);
decl b_3= (delta_11001^2)/((delta_2^2)*(delta_20002-1));

decl A_1= ((12*b_0)*(trace(mZd*mZ2d)-trace(mZ2d*mZ2d)))+(((12*b_1)/N)*ib*k)
-(((6*b_2)/N)*ib);
decl A_2= -((9*b_0)*(trace(mZd*mZd)-(2*trace(mZd*mZ2d))+trace(mZ2d*mZ2d)))
-(((12*b_3)/N)*ib*(ib+2));

```

```

decl A_3=0;

decl a= 0;
decl b= (A_2-(2*A_3))/(12*ib*(ib+2));
decl c= (A_1-A_2)/(12*ib);

for(i=0;i<INREP;++i)
{

// s_vy= mup + dphi*rann(N,1); // normal
// s_vy= mup + dphi*rant(N,1,1); //T-Student v=1
s_vy= mup + dphi*rant(N,1,2); //T-Student v=2
// s_vy= mup + dphi*rant(N,1,4); //T-Student v=4
// s_vy= mup + dphi*ranlogistic(N,1); //Logistica
// // Quasi-Newton maximization: restricted

vpr_1=0|vbeta|dphi;
vprnull_1=vbeta|dphi;

//Maximizacoes para \beta1
///Maximizacoes para Normal
// irr1 = MaxBFGS(flognormalr1, &vpr_1, &dfuncr_1, 0,TRUE);
// if (irr1 != MAX_CONV && irr1 != MAX_WEAK_CONV){
// ++k_ML1; // check convergence (H1)
// --i;
// continue;
// }
//
// irrnull1 = MaxBFGS(flognormalrnull1, &vprnull_1, &dfuncrnull_1, 0,TRUE);
// if (irrnull1 != MAX_CONV && irrnull1 != MAX_WEAK_CONV){
// ++k_ML1; // check convergence (H0) Lmpr
// --i;
// continue
// }

///Maximizacoes para T-Student v=1
// irr1 = MaxBFGS(flogtstudent1r1, &vpr_1, &dfuncr_1, 0,TRUE);
// if (irr1 != MAX_CONV && irr1 != MAX_WEAK_CONV){
// ++k_ML1; // check convergence (H1)
// --i;
// continue;
// }
// irrnull1 = MaxBFGS(flogtstudent1rnull1, &vprnull_1, &dfuncrnull_1, 0,TRUE);
// if (irrnull1 != MAX_CONV && irrnull1 != MAX_WEAK_CONV){
// ++k_ML1; // check convergence (H0) Lmpr
// --i;
// continue
// }

```

```

//Maximizacoes para T-Student v=2
    irr1 = MaxBFGS(flogtstudent2r1, &vpr_1, &dfuncr_1, 0,TRUE);
    if (irr1 != MAX_CONV && irr1 != MAX_WEAK_CONV){
        ++k_ML1;           // check convergence (H1)
        --i;
        continue;
    }
    irrnull1 = MaxBFGS(flogtstudent2rnull1, &vprnull_1, &dfuncrnull_1, 0,TRUE);
    if (irrnull1 != MAX_CONV && irrnull1 != MAX_WEAK_CONV){
        ++k_ML1;           // check convergence (H0) Lmpr
        --i;
        continue;
    }

////Maximizacoes para T-Student v=4
//    irr1 = MaxBFGS(flogtstudent4r1, &vpr_1, &dfuncr_1, 0,TRUE);
//    if (irr1 != MAX_CONV && irr1 != MAX_WEAK_CONV){
//        ++k_ML1;           // check convergence (H1)
//        --i;
//        continue;
//    }
//    irrnull1 = MaxBFGS(flogtstudent4rnull1, &vprnull_1, &dfuncrnull_1, 0,TRUE);
//    if (irrnull1 != MAX_CONV && irrnull1 != MAX_WEAK_CONV){
//        ++k_ML1;           // check convergence (H0) Lmpr
//        --i;
//        continue;
//    }

////Maximizacoes para Logistica
//    irr1 = MaxBFGS(flogisticar1, &vpr_1, &dfuncr_1, 0,TRUE);
//    if (irr1 != MAX_CONV && irr1 != MAX_WEAK_CONV){
//        ++k_ML1;           // check convergence (H1)
//        --i;
//        continue;
//    }
//    irrnull1 = MaxBFGS(flogisticarnull1, &vprnull_1, &dfuncrnull_1, 0,TRUE);
//    if (irrnull1 != MAX_CONV && irrnull1 != MAX_WEAK_CONV){
//        ++k_ML1;           // check convergence (H0) Lmpr
//        --i;
//        continue;
//    }

//Declaracao de variaveis para o testes verossimilhanca usual
dMaxLog_1=dfuncr_1;
dMaxLogNull_1=dfuncrnull_1;

//Declaracao dos parametros
vbetaMLE1[][i] = vpr_1[0:1];
dphiMLE1[][i]=vpr_1[2];
vbetaMLEnull1[][i] = 0|vprnull_1[0];
dphiMLEnull1[][i]=vprnull_1[1];

```



```

/*Declaracao da parte linear e nao-linear do fator de correcao */
//de bartlett para \phi conhecido
decl h0= (trace((mZd.^2))- (trace((mZ2d.^2))));
decl L_p= ((d0) *h0)+d1/N +(((2*(ib+k))-ib)/(2*N))*d2;
decl NL_p= dphiMLEnull1[] [i]^2*((d3*(trace((2*mBd)-(mD^2))-
(trace((2*mB2d)-(mD2^2)))))+(d4*((1/2*(sumr(sumc(mD.*mZ.*mD))))
-(trace(mZ*mB)))-(1/2*(sumr(sumc(mD2.*mZ2.*mD2))))+trace(mZ2*mB2)))));

decl e_p= L_p +NL_p;
decl c_b1=1+((e_p)/ib);
decl c_b2= 1-((e_p)/ib);

////Teste da razao de Verossimilhanças Bootstrap
//      //Calculando o LRBoot em fBoot
//      //Chute inicial em Ho
vp_Ho_B= vbeta|dphi;
//
////Maximizando a funcao de log-verossimilhança em Ho
// Para Normal
MaxBFGS(flognormalboot, &vp_Ho_B, &dfuncrnullBOOT, 0,TRUE);
// Para T-Student 1 grau de liberdade
MaxBFGS(flogstudent1boot, &vp_Ho_B, &dfuncrnullBOOT, 0,TRUE);
// Para T-Student 2 grau de liberdade
MaxBFGS(flogtstudent2boot, &vp_Ho_B, &dfuncrnullBOOT, 0,TRUE);
// Para T-Student 4 grau de liberdade
MaxBFGS(flogstudent4boot, &vp_Ho_B, &dfuncrnullBOOT, 0,TRUE)
// Para Logistica;
MaxBFGS(flogisticaboot, &vp_Ho_B, &dfuncrnullBOOT, 0,TRUE);
decl vparBP = 0|vp_Ho_B;
//
////Funcao de loop para verossimilhança por bootstrap
//decl Pcrit = fBoot(vparBP, N, NBOOT);
//
////Funcao de loop para escore por bootstrap
//decl Wcrit = WBoot(vparBP, N, NBOOT);
//
////Calculando os pontos criticos da razao de verossimilhança por Bootstrap
//vcvBoot[0][i]=quantiler(Pcrit, 0.90);
//vcvBoot[1][i]=quantiler(Pcrit,0.95);
//vcvBoot[2][i]=quantiler(Pcrit,0.99);
//vcvBoot[3][i]=quantiler(Pcrit,0.995);
//
////Calculando os pontos criticos do escore por Bootstrap
//pcwBoot[0][i]=quantiler(Wcrit, 0.90);
//pcwBoot[1][i]=quantiler(Wcrit,0.95);
//pcwBoot[2][i]=quantiler(Wcrit,0.99);
//pcwBoot[3][i]=quantiler(Wcrit,0.995);
//

//fim das maximizacoes

```

```

// //Declaracao da matrix X hat
decl x1hat=s_mX[] [0];
decl x2hat=s_mX[] [1];
decl x3hat=s_mX[] [2];
decl x4hat=exp(vbetaMLE1[0][i].*s_mX[] [3]);
decl x5hat=vbetaMLE1[1][i]*s_mX[] [3].*exp(vbetaMLE1[0][i].*s_mX[] [3]);
decl X1hat=x5hat;
decl X2hat=x4hat;

// //declaracao da matriz X til
decl x1til=s_mX[] [0];
decl x2til=s_mX[] [1];
decl x3til=s_mX[] [2];
decl x4til=exp(vbetaMLEnull1[0][i].*s_mX[] [3]);
decl x5til=vbetaMLEnull1[1][i]*s_mX[] [3].*exp(vbetaMLEnull1[0][i].*s_mX[] [3]);
decl X1til= x5til;
decl X2til= x4til;

// declaracao de mu(til)\beta_1 interesse
// 2 beta
decl vlpredtil1=vbetaMLEnull1[1][i]*exp(vbetaMLEnull1[0][i].*s_mX[] [3]);
// decl vlpredtil1=vbetaMLEnull1[1][i]*exp(vbetaMLEnull1[0][i].*s_mX[] [3])
+vbetaMLEnull1[2][i].*s_mX[] [2]; //3 betas
// decl vlpredtil1=vbetaMLEnull1[1][i]*exp(vbetaMLEnull1[0][i].*s_mX[] [3])
+vbetaMLEnull1[2][i].*s_mX[] [2]
// + vbetaMLEnull1[3][i].*s_mX[] [1]; //4 betas
// decl vlpredtil1=vbetaMLEnull1[1][i]*exp(vbetaMLEnull1[0][i].*s_mX[] [3])
+vbetaMLEnull1[2][i].*s_mX[] [2]+
//vbetaMLEnull1[3][i].*s_mX[] [1]+ vbetaMLEnull1[4][i].*s_mX[] [0]; //5 betas
decl mutil1=vlpredtil1;
decl s_vztil= (s_vy - mutil1)/dphiMLEnull1[0][i];

// declaracao de mu(hat) \beta interesse
decl vlpredhat1=vbetaMLE1[1][i]*exp(vbetaMLE1[0][i].*s_mX[] [3]); // 2 beta
// decl vlpredhat1=vbetaMLE1[1][i]*exp(vbetaMLE1[0][i].*s_mX[] [3])
//+vbetaMLE1[2][i].*s_mX[] [2]; //3 betas
// decl vlpredhat1=vbetaMLE1[1][i]*exp(vbetaMLE1[0][i].*s_mX[] [3])
//+vbetaMLE1[2][i].*s_mX[] [2]+ vbetaMLE1[3][i].*s_mX[] [1]; //4 betas
// decl vlpredhat1=vbetaMLE1[1][i]*exp(vbetaMLE1[0][i].*s_mX[] [3])
//+vbetaMLE1[2][i].*s_mX[] [2]+ vbetaMLE1[3][i].*s_mX[] [1]+ vbetaMLE1[4][i]
//.*s_mX[] [0]; //5 betas
decl muhat1=vlpredhat1;

// //Declaracao das derivadas da funcao g(z) em relacao a z
// // Normal
// decl g_1hat= -((s_vy-muhat1)./dphiMLE1[] [i]);
// decl g_1til= -((s_vy-mutil1)./dphiMLEnull1[] [i]);

// //T-Student v =1
//decl g_1hat= -2*((s_vy-muhat1)./dphiMLE1[] [i])./(ones(N,1))

```

```

+((s_vy-muhat1)./dphiMLE1[] [i]).^2); //g^1 sob H1
//decl g_1til=-2*((s_vy-muhat1)./dphiMLEnull1[] [i])./(ones(N,1)
+((s_vy-muhat1)./dphiMLEnull1[] [i]).^2); //g^1 sob H0

//T-Student v=2
decl g_1hat= -3*((s_vy-muhat1)./dphiMLE1[] [i])./(2*ones(N,1)
+((s_vy-muhat1)./dphiMLE1[] [i]).^2); //g^1 sob H1
decl g_1til=-3*((s_vy-muhat1)./dphiMLEnull1[] [i])./(2*ones(N,1)
+((s_vy-muhat1)./dphiMLEnull1[] [i]).^2); //g^1 sob H0

// //T-Student v=4
//decl g_1hat= -5*((s_vy-muhat1)./dphiMLE1[] [i])./(4*ones(N,1)
+((s_vy-muhat1)./dphiMLE1[] [i]).^2); //g^1 sob H1 da T-student v=4
//decl g_1til=-5*((s_vy-muhat1)./dphiMLEnull1[] [i])./(4*ones(N,1)
+((s_vy-muhat1)./dphiMLEnull1[] [i]).^2); //g^1 sob H0 da T-student v=4
//
////

// //Logistica
//decl g_1hat= -ones(N,1)+((2*exp(-(s_vy-muhat1)./dphiMLE1[] [i]))./(ones(N,1)
+exp(-(s_vy-muhat1)./dphiMLE1[] [i])));
//decl g_1til=-ones(N,1)+((2*exp(-(s_vy-muhat1)./dphiMLEnull1[] [i]))./(ones(N,1)
+exp(-(s_vy-muhat1)./dphiMLEnull1[] [i])));

//
//declaracao da matriz de informacao I hat
decl i1_1= (-1/(dphiMLE1[] [i].^2))* (X1hat'*delta_01*X1hat);
decl i2_1= (-1/(dphiMLE1[] [i].^2))* (X1hat'*delta_01*X2hat);
decl i3_1= (-1/(dphiMLE1[] [i].^2))*delta_01001 .* (X1hat'*ones(N,1));
decl i4_1= (-1/(dphiMLE1[] [i].^2))* (X2hat'*delta_01*X1hat);
decl i5_1= (-1/(dphiMLE1[] [i].^2))* (X2hat'*delta_01*X2hat);
decl i6_1= zeros(k,1);
decl i7_1= zeros(1,ib);
decl i8_1= zeros(1,k);
decl i9_1= (-N)/(dphiMLE1[] [i].^2).*(delta_01002 -1);
decl Ihat_1=(i1_1~i2_1~i3_1)|(i4_1~i5_1~i6_1)|(i7_1~i8_1~i9_1);
decl Ihatinv_1= invert(Ihat_1);
decl Ihatdet_1= determinant(Ihat_1);
decl IIhatdet_1= (Ihatdet_1).^(0.5);
//
//
//declaracao da matriz de informacao I til
decl i11_1= (-1/(dphiMLEnull1[] [i].^2))* (X1til'*delta_01*X1til);
decl i12_1= (-1/(dphiMLEnull1[] [i].^2))* (X1til'*delta_01*X2til);
decl i13_1= (-1/(dphiMLEnull1[] [i].^2))*delta_01001 .* (X1til'*ones(N,1));
decl i14_1= (-1/(dphiMLEnull1[] [i].^2))* (X2til'*delta_01*X1til);
decl i15_1= (-1/(dphiMLEnull1[] [i].^2))* (X2til'*delta_01*X2til);
decl i16_1= zeros(k,1);
decl i17_1= zeros(1,ib);
decl i18_1= zeros(1,k);

```

```

decl i19_1= (-N)/(dphiMLEnull1[] [i].^2).*(delta_01002 -1);
decl Itil_1=(i11_1~i12_1~i13_1)|(i14_1~i15_1~i16_1)|(i17_1~i18_1~i19_1);
decl Itilinv_1= invert(Itil_1);
decl I= Itilinv_1[0:ib-1][0:ib-1];
decl Itildet_1= determinant(Itil_1);
decl IItildet_1= (Itildet_1).^(0.5);

//declaracao do vetor escore U til
decl u11_1= (-1/dphiMLEnull1[] [i]) .*X1til'*g_1til;
decl u12_1= (-1/dphiMLEnull1[] [i]^2) .*(X2til'*g_1til);
decl u13_1= (-N/dphiMLEnull1[] [i]) - (1/dphiMLEnull1[] [i]^2)
.* (s_vy - mutil1)'*g_1til;
decl U_1= (u11_1)|(u12_1)|(u13_1);

//Declaracao das variaveis para a escore
decl vw= -(g_1til./s_vztil);
decl vs= -(delta_2^(-1/2).* g_1til);
decl mX2inv= invert(X2til'*X2til);
decl mC= mX2inv*X2til'*X1til;
decl mR= X1til- X2til*mC;
decl mRwR= invert(mR'*mR);

//Declaracao da estatistica Score
decl W= vs'*X1til*mRwR*X1til'*vs;
//
// Calculo da Estatisticas:
//Estatistica da razao de verossimilhanças usual
vLR_1[0][i] = 2*(dMaxLog_1 - dMaxLogNull_1);

//Estatistica da razao de verossimilhanças com correcao de Bartlett 1
vLRc[0][i] = (vLR_1[0][i])/c_b1;

//Estatistica da razao de verossimilhanças com correcao de Bartlett 2
vLRcc[0][i] = (vLR_1[0][i])*c_b2;

//Estatistica score
W_s[0][i]= W;
//
//Estatistica escore com correcao tipo Bartlett
Wc[0][i]= W*(1-(c+(b*W)+(a*(W.^2))));

// Estatistica escore com transformacao monotona
if(a==0 && b==0 )
W_tm[0][i]= W;

else
{
if(a==0 && b!=0)
W_tm[0][i]= ((1/(2*b))*exp(-c))*(1-exp(-(2*b*W_s[0][i])));

```

```

else
{
  if(a>0 && b!=0)
    W_tm[0][i]= ((sqrt(M_PI/(3*a)))*exp((b^2/(3*a))-c))*(probn((sqrt(6*a)
    *W_s[0][i])+(sqrt(2/(3*a))*b))-probn(sqrt(2/(3*a))*b));
  }
}

if (imod(i, cprint) == 0)
  eprint(i, "\n MC replications performed: ", (i/INREP)*100,
  " percent\n"); // print progress status
}

// Taxas de rejeicao: para \beta_1 como interesse
mRej_1[0][0] = (sumr((vLR_1 .> mEstQuant[0][0]))/(INREP))*100;
mRej_1[0][1] = (sumr((vLR_1 .> mEstQuant[0][1]))/(INREP))*100;
mRej_1[0][2] = (sumr((vLR_1 .> mEstQuant[0][2]))/(INREP))*100;
mRej_1[0][3] = (sumr((vLR_1 .> mEstQuant[0][3]))/(INREP))*100;

mRej_1[1][0] = (sumr((W_s .> mEstQuant[0][0]))/(INREP))*100;
mRej_1[1][1] = (sumr((W_s .> mEstQuant[0][1]))/(INREP))*100;
mRej_1[1][2] = (sumr((W_s .> mEstQuant[0][2]))/(INREP))*100;
mRej_1[1][3] = (sumr((W_s .> mEstQuant[0][3]))/(INREP))*100;

// mRej_1[2][0] = (sumr(vLR_1 .> vcvBoot[0][ ]))/(INREP))*100;
// mRej_1[2][1] = (sumr(vLR_1 .> vcvBoot[1][ ]))/(INREP))*100;
// mRej_1[2][2] = (sumr(vLR_1 .> vcvBoot[2][ ]))/(INREP))*100;
// mRej_1[2][3] = (sumr(vLR_1 .> vcvBoot[3][ ]))/(INREP))*100;
//
// mRej_1[3][0] = (sumr((W_s .> pcwBoot[0][ ]))/(INREP))*100;
// mRej_1[3][1] = (sumr((W_s .> pcwBoot[1][ ]))/(INREP))*100;
// mRej_1[3][2] = (sumr((W_s .> pcwBoot[2][ ]))/(INREP))*100;
// mRej_1[3][3] = (sumr((W_s .> pcwBoot[3][ ]))/(INREP))*100;

mRej_1[4][0] = (sumr((vLRc .> mEstQuant[0][0]))/(INREP))*100;
mRej_1[4][1] = (sumr((vLRc .> mEstQuant[0][1]))/(INREP))*100;
mRej_1[4][2] = (sumr((vLRc .> mEstQuant[0][2]))/(INREP))*100;
mRej_1[4][3] = (sumr((vLRc .> mEstQuant[0][3]))/(INREP))*100;

mRej_1[5][0] = (sumr((vLRcc .> mEstQuant[0][0]))/(INREP))*100;
mRej_1[5][1] = (sumr((vLRcc .> mEstQuant[0][1]))/(INREP))*100;
mRej_1[5][2] = (sumr((vLRcc .> mEstQuant[0][2]))/(INREP))*100;
mRej_1[5][3] = (sumr((vLRcc .> mEstQuant[0][3]))/(INREP))*100;
//
mRej_1[6][0] = (sumr((Wc .> mEstQuant[0][0]))/(INREP))*100;
mRej_1[6][1] = (sumr((Wc .> mEstQuant[0][1]))/(INREP))*100;
mRej_1[6][2] = (sumr((Wc .> mEstQuant[0][2]))/(INREP))*100;
mRej_1[6][3] = (sumr((Wc .> mEstQuant[0][3]))/(INREP))*100;

```

```

mRej_1[7][0] = (sumr((W_tm .> mEstQuant[0][0]))/(INREP))*100;
mRej_1[7][1] = (sumr((W_tm .> mEstQuant[0][1]))/(INREP))*100;
mRej_1[7][2] = (sumr((W_tm .> mEstQuant[0][2]))/(INREP))*100;
mRej_1[7][3] = (sumr((W_tm .> mEstQuant[0][3]))/(INREP))*100;
//

//

mMoments[0][0] = (ib);
mMoments[0][1] = (2*ib);
mMoments[1][0] = (meanr(vLR_1));
mMoments[1][1] = (varr(vLR_1));
mMoments[2][0] = (meanr(W_s));
mMoments[2][1] = (varr(W_s));
mMoments[3][0] = (meanr(vLRc));
mMoments[3][1] = (varr(vLRc));
mMoments[4][0] = (meanr(vLRcc));
mMoments[4][1] = (varr(vLRcc));
mMoments[5][0] = (meanr(Wc));
mMoments[5][1] = (varr(Wc));
mMoments[6][0] = (meanr(W_tm));
mMoments[6][1] = (varr(W_tm));

//

// Empirical CVs:
mEstQuant[1][]=(quantiler(vLR_1, <0.9, 0.95, 0.99, 0.995>));
mEstQuant[2][]=(quantiler(W_s, <0.9, 0.95, 0.99, 0.995>));
mEstQuant[3][]=(quantiler(vLRc, <0.9, 0.95, 0.99, 0.995>));
mEstQuant[4][]=(quantiler(vLRcc, <0.9, 0.95, 0.99, 0.995>));
mEstQuant[5][]=(quantiler(Wc, <0.9, 0.95, 0.99, 0.995>));
mEstQuant[6][]=(quantiler(W_tm, <0.9, 0.95, 0.99, 0.995>));

//// Calculando o poder do teste
// mRej_1[8][0] = (sumr((vLR_1 .> mEstQuant[1][0]))/(INREP))*100;
// mRej_1[8][1] = (sumr((vLR_1 .> mEstQuant[1][1]))/(INREP))*100;
// mRej_1[8][2] = (sumr((vLR_1 .> mEstQuant[1][2]))/(INREP))*100;
// mRej_1[8][3] = (sumr((vLR_1 .> mEstQuant[1][3]))/(INREP))*100;
//
//
// mRej_1[9][0] = (sumr((W_s .> mEstQuant[2][0]))/(INREP))*100;
// mRej_1[9][1] = (sumr((W_s .> mEstQuant[2][1]))/(INREP))*100;
// mRej_1[9][2] = (sumr((W_s .> mEstQuant[2][2]))/(INREP))*100;
// mRej_1[9][3] = (sumr((W_s .> mEstQuant[2][3]))/(INREP))*100;
//
// mRej_1[10][0] = (sumr((vLRc .> mEstQuant[3][0]))/(INREP))*100;
// mRej_1[10][1] = (sumr((vLRc .> mEstQuant[3][1]))/(INREP))*100;
// mRej_1[10][2] = (sumr((vLRc .> mEstQuant[3][2]))/(INREP))*100;
// mRej_1[10][3] = (sumr((vLRc .> mEstQuant[3][3]))/(INREP))*100;
//
// mRej_1[11][0] = (sumr((vLRcc .> mEstQuant[4][0]))/(INREP))*100;
// mRej_1[11][1] = (sumr((vLRcc .> mEstQuant[4][1]))/(INREP))*100;

```

```

// mRej_1[11][2] = (sumr((vLRcc .> mEstQuant[4][2]))/(INREP))*100;
// mRej_1[11][3] = (sumr((vLRcc .> mEstQuant[4][3]))/(INREP))*100;
//
// mRej_1[12][0] = (sumr((Wc .> mEstQuant[5][0]))/(INREP))*100;
// mRej_1[12][1] = (sumr((Wc .> mEstQuant[5][1]))/(INREP))*100;
// mRej_1[12][2] = (sumr((Wc .> mEstQuant[5][2]))/(INREP))*100;
// mRej_1[12][3] = (sumr((Wc .> mEstQuant[5][3]))/(INREP))*100;
//
// mRej_1[13][0] = (sumr((W_tm .> mEstQuant[6][0]))/(INREP))*100;
// mRej_1[13][1] = (sumr((W_tm .> mEstQuant[6][1]))/(INREP))*100;
// mRej_1[13][2] = (sumr((W_tm .> mEstQuant[6][2]))/(INREP))*100;
// mRej_1[13][3] = (sumr((W_tm .> mEstQuant[6][3]))/(INREP))*100;

fprintf(file,"*****\n");
fprintf(file,"*****\n");
fprintf(file,"Resultados quando \beta1 e o parametro de interesse \n.");
if (k_ML1)
    fprintf( file, "\nWARNING: ", k_ML1, " experiments failed (did not converge)
    under H1 .\n");
fprintfln(file,"0 numero de observacoes utilizadas foi:",rows(s_vy));
fprintfln(file, "\n *** Estimated size of Likelihood test ***\n", "%r", {"LR", "W_s",
"LRB", "W_sB", "LRc", "LRcc", "Wc", "W_tm"},
"%c", {"10%", "5%", "1%", "0.5%"}, "%10.3f", mRej_1);

fprintfln(file, "\nEmpirical and nominal critical values:" , "%r", {"chi-squared",
"LR", "W_s", "LRc", "LRcc", "Wc", "W_tm"}, "%c", {"10%", "5%", "1%", "0.5%"},
"%9.1f"
, mEstQuant);
//
fprintfln(file, "Momentos dos testes:" , "%r", {"chi-squared", "LR", "W_s",
"LR*", "LR**", "WC", "W_tm"}, "%c", {"Mean", "Variance"}, "%9.1f", mMoments);
//

gm= gm + 0.5; //fim do loop do jj
// fprintf(fpout, vLR_1');
// fprintf(fpout1, Wc');
// fprintf(fpout2, vLRc');
// fprintf(fpout3, vLRcc');
// fprintf(fpout4, Wc');
// fprintf(fpout5, W_tm');
}
}
fprintf(file, "\nExecution time: ", timespan(time), "\n");
fclose(file);

}

```

Programa da funções

```

/*****
Autor: Katya Silene Porto Rodrigues Braga
Programa: Funcoes.ox
*****/
#include <oxstd.h>
#include <oxfloat.h>

//Declaracao das funcoes para a Normal

//Funcao log-verossimilhanca da distribuicao Normal Maximizada
// em H1 - para o teste da razao de verossimilhaca tradicional
flognormalr1(const vp, const adfunc, const avscore, const amhess)
{
    decl dphi=vp[2];
    decl iq = 1 ;
    //declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(vp[0] .s_mX[] [3]);
    decl x5=vp[1]*s_mX[] [3].*exp(vp[0] .s_mX[] [3]);
    decl X1=x5;
    //fim da declaracao da matriz de derivadas (X til);

    decl vlpred=(vp[1]*exp(vp[0] .s_mX[] [3])); //2betas
    // decl vlpred=(vp[1]*exp(vp[0] .s_mX[] [3]))+vp[2]*s_mX[] [2]; // 3 betas
    // decl vlpred=(vp[1]*exp(vp[0] .s_mX[] [3]))+vp[2]*s_mX[] [2]
    //+vp[3]*s_mX[] [1]; // 4 betas
    // decl vlpred=(vp[1]*exp(vp[0] .s_mX[] [3]))+vp[2]*s_mX[] [2]+vp[3]*s_mX[] [1]
    //+vp[4]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

    adfunc[0]=-(N*(log(vp[2])))-(N*0.5*(log(M_2PI)))-((1.0/(2*(vp[2].^2)))
    .*((s_vy-mu)*(s_vy-mu)));
    // adfunc[0]=sumc(-log(vp[5])-0.5*log(M_2PI)-(0.5/(vp[5]^2)).*((s_vy-mu).^2));

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim de flognormalr1

//Funcao log-verossimilhanca da distribuicao Normal Maximizada
// em H0 - para o teste da razao de verossimilhaca tradicional
flognormalrnull1(const vp, const adfunc, const avscore, const amhess)
{
    decl dphi= vp[1];
    //declaracao da matriz de derivadas (X til)

```



```

decl x1=s_mX[] [0];
decl x2=s_mX[] [1];
decl x3=s_mX[] [2];
decl x4=exp(0 .*s_mX[] [3]);
decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]; // 3 betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]+vp[2]
//s_mX[] [1]; // 4 betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]+vp[2]*s_mX[] [1]
//+vp[3]*s_mX[] [0]; // 5 betas
decl mu=vlpred;

adfunc[0]=-(N*(log(vp[1])))-(N*0.5*(log(M_2PI)))-((1.0/(2*(vp[1].^2)))
//.*((s_vy-mu)*(s_vy-mu)));

if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
    return 0;
else
    return 1; // 1 indicates success
}
// Fim de flognormalrnull1

//Funcao log-verossimilhanca da distribuicao Normal Maximizada
// em Ho - para o teste da razao de verossimilhaca Bootstrap
flognormalboot(const vp, const adfunc, const avscore, const amhess)
{
    decl dphi= vp[4];
    //declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);
    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
    //fim da declaracao da matriz de derivadas (X til);

    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2betas
    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]; // 3 betas
    // decl vlpred=(vp[3]*exp(vp[0] .*s_mX[] [3]))+vp[1]*s_mX[] [2]
    //+vp[2]*s_mX[] [1]; // 4 betas
    decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]+vp[2]*s_mX[] [1]
    //+vp[3]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

    adfunc[0]=(-N*log(vp[4]))-(N*0.5*log(M_2PI)))-((1.0/(2*(vp[4].^2)))

```

```

        .*((s_vy-mu)*(s_vy-mu)));

if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
    return 0;
else
    return 1; // 1 indicates success
}
//Fim de flognormalboot

//Funcao log-verossimilhanca da distribuicao Normal Maximizada
// em H1 - para o teste da razao de verossimilhaca Bootstrap
flognormalrBOOT(const vp, const adfunc, const avscore, const amhess)
{
    decl dphi=vp[5];
    decl iq = 1 ;
//declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(vp[0] .*s_mX[] [3]);
    decl x5=vp[1]*s_mX[] [3].*exp(vp[0] .*s_mX[] [3]);
    decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])); //2betas
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])+vp[2]*s_mX[] [2]; // 3 betas
// decl vlpred=(vp[3]*exp(vp[0] .*s_mX[] [3])+vp[1]*s_mX[] [2]
//+vp[2]*s_mX[] [1]; // 4 betas
    decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])+vp[2]*s_mX[] [2]+vp[3]*s_mX[] [1]
//+vp[4]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

    adfunc[0]=(-N*log(vp[5]))-(N*0.5*log(M_2PI))-((1.0/(2*(vp[5].^2)))
//.*((s_vyboot-mu)*(s_vyboot-mu)));
// adfunc[0]=sumc(-log(vp[5])-0.5*log(M_2PI)-(0.5/(vp[5]^2)).*((s_vy-mu).^2));

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
// Fim de flognormalrBOOT

//Funcao log-verossimilhanca da distribuicao Normal Maximizada
// em Ho - para o teste da razao de verossimilhaca Bootstrap
flognormalrnullBOOT(const vp, const adfunc, const avscore, const amhess)
{
    decl dphi= vp[4];
//declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];

```

```

decl x2=s_mX[] [1];
decl x3=s_mX[] [2];
decl x4=exp(0 .*s_mX[] [3]);
decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]; // 3 betas
// decl vlpred=(vp[3]*exp(vp[0] .*s_mX[] [3]))+vp[1]*s_mX[] [2]
//+vp[2]*s_mX[] [1]; // 4 betas
decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]+vp[2]*s_mX[] [1]
//+vp[3]*s_mX[] [0]; // 5 betas
decl mu=vlpred;

adfunc[0]=(-N*log(vp[4]))-(N*0.5*log(M_2PI))-((1.0/(2*(vp[4].^2)))
.*(s_vyboot-mu)*(s_vyboot-mu));

if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
    return 0;
else
    return 1; // 1 indicates success
}
//Fim de flognormalrnullBOOT

//Fim das declaraes para a Normal
//*****//

//Declaraes das funes para a T-Student 1 grau de liberdade

// Funcao log-verossimilhanca da distribuicao T-Student com 1 grau de
//liberdade Maximizada
// em H1 - para o teste da razao de verossimilhaca tradicional
flogtstudent1r1(const vp, const adfunc, const avscore, const amhess)
{
    decl dphi=vp[2];
    decl iq = 1;
    //declaracao da matriz de derivadas (X til)
    // decl x1=s_mX[] [0];
    // decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(vp[0] .*s_mX[] [3]);
    decl x5=vp[1]*s_mX[] [3].*exp(vp[0] .*s_mX[] [3]);
    decl X1=x5;
    //fim da declaracao da matriz de derivadas (X til);

    decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])); //2betas
    // decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]; // 3 betas

```

```

// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]
//+vp[3]*s_mX[] [1]; // 4 betas
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]
//+vp[3]*s_mX[] [1]+vp[4]*s_mX[] [0]; // 5 betas
decl mu=vlpred;

adfunc[0]=-N*log(vp[2])-N*log(M_PI)-((log(ones(N,1)
+((s_vy-mu)./vp[2]).^2)))'*ones(N,1);
//print("adfunc",adfunc);

if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
    return 0;
else
    return 1; // 1 indicates success
}
//fim flogtstudent1r1

// Funcao log-verossimilhanca da distribuicao T-Student com 1 grau
//de liberdade Maximizada
// em H0 - para o teste da razao de verossimilhaca tradicional
flogtstudent1rnull1(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[1];
    decl iq = 1;
    //declaracao da matriz de derivadas (X til)
    // decl x1=s_mX[] [0];
    // decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);
    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
    //fim da declaracao da matriz de derivadas (X til);

    decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2betas
    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]; // 3 betas
    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
    //+vp[2]*s_mX[] [1]; // 4 betas
    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
    //+vp[2]*s_mX[] [1]+vp[3]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

    adfunc[0]=-N*log(vp[1])-N*log(M_PI)-((log(ones(N,1)+
    ((s_vy-mu)./vp[1]).^2)))'*ones(N,1);

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}

```

```

//Fim flogtstudent1rnull1

// Funcao log-verossimilhanca da distribuicao T-Student com 1 grau
//de liberdade Maximizada
// em H0 - para o teste da razao de verossimilhaca por Bootstrap
flogstudent1boot(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[2];
    decl iq = 1;
    //declaracao da matriz de derivadas (X til)
    // decl x1=s_mX[] [0];
    // decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);
    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
    //fim da declaracao da matriz de derivadas (X til);

    // decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])) //2betas
    decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]; // 3 betas
    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
    //+vp[2]*s_mX[] [1]; // 4 betas
    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
    //+vp[2]*s_mX[] [1]+vp[3]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

    adfunc[0]=-N*log(vp[2])-N*log(M_PI)-((log(ones(N,1)
    +((s_vy-mu)./vp[2]).^2)))'*ones(N,1);

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim flogtstudent1boot

// Funcao log-verossimilhanca da distribuicao T-Student com 1 grau de
//liberdade Maximizada
// em H1 - para o teste da razao de verossimilhaca por Bootstrap
flogstudent1BOOT(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[3];
    decl iq = 1;
    //declaracao da matriz de derivadas (X til)
    // decl x1=s_mX[] [0];
    // decl x2=s_mX[] [1];

```

```

    decl x3=s_mX[] [2];
    decl x4=exp(vp[0] .*s_mX[] [3]);
    decl x5=vp[1]*s_mX[] [3].*exp(vp[0] .*s_mX[] [3]);
    decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

    // decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])) //2betas
    decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]; // 3 betas
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]
//+vp[3]*s_mX[] [1]; // 4 betas
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]
//+vp[3]*s_mX[] [1]+vp[4]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

    adfunc[0]=-N*log(vp[3])-N*log(M_PI)-((log(ones(N,1)+
    ((s_vyboot-mu)./vp[3]).^2)))'*ones(N,1);
//print("adfunc",adfunc);

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim flogtstudent1BOOT

// Funcao log-verossimilhanca da distribuicao T-Student com 1 grau
//de liberdade Maximizada
// em H0 - para o teste da razao de verossimilhaca tradicional
flogstudentr1nullBOOT(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[2];
    decl iq = 1;
//declaracao da matriz de derivadas (X til)
// decl x1=s_mX[] [0];
// decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);
    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

    // decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])) //2betas
    decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]; // 3 betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
//+vp[2]*s_mX[] [1]; // 4 betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
//+vp[2]*s_mX[] [1]+vp[3]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

```

```

adfunc[0]=-N*log(vp[2])-N*log(M_PI)-((log(ones(N,1)+
((s_vyboot-mu)./vp[2]).^2)))*ones(N,1);
//print("adfunc",adfunc);

if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
    return 0;
else
    return 1; // 1 indicates success
}
//Fim flogtstudent1nullBOOT

//Fim das declaracoes para T-Student 1 grau de liberdade
//*****//

//Declaraes das funes para a T-Student 2 grau de liberdade

// Funcao log-verossimilhanca da distribuicao T-Student com 2 grau
//de liberdade Maximizada
// em H1 - para o teste da razao de verossimilhaca tradicional
flogtstudent2r1(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[2];
    decl iq = 1;
//declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(vp[0] .*s_mX[] [3]);
    decl x5=vp[1]*s_mX[] [3].*exp(vp[0] .*s_mX[] [3]);
    decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

    decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])); //2 beta
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])+vp[2].*s_mX[] [2]; //3 beta
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])+vp[2].*s_mX[] [2]
//+vp[3].*s_mX[] [1]; //4 beta
//decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])+vp[2].*s_mX[] [2]
//+vp[3].*s_mX[] [1]+vp[4].*s_mX[] [0]; //5 beta
    decl mu=vlpred;

    adfunc[0]=-N*log(vp[2])-3/2*((log(2*ones(N,1)+
((s_vyboot-mu)./vp[2]).^2))*ones(N,1));

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim do flogtstudent2r1

```

```

// Funcao log-verossimilhanca da distribuicao T-Student com 2 grau de
//liberdade Maximizada
// em H0 - para o teste da razao de verossimilhaca tradicional
flogtstudent2rnull1(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[1];
    decl iq = 1;
//declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);
    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

    decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2 beta
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]; //3 beta
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]
//+vp[2].*s_mX[] [1]; //4 beta
//decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]
//+vp[2].*s_mX[] [1]+vp[3].*s_mX[] [0]; //5 beta
    decl mu=vlpred;

    adfunc[0]=-N*log(vp[1])-3/2*(((log(2*ones(N,1)+
((s_vy-mu)./vp[1]).^2)))'*ones(N,1));

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim do flogtstudent2rnull1

// Funcao log-verossimilhanca da distribuicao T-Student com 2 grau de
//liberdade Maximizada
// em H0 - para o teste da razao de verossimilhaca por Bootstrap
flogtstudent2boot(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[4];
    decl iq = 1;
//declaracao da matriz de derivadas (X til)
//declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);

```



```

    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2 beta
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]; //3 beta
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]
//+vp[2].*s_mX[] [1]; //4 beta
decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]
//+vp[2].*s_mX[] [1]+vp[3].*s_mX[] [0]; //5 beta
    decl mu=vlpred;

    adfunc[0]=-N*log(vp[4])-3/2*(((log(2*ones(N,1)+
((s_vy-mu)./vp[4]).^2)))'*ones(N,1));

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim flogtstudent2boot

// Funcao log-verossimilhanca da distribuicao T-Student com 2 grau de
//liberdade Maximizada
// em H1 - para o teste da razao de verossimilhaca por Bootstrap
flogstudentr2B00T(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[5];
    decl iq = 1;
//declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(vp[0] .*s_mX[] [3]);
    decl x5=vp[1]*s_mX[] [3].*exp(vp[0] .*s_mX[] [3]);
    decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])); //2 beta
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2].*s_mX[] [2]; //3 beta
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2].*s_mX[] [2]
//+vp[3].*s_mX[] [1]; //4 beta
decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2].*s_mX[] [2]
//+vp[3].*s_mX[] [1]+vp[4].*s_mX[] [0]; //5 beta
    decl mu=vlpred;

    adfunc[0]=-N*log(vp[5])-3/2*(((log(2*ones(N,1)+
((s_vyboot-mu)./vp[5]).^2)))'*ones(N,1));

```

```

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim flogtstudent2B00T

// Funcao log-verossimilhanca da distribuicao T-Student com 2 grau de
//liberdade Maximizada
// em H0 - para o teste da razao de verossimilhaca tradicional
flogstudentr2nullB00T(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[4];
    decl iq = 1;
//declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);
    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2 beta
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]; //3 beta
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]
//+vp[2].*s_mX[] [1]; //4 beta
decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]
//+vp[2].*s_mX[] [1]+vp[3].*s_mX[] [0]; //5 beta
    decl mu=vlpred;

    adfunc[0]=-N*log(vp[4])-3/2*(((log(2*ones(N,1)+
((s_vyboot-mu)./vp[4]).^2)))'*ones(N,1));

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim flogtstudent2nullB00T

//Fim das declaracoes para T-Student 2 grau de liberdade
//*****//

// Funcao log-verossimilhanca da distribuicao T-Student com 4 grau de
//liberdade sob H1
flogtstudent4r1(const vp, const adfunc, const avscore, const amhess)

```

```

{

    decl dphi=vp[2];
    decl iq = 1;
    //declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(vp[0] .*s_mX[] [3]);
    decl x5=vp[1]*s_mX[] [3].*exp(vp[0] .*s_mX[] [3]);
    decl X1=x5;
    //fim da declaracao da matriz de derivadas (X til);

    decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])); //2 beta
    // decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2].*s_mX[] [2]; //3 beta
    // decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2].*s_mX[] [2]
    //+vp[3].*s_mX[] [1]; //4 beta
    //decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2].*s_mX[] [2]
    //+vp[3].*s_mX[] [1]+vp[4].*s_mX[] [0]; //5 beta
    decl mu=vlpred;

    adfunc[0]=-N*log(vp[2])+N*log(12)-5/2*((log(4*ones(N,1)+
    ((s_vy-mu)./vp[2]).^2)))'*ones(N,1);

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim flogtstudent4r1

// Funcao log-verossimilhanca da distribuicao T-Student com 4 grau de
//liberdade sob H0
flogtstudent4rnull1(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[1];
    decl iq = 1;
    //declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);
    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
    //fim da declaracao da matriz de derivadas (X til);

    decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2 beta

```

```

// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]; //3 beta
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]
//+vp[2].*s_mX[] [1]; //4 beta
//decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1].*s_mX[] [2]
//+vp[2].*s_mX[] [1]+vp[3].*s_mX[] [0]; //5 beta
decl mu=vlpred;

adfunc[0]=-N*log(vp[1])+N*log(12)-5/2*((log(4*ones(N,1)
+((s_vy-mu)./vp[1]).^2)))'*ones(N,1);

if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
    return 0;
else
    return 1; // 1 indicates success
}
//Fim flogtstudent4rnull1

// Funcao log-verossimilhanca da distribuicao T-Student com 4 grau de
//liberdade Maximizada
// em H0 - para o teste da razao de verossimilhaca por Bootstrap
flogtstudent4boot(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[3];
    decl iq = 1;
//declaracao da matriz de derivadas (X til)
// decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);
    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]; // 3 betas
    decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
    //+vp[2]*s_mX[] [1]; // 4 betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
//+vp[2]*s_mX[] [1]+vp[3]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

adfunc[0]=-N*log(vp[3])+N*log(12)-5/2*((log(4*ones(N,1)+
((s_vy-mu)./vp[3]).^2)))'*ones(N,1);

if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
    return 0;
else
    return 1; // 1 indicates success

```

```

}
//Fim flogtstudent4boot

// Funcao log-verossimilhanca da distribuicao T-Student com 4 grau de
//liberdade Maximizada
// em H1 - para o teste da razao de verossimilhaca por Bootstrap
flogstudentr4B00T(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[4];
    decl iq = 1;
    //declaracao da matriz de derivadas (X til)
    // decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(vp[0] .*s_mX[] [3]);
    decl x5=vp[1]*s_mX[] [3].*exp(vp[0] .*s_mX[] [3]);
    decl X1=x5;
    //fim da declaracao da matriz de derivadas (X til);

    // decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])); //2betas
    // decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]; // 3 betas
    decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]
    //+vp[3]*s_mX[] [1]; // 4 betas
    // decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]
    //+vp[3]*s_mX[] [1]+vp[4]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

    adfunc[0]=-N*log(vp[4])+N*log(12)-5/2*((log(4*ones(N,1)+
    ((s_vyboot-mu)./vp[4]).^2)))'*ones(N,1);
    //print("adfunc",adfunc);

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim flogtstudent4B00T

// Funcao log-verossimilhanca da distribuicao T-Student com 4 grau de
//liberdade Maximizada
// em H0 - para o teste da razao de verossimilhaca tradicional
flogstudentr4nullB00T(const vp, const adfunc, const avscore, const amhess)
{

    decl dphi=vp[3];
    decl iq = 1;
    //declaracao da matriz de derivadas (X til)
    // decl x1=s_mX[] [0];

```

```

    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);
    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]; // 3 betas
    decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
    //+vp[2]*s_mX[] [1]; // 4 betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
//+vp[2]*s_mX[] [1]+vp[3]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

    adfunc[0]=-N*log(vp[3])+N*log(12)-5/2*((log(4*ones(N,1)+
    ((s_vyboot-mu)./vp[3]).^2)))'*ones(N,1);
//print("adfunc",adfunc);

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim flogtstudent4nullB00T

//Fim das declaracoes para T-Student 4 grau de liberdade
//*****//
//Declaracao das funcoes para a logistica

//Funcao log-verossimilhanca da distribuicao logistica Maximizada
// em H1 - para o teste da razao de verossimilhaca tradicional
flogisticar1(const vp, const adfunc, const avscore, const amhess)
{
    decl dphi=vp[2];
    decl iq = 1 ;
//declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(vp[0] .*s_mX[] [3]);
    decl x5=vp[1]*s_mX[] [3].*exp(vp[0] .*s_mX[] [3]);
    decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

    decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])); //2betas
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]; // 3 betas
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))
//+vp[2]*s_mX[] [2]+vp[3]*s_mX[] [1]; // 4 betas
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]

```

```

//+vp[3]*s_mX[] [1]+vp[4]*s_mX[] [0]; // 5 betas
decl mu=vlpred;

adfunc[0]=(-N*log(vp[2]))+(((-(s_vy-mu)/vp[2])-2*(log(ones(N,1)
+exp(-(s_vy-mu)/vp[2])))))*ones(N,1));

if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
    return 0;
else
    return 1; // 1 indicates success
}
//Fim de flogisticar1

//Funcao log-verossimilhanca da distribuicao logistica Maximizada
// em H0 - para o teste da razao de verossimilhaca tradicional
flogisticarnull1(const vp, const adfunc, const avscore, const amhess)
{
    decl dphi= vp[1];
    //declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);
    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
    //fim da declaracao da matriz de derivadas (X til);

    decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2betas
    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]; // 3 betas
    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
    //+vp[2]*s_mX[] [1]; // 4 betas
    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
    //+vp[2]*s_mX[] [1]+vp[3]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

    adfunc[0]=(-N*log(vp[1]))+(((-(s_vy-mu)/vp[1])-2*(log(ones(N,1)
+exp(-(s_vy-mu)/vp[1])))))*ones(N,1));

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
// Fim de flogisticarnull1

//Funcao log-verossimilhanca da distribuicao logistica Maximizada
// em Ho - para o teste da razao de verossimilhaca Bootstrap
flogisticaboot(const vp, const adfunc, const avscore, const amhess)
{
    decl dphi= vp[4];

```

```

//declaracao da matriz de derivadas (X til)
decl x1=s_mX[] [0];
decl x2=s_mX[] [1];
decl x3=s_mX[] [2];
decl x4=exp(0 .*s_mX[] [3]);
decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]; // 3 betas
// decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
//+vp[2]*s_mX[] [1]; // 4 betas
decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
//+vp[2]*s_mX[] [1]+vp[3]*s_mX[] [0]; // 5 betas
decl mu=vlpred;

adfunc[0]=(-N*log(vp[4]))+(((-(s_vy-mu)/vp[4])-2*(log(ones(N,1)
+exp(-(s_vy-mu)/vp[4]))))'*ones(N,1));

if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
    return 0;
else
    return 1; // 1 indicates success
}
//Fim de flogisticaboot

//Funcao log-verossimilhanca da distribuicao Normal Maximizada
// em H1 - para o teste da razao de verossimilhaca Bootstrap
flogisticarBOOT(const vp, const adfunc, const avscore, const amhess)
{
    decl dphi=vp[5];
    decl iq = 1 ;
//declaracao da matriz de derivadas (X til)
decl x1=s_mX[] [0];
decl x2=s_mX[] [1];
decl x3=s_mX[] [2];
decl x4=exp(vp[0] .*s_mX[] [3]);
decl x5=vp[1]*s_mX[] [3].*exp(vp[0] .*s_mX[] [3]);
decl X1=x5;
//fim da declaracao da matriz de derivadas (X til);

// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3])); //2betas
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]; // 3 betas
// decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]
//+vp[3]*s_mX[] [1]; // 4 betas
decl vlpred=(vp[1]*exp(vp[0] .*s_mX[] [3]))+vp[2]*s_mX[] [2]
//+vp[3]*s_mX[] [1]+vp[4]*s_mX[] [0]; // 5 betas
decl mu=vlpred;

```



```

adfunc[0]=(-N*log(vp[5]))+(((-(s_vyboot-mu)/vp[5])-2*(log(ones(N,1)
+exp(-(s_vyboot-mu)/vp[5])))))*ones(N,1));
// adfunc[0]=sumc(-log(vp[5])-0.5*log(M_2PI)-(0.5/(vp[5]^2)).*((s_vy-mu).^2));

if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
    return 0;
else
    return 1; // 1 indicates success
}
// Fim de flogisticarBOOT

//Funcao log-verossimilhanca da distribuicao logistica Maximizada
// em Ho - para o teste da razao de verossimilhaca Bootstrap
flogisticarnullBOOT(const vp, const adfunc, const avscore, const amhess)
{
    decl dphi= vp[4];
    //declaracao da matriz de derivadas (X til)
    decl x1=s_mX[] [0];
    decl x2=s_mX[] [1];
    decl x3=s_mX[] [2];
    decl x4=exp(0 .*s_mX[] [3]);
    decl x5=vp[0]*s_mX[] [3].*exp(0 .*s_mX[] [3]);
    decl X1=x5;
    //fim da declaracao da matriz de derivadas (X til);

    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3])); //2betas
    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]; // 3 betas
    // decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
    //+vp[2]*s_mX[] [1]; // 4 betas
    decl vlpred=(vp[0]*exp(0 .*s_mX[] [3]))+vp[1]*s_mX[] [2]
    //+vp[2]*s_mX[] [1]+vp[3]*s_mX[] [0]; // 5 betas
    decl mu=vlpred;

    adfunc[0]=(-N*log(vp[4]))+(((-(s_vyboot-mu)/vp[4])-2*(log(ones(N,1)
+exp(-(s_vyboot-mu)/vp[4])))))*ones(N,1));

    if(isnan(adfunc[0]) | isdotinf(adfunc[0]))
        return 0;
    else
        return 1; // 1 indicates success
}
//Fim de flogisticarnullBOOT

```

```

/* Funcao para gerar o vetor de reamostragem Bootstrap em sua forma paramtrica */
fSample(const func, const vp, const cn)
{
    decl x1=ones(N,1);
    decl ci;
    decl v_Y = zeros(cn,1);
    decl vlpred1= zeros(cn,1);
    decl mu1 = zeros(cn,1);
    decl dphi1= 0;
    for (ci=0; ci < cn; ci++)
    {
        //      vlpred1[ci]=(vp[1]*exp(vp[0] .*s_mX[ci][3])); //2 beta
        // vlpred1[ci]=(vp[1]*exp(vp[0] .*s_mX[ci][3]))+(vp[2] .*s_mX[ci][2]); //3 beta
        //      vlpred1[ci]=(vp[1]*exp(vp[0] .*s_mX[ci][3]))+(vp[2] .*s_mX[ci][2])
        //+(vp[3] .*s_mX[ci][1]); //4 betas
        vlpred1[ci]=(vp[1]*exp(vp[0] .*s_mX[ci][3]))+(vp[2] .*s_mX[ci][2])
        //+(vp[3] .*s_mX[ci][1])+(vp[4].*s_mX[ci][0]); //5 betas
        mu1[ci]=vlpred1[ci];
        // Para gerar numeros normais
        //      v_Y[ci] = mu1[ci] + vp[2]*rann(1,1); // 2 betas
        //      v_Y[ci] = mu1[ci] + vp[3]*rann(1,1); // 3 betas
        //v_Y[ci] = mu1[ci] + vp[4]*rann(1,1); // 4 betas
        //v_Y[ci] = mu1[ci] + vp[5]*rann(1,1); // 5 betas

        //Para gerar numeros t-student 1
        //v_Y[ci] = mu1[ci] + vp[2]*rant(1,1,1); // 2 betas
        //v_Y[ci] = mu1[ci] + vp[3]*rant(1,1,1); // 3 betas
        //v_Y[ci] = mu1[ci] + vp[4]*rant(1,1,1); // 4 betas
        //v_Y[ci] = mu1[ci] + vp[5]*rant(1,1,1); // 5 betas

        //Para gerar numeros t-student 2
        //v_Y[ci] = mu1[ci] + vp[1]*rant(1,1,2); // 1 betas
        //v_Y[ci] = mu1[ci] + vp[2]*rant(1,1,2); // 2 betas
        //v_Y[ci] = mu1[ci] + vp[3]*rant(1,1,2); // 3 betas
        //v_Y[ci] = mu1[ci] + vp[4]*rant(1,1,2); // 4 betas
        //v_Y[ci] = mu1[ci] + vp[5]*rant(1,1,2); // 5 betas

        //Para gerar numeros t-student 4
        //v_Y[ci] = mu1[ci] + vp[2]*rant(1,1,4); // 2 betas
        //v_Y[ci] = mu1[ci] + vp[3]*rant(1,1,4); // 3 betas
        //v_Y[ci] = mu1[ci] + vp[4]*rant(1,1,4); // 4 betas
        //v_Y[ci] = mu1[ci] + vp[5]*rant(1,1,4); // 5 betas

        //Para gerar numeros Logistica
        //v_Y[ci] = mu1[ci] + vp[2]*ranlogistic(1,1); // 2 betas
        //v_Y[ci] = mu1[ci] + vp[3]*ranlogistic(1,1); // 3 betas
        //v_Y[ci] = mu1[ci] + vp[4]*ranlogistic(1,1); // 4 betas
        v_Y[ci] = mu1[ci] + vp[5]*ranlogistic(1,1); // 5 betas

    }
    func[0] = v_Y; //amostra bootstrap
    return 1;
}

```

```

}
//Fim de fSample

/* Funcao para estimar a distribuicao empirica por Bootstrap do teste
da razao de verossimilhanças tradicional */
fBoot(const vpar, const tn, const nB)
{
    decl ib= 1; // numeros de betas de interesse
    decl k= 4; //numeros de betas de perturbacao
    decl cj, irrBOOT, irrnullBOOT, dfuncrBOOT, dfuncrnullBOOT,
    amostra, dMaxLogBOOT, dMaxLogNullBOOT ;
    decl mrazaoverossimilhancaboot= zeros(1,nB);
    decl vprBOOT = zeros(ib+k+1, 1);
    decl vprnullBOOT = zeros(k+1,1);
    decl k_MLBOOT = 0;
    decl vbeta1=zeros(ib+k,1);
    decl phi1=1;
    //Maximizacao das funcoes por BOOSTRAP
    for (cj=0; cj < NBOOT; ++cj) //inicio do loop bootstrap
    {

        //Selecionando amostra
        fSample(&amostra,vpar,N);
        s_vyboot= amostra;

        //Chute inicial
        vprBOOT=vbeta1|phi1;
        vprnullBOOT=0|0|0|phi1;
        //Maximizacoes
        //// Maximizacoes para Normal
        // irrBOOT = MaxBFGS(flognormalrBOOT, &vprBOOT, &dfuncrBOOT, 0,TRUE);
        // if (irrBOOT != MAX_CONV && irrBOOT != MAX_WEAK_CONV){
        // ++k_MLBOOT; // check convergence (H1)
        // --cj;
        // continue;
        // }
        //
        // irrnullBOOT = MaxBFGS(flognormalrnullBOOT, &vprnullBOOT,
        //&dfuncrnullBOOT, 0,TRUE);
        // if (irrnullBOOT != MAX_CONV && irrnullBOOT != MAX_WEAK_CONV){
        // ++k_MLBOOT; // check convergence (H0) Lmpr
        // --cj;
        // continue
        // }

        //Maximizacoes para t-student 1
        //irrBOOT = MaxBFGS(flogstudentrBOOT, &vprBOOT, &dfuncrBOOT, 0,TRUE);
        // if (irrBOOT != MAX_CONV && irrBOOT != MAX_WEAK_CONV){
        // ++k_MLBOOT; // check convergence (H1)
        // --cj;
        // continue;
        // }
        //

```

```

// irrnullBOOT = MaxBFGS(flogstudentr1nullBOOT, &vprnullBOOT,
//&dfuncrnullBOOT, 0,TRUE);
// if (irrnullBOOT != MAX_CONV && irrnullBOOT != MAX_WEAK_CONV){
// ++k_MLBOOT;          // check convergence (H0) Lmpr
// --cj;
// continue
// }
//
////Maximizacoes para t-student 2
//irrBOOT = MaxBFGS(flogstudentr2BOOT, &vprBOOT, &dfuncrBOOT, 0,TRUE);
// if (irrBOOT != MAX_CONV && irrBOOT != MAX_WEAK_CONV){
// ++k_MLBOOT;          // check convergence (H1)
// --cj;
// continue;
// }
//
// irrnullBOOT = MaxBFGS(flogstudentr2nullBOOT, &vprnullBOOT,
//&dfuncrnullBOOT, 0,TRUE);
// if (irrnullBOOT != MAX_CONV && irrnullBOOT != MAX_WEAK_CONV){
// ++k_MLBOOT;          // check convergence (H0) Lmpr
// --cj;
// continue
// }

////Maximizacoes para t-student 4
//irrBOOT = MaxBFGS(flogstudentr4BOOT, &vprBOOT, &dfuncrBOOT, 0,TRUE);
// if (irrBOOT != MAX_CONV && irrBOOT != MAX_WEAK_CONV){
// ++k_MLBOOT;          // check convergence (H1)
// --cj;
// continue;
// }
//
// irrnullBOOT = MaxBFGS(flogstudentr4nullBOOT, &vprnullBOOT,
//&dfuncrnullBOOT, 0,TRUE);
// if (irrnullBOOT != MAX_CONV && irrnullBOOT != MAX_WEAK_CONV){
// ++k_MLBOOT;          // check convergence (H0) Lmpr
// --cj;
// continue
// }

///Maximizacoes para Logistica
irrBOOT = MaxBFGS(flogisticarBOOT, &vprBOOT, &dfuncrBOOT, 0,TRUE);
if (irrBOOT != MAX_CONV && irrBOOT != MAX_WEAK_CONV){
++k_MLBOOT;          // check convergence (H1)
--cj;
continue;
}

irrnullBOOT = MaxBFGS(flogisticarnullBOOT, &vprnullBOOT,
&dfuncrnullBOOT, 0,TRUE);
if (irrnullBOOT != MAX_CONV && irrnullBOOT != MAX_WEAK_CONV){
++k_MLBOOT;          // check convergence (H0) Lmpr

```

```

--cj;
continue
}

//Declaracao de variaveis para o teste por Bootstrap
dMaxLogBOOT=dfuncrBOOT;
dMaxLogNullBOOT=dfuncrnullBOOT;

//Estatistica de teste de verossimilhanca por Bootstrap
mrazaoverossimilhancaboot[0][cj]= 2*(dMaxLogBOOT - dMaxLogNullBOOT);
} //final do bootstrap
return (mrazaoverossimilhancaboot) ;
}
//Fim de fBoot

/* Funcao para estimar a distribuicao empirica por Bootstrap do teste
escore tradicional */
WBoot(const vpar, const N, const nB)
{
    decl x1=ones(N,1);
    decl k=3; //numeros de betas de perturbacao
    decl cw, irrnullBOOT,dfuncrnullBOOT, amostra ;
    decl mescoreboot= zeros(1,nB);
    decl k_MLBOOT= 0;
    decl vprnullBOOT = zeros(k+1,1);
    decl phi1=1;
    decl betaMLE_H0= zeros(k+1,nB);
    decl phiMLE_H0=zeros(1,nB);
    decl fboot= 0; //esperanca de  $g^2 \cdot z_1$  para Normal,
    //t-student v=1,2,4 , logistica
    // decl gboot= -diag(1); //esperanca de  $g^2$  para Normal
    // decl gboot= -1/2; //esperanca de  $g^2$  para t-student v=1
    // decl gboot= -3/5; //esperanca de  $g^2$  para t-student v=2
    // decl gboot= -5/7; //esperanca de  $g^2$  para t-student v=4
    decl gboot= -1/3; //esperanca de  $g^2$  para Logistica
    // decl eboot=-1; //esperanca de  $g^2 \cdot z_1^2$  para Normal
    // decl eboot= 1/2; //esperanca de  $g^2 \cdot z_1^2$  para T-student v=1
    // decl eboot= 1/5; //esperanca de  $g^2 \cdot z_1^2$  para T-student v=2
    // decl eboot=-1/7; //esperanca de  $g^2 \cdot z_1^2$ 
    //para T-student v=4
    decl eboot=-0.42996; //esperanca de  $g^2 \cdot z_1^2$  para
    //Logistica
    // decl delta_2boot= 1/2; //para a distribuicao nT-Student v=1
    //decl delta_2boot= 3/5; //para a distribuicao T-Student v=2
    // decl delta_2boot= 5/7; //para a distribuicao T-Student v=4
    decl delta_2boot= 1/3; //para a distribuicao Logistica
    decl kboot= 3;
    decl ibboot= 1;
    //Maximizacao das funcoes por BOOTSTRAP
    for (cw=0; cw < NBOOT; ++cw) //inicio do loop bootstrap
    {

```

```

//Selecionando amostra
fSample(&amostra,vpar,N);
s_vyboot= amostra;

//Chute inicial
vprnullBOOT=0|0|0|phi1;
//Maximizacoes
////Maximizacoes para Normal
// irrnullBOOT = MaxBFGS(flognormalrnullBOOT, &vprnullBOOT,
//&dfuncrnullBOOT, 0,TRUE);
// if (irrnullBOOT != MAX_CONV && irrnullBOOT != MAX_WEAK_CONV){
// ++k_MLBOOT;          // check convergence (H0) Lmpr
// --cw;
// continue
// }

////Maximizacoes para t-student 1
//irrnullBOOT = MaxBFGS(flogstudentr1nullBOOT, &vprnullBOOT,
//&dfuncrnullBOOT, 0,TRUE);
// if (irrnullBOOT != MAX_CONV && irrnullBOOT != MAX_WEAK_CONV){
// ++k_MLBOOT;          // check convergence (H0) Lmpr
// --cw;
// continue
// }

////Maximizacoes para t-student 2
//irrnullBOOT = MaxBFGS(flogstudentr2nullBOOT, &vprnullBOOT,
//&dfuncrnullBOOT, 0,TRUE);
// if (irrnullBOOT != MAX_CONV && irrnullBOOT != MAX_WEAK_CONV){
// ++k_MLBOOT;          // check convergence (H0) Lmpr
// --cw;
// continue
// }

////Maximizacoes para t-student 4
//irrnullBOOT = MaxBFGS(flogstudentr4nullBOOT, &vprnullBOOT,
//&dfuncrnullBOOT, 0,TRUE);
// if (irrnullBOOT != MAX_CONV && irrnullBOOT != MAX_WEAK_CONV){
// ++k_MLBOOT;          // check convergence (H0) Lmpr
// --cw;
// continue
// }
//
//Maximizacoes para Logistica
irrnullBOOT = MaxBFGS(flogisticarnullBOOT, &vprnullBOOT,
&dfuncrnullBOOT, 0,TRUE);
if (irrnullBOOT != MAX_CONV && irrnullBOOT != MAX_WEAK_CONV){
++k_MLBOOT;          // check convergence (H0) Lmpr
--cw;
continue
}

```

```

    betaMLE_H0[] [cw] = 0|vprnullBOOT[0:2];
    phiMLE_H0[] [cw]=vprnullBOOT[3];
// print("betaMLE_H0[]:",betaMLE_H0[0]);
//Declaracao da matriz X1til
// //declaracao da matriz X til
// decl x1tilboot=s_mX[] [0];
    decl x2tilboot=s_mX[] [1];
    decl x3tilboot=s_mX[] [2];
    decl x4tilboot=exp(betaMLE_H0[0] [cw] .*s_mX[] [3]);
    decl x5tilboot=betaMLE_H0[1] [cw]*s_mX[] [3].*exp(betaMLE_H0[0] [cw] .*s_mX[] [3]);
    decl X1tilboot= x5tilboot;
    decl X2tilboot= x4tilboot~x3tilboot~x2tilboot;

// declaracao de mutil
//decl vlpredtil1boot=betaMLE_H0[1] [cw]*exp(betaMLE_H0[0] [cw]
//.*s_mX[] [3]); //2 beta
// decl vlpredtil1boot=betaMLE_H0[1] [cw]*exp(betaMLE_H0[0] [cw] .*s_mX[] [3])
//+(betaMLE_H0[2] [cw] .*s_mX[] [2]); //3 betas
    decl vlpredtil1boot=betaMLE_H0[1] [cw]*exp(betaMLE_H0[0] [cw] .*s_mX[] [3])
    +(betaMLE_H0[2] [cw] .*s_mX[] [2])
    +(betaMLE_H0[3] [cw] .*s_mX[] [1]); // 4 betas
// decl vlpredtil1boot=betaMLE_H0[1] [cw]*exp(betaMLE_H0[0] [cw] .*s_mX[] [3])
//+(betaMLE_H0[2] [cw] .*s_mX[] [2])+(betaMLE_H0[3] [cw] .*s_mX[] [1])
//+(betaMLE_H0[4] [cw] .*s_mX[] [0]); // 5 betas

    decl muboot=vlpredtil1boot;
    decl s_vztilboot= (s_vyboot - muboot)/phiMLE_H0[] [cw];
// decl g_1boot= -((s_vyboot-muboot)./phiMLE_H0[] [cw]); // para Normal
// decl g_1boot=-2*((s_vyboot-muboot)./phiMLE_H0[] [cw])./(ones(N,1)
//+((s_vyboot-muboot)./phiMLE_H0[] [cw]).^2); //g^1 sob H0 da T-student v=1
// decl g_1boot=-3*((s_vyboot-muboot)./phiMLE_H0[] [cw])./(2*ones(N,1)+
//((s_vyboot-muboot)./phiMLE_H0[] [cw]).^2); //g^1 sob H0 da T-student v=2
// decl g_1boot=-5*((s_vyboot-muboot)./phiMLE_H0[] [cw])./(4*ones(N,1)+
//((s_vyboot-muboot)./phiMLE_H0[] [cw]).^2); //g^1 sob H0 da T-student v=4
    decl g_1boot= -ones(N,1)+((2*exp(-(s_vyboot-muboot)./phiMLE_H0[] [cw]))
./(ones(N,1)+exp(-(s_vyboot-muboot)./phiMLE_H0[] [cw])));
//g^1 sob H0 da logistica

//declaracao da matriz de informacao I til
    decl i1boot= (-1/(phiMLE_H0[] [cw].^2))* (X1tilboot'*gboot*X1tilboot);
    decl i2boot= (-1/(phiMLE_H0[] [cw].^2))* (X1tilboot'*gboot*X2tilboot);
    decl i3boot= (-1/(phiMLE_H0[] [cw].^2))*fboot .* (X1tilboot'*ones(N,1));
    decl i4boot= (-1/(phiMLE_H0[] [cw].^2))* (X2tilboot'*gboot*X1tilboot);
    decl i5boot= (-1/(phiMLE_H0[] [cw].^2))* (X2tilboot'*gboot*X2tilboot);
    decl i6boot= zeros(kboot,1);
    decl i7boot= zeros(1,iboot);
    decl i8boot= zeros(1,kboot);
    decl i9boot= (-N)/(phiMLE_H0[] [cw].^2).*(eboot -1);
    decl Iboot_1=(i1boot~i2boot~i3boot)|(i4boot~i5boot~i6boot)
    |(i7boot~i8boot~i9boot);

```

```

    decl Itilinvboot= invert(Iboot_1);
    decl Iboot= Itilinvboot[0][0];

//Declaracao das variaveis para a escore
decl vwboot= -(g_1boot./s_vztilboot);
decl vsboot= -(delta_2boot^(-1/2).* g_1boot);
decl mX2invboot= invert(X2tilboot'*X2tilboot);
decl mCboot= mX2invboot*X2tilboot'*X1tilboot;
decl mRboot= X1tilboot- X2tilboot*mCboot;
decl mRwRboot= invert(mRboot'*mRboot);

//Declaracao da estatistica Score
    decl Wboot= vsboot'*X1tilboot*mRwRboot*X1tilboot'*vsboot;

//Estatistica de teste de verossimilhanca por Bootstrap
mescoreboot[0][cw]= Wboot;
} //final do bootstrap
return (mescoreboot) ;
}

```

Programa de ligação

```

/* Autor: Katya Silene Porto
   Programa: Programa de ligacao - ligacao.h */

// Declaracao das funcoes usadas nos testes verossimilhanca usual e skovgaard
//Funcao log-verossimilhanca da distribuicao Normal sob H1
flognormalr1(const vp, const adfunc, const avscore, const amhess);

//Funcao log-verossimilhanca da distribuicao Normal sob H0
flognormalrnull1(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao T-Student com 1 grau de liberdade
//sob H1
flogtstudent1r1(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao T-Student com 1 grau de liberdade
//sob H0
flogtstudent1rnull1(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao T-Student com 2 grau de liberdade
//sob H1
flogtstudent2r1(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao T-Student com 2 grau de liberdade
//sob H0
flogtstudent2rnull1(const vp, const adfunc, const avscore, const amhess);

```



```

// Funcao log-verossimilhanca da distribuicao T-Student com 4 grau de liberdade
//sob H1
flogtstudent4r1(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao T-Student com 4 grau de liberdade
//sob H0
flogtstudent4rnull1(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao Logistica sob H1
flogisticar1(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao Logistica sob H0
flogisticarnull1(const vp, const adfunc, const avscore, const amhess);
//Fim das declaracoes


// Declaracao das funcoes usadas nos testes verossimilhanca bootstrap
//Funcao para gerar a amostra
fSample(const func, const vP, const cn);

//Funcao de loop de bootstrap para teste da razao de verossimilhanca tradicional
fBoot(const vpar, const tn, const nB);

//Funcao de loop de bootstrap para teste escore tradicional
WBoot(const vpar, const N, const nB);

//Funcao log-verossimilhanca da distribuicao Normal sob H1
flognormalrBOOT(const vp, const adfunc, const avscore, const amhess);

//Funcao log-verossimilhanca da distribuicao Normal sob H0
flognormalboot(const vp, const adfunc, const avscore, const amhess);

//Funcao log-verossimilhanca da distribuicao Normal sob H0
flognormalrnullBOOT(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao T-Student com 1 grau de liberdade
//sob H0
flogstudent1boot(const vp, const adfunc, const avscore, const amhess);

//Funcao log-verossimilhanca da distribuicao T-Student com 1 grau de liberdade
//sob H1
flogstudentr1BOOT(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao T-Student com 1 grau de liberdade
//sob H0
flogstudentr1nullBOOT(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao T-Student com 2 grau de liberdade
//sob H0
flogtstudent2boot(const vp, const adfunc, const avscore, const amhess);

```

```

//Funcao log-verossimilhanca da distribuicao T-Student com 2 grau de liberdade
//sob H1
flogstudentr2BOOT(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao T-Student com 2 grau de liberdade
//sob H0
flogstudentr4nullBOOT(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao T-Student com 4 grau de liberdade
//sob H0
flogstudent4boot(const vp, const adfunc, const avscore, const amhess);

//Funcao log-verossimilhanca da distribuicao T-Student com 4 grau de liberdade
//sob H1
flogstudentr4BOOT(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao T-Student com 4 grau de liberdade
//sob H0
flogstudentr4nullBOOT(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao Logistica sob H0
flogisticaboot(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao Logistica sob H1
flogisticarBOOT(const vp, const adfunc, const avscore, const amhess);

// Funcao log-verossimilhanca da distribuicao Logistica sob H0
flogisticarnullBOOT(const vp, const adfunc, const avscore, const amhess);


//Fim das declaracoes

```

Referências

- [1] Anderson, T. W., Fang, K. T., (1987). Cochran's theorem for elliptically contoured distributions. *Sankhya A*, **49**, 305-315.
- [2] 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.
- [3] Bartlett, M.S (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London A*, **160**, 268-282.
- [4] Berkane, M., Bentler, P.M. (1986). Moments of elliptically distributed random variates. *Statistics and Probability Letters*, **4**, 333-335.
- [5] Buse, A. (1982). The likelihood ratio, Wald, and Lagrange multiplier tests: an expository note. *The American Statistician*, **36**, 153-157.
- [6] Cambanis, S., Huang, S., Simons, G. (1981). On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis*, **11**, 368-385.
- [7] Chesher, A., Smith, R.. (1995). Bartlett corrections to likelihood ratio tests. *Biometrika*, **82**, 433-436.
- [8] Chmielewski, M.A. (1981). Elliptically symmetric distributions: a review and bibliography. *International Statistical Review*, **49**, 67-74.
- [9] Cordeiro, G.M. (1983). Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistical Society B*, **45**, 404-413.
- [10] Cordeiro, G.M. (1987). On the corrections to the likelihood ratio statistics. *Biometrika*, **74**, 265-274.
- [11] Cordeiro, G.M. (1999). *Introdução à Teoria Assintótica*. IMPA - Instituto de Matemática Pura e Aplicada, Rio de Janeiro.
- [12] Cordeiro, G.M. (2004). Corrected likelihood ratio tests in symmetric nonlinear regression models. *Journal of Statistical Computation and Simulation*, **74**, 600-620.
- [13] Cordeiro, G.M., Cribari-Neto, F., Aubin, E.C.Q., Ferrari, S.L.P. (1995). Bartlett correction for one-parameter exponential family models. *Journal of Statistical Computation and Simulation*, **53**, 211-231.

- [14] Cordeiro, G.M., Ferrari, S.L.P (1991). A modified score statistic having chi-squared distribution to order n^{-1} . *Biometrika*, **78**, 573-582.
- [15] Cordeiro, G. M., Ferrari, S.L.P., Cysneiros, A.H. (1998). A formula to improve score test statistics. *Journal of Statistical Computation and Simulation*, **61**, 123-136.
- [16] Cordeiro, G. M., Ferrari, S.L.P., Paula, G.A. (1993). Improved score tests for generalized linear models. *Journal of Royal Statistical Society B*, **55**, 661-674.
- [17] Cordeiro, G. M., Ferrari, S.L.P., Uribe-Opazo, M., Vasconcellos, K. (2000). Corrected maximum likelihood estimation in a class of symmetric nonlinear regression models. *Statistic and Probability Letters*.
- [18] Cordeiro, G.M., Paula, G.A. (1989). Improved likelihood ratio statistic for exponential family nonlinear models. *Biometrika*, **76**, 93-100.
- [19] Cox, D.R., Hinkley, D.V (1974). *Theoretical Statistics*. New York: Wiley.
- [20] Cox, D.R., Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*, **49**, 1-39.
- [21] Cribari-Neto, F., Cordeiro, G.M. (1997). On Bartlett and Bartlett-type corrections. *Econometric Reviews*, **15**, 339-367.
- [22] Cribari-Neto, F., Ferrari, S.L.P. (1995a). Bartlett corrected tests for heteroskedastic linear models. *Economics Letters*, **48**, 113-118.
- [23] Cribari-Neto, F., Ferrari, S.L.P. (1995b). Second order asymptotics for score tests in generalized linear models. *Biometrika*, **82**, 426-432.
- [24] Cribari-Neto, F., Ferrari, S.L.P. (1995c). An improved Lagrange multiplier test for heteroskedasticity. *Communications in Statistics - Simulation and Computation*, **24**, 31-44.
- [25] Cribari-Neto, F., Zarkos, S.G. (1995). Improved test statistics for multivariate regression. *Economics Letters*, **49**, 113-120.
- [26] Cribari-Neto, F., Zarkos, S.G. (1999). Bootstrap methods for heteroskedastic regression models: evidence on estimation and testing. *Econometric Reviews*, **1989**, 211-228.
- [27] Cysneiros, A.H.M.A., Cordeiro, G. (2002). On improving the χ^2 approxi-

- mation of score tests in location-scale nonlinear models. *Communications in Statistics, Theory and Methods*, **31**, 1709-1732.
- [28] 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.
 - [29] Devlin, S.J., Gnanadesikan, R., Kettenring, J.R. (1976). Some multivariate applications of elliptical distributions. *Essays in Probability and Statistics*, **24**, 365-393. Tokyo: Shinko Tsusho Co.
 - [30] DiCiccio, T.J. (1986). Approximate conditional inference for location families. *The Canadian Journal of Statistics*, **14**, 5-18.
 - [31] Doornik, J.A. (2001). *Ox: an Object-Oriented Matrix Language*, 4th ed. London: Timberlake Consultants Press; Oxford: <http://www.doornik.com>.
 - [32] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, **7**, 1-26.
 - [33] Efron, B. (1981). Nonparametric standard errors and confidence intervals. *The Canadian Journal of Statistics*, **9**, 139-172.
 - [34] Efron, B., Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. New York: Chapman and Hall.
 - [35] Fang, K.T., Anderson, T.W. (1990). *Statistical Inference in Elliptical Contoured and Related Distributions*. New York: Allerton Press.
 - [36] Fang, K.T., Kotz, S., Ng, K.W. (1990). *Symmetric Multivariate and Related Distributions*. London: Chapman and Hall.
 - [37] Fang, K.T., Zhang, Y.T. (1990). *Generalized Multivariate Analysis*. London: Springer-Verlag.
 - [38] Ferrari, S.L.P., Arellano-Valle, R.B. (1996). Bartlett corrected tests for regression models with Student- t independent errors. *Brazilian Journal of Probability and Statistics*, **10**, 15-33.
 - [39] Ferrari, S.L.P., Cysneiros, A.H.A.M., Cribari-Neto, F. (2004). An Improved Test for Heteroskedasticity Using Adjusted Modified Profile Likelihood Inference. *Journal of Statistical Planning and Inferences*, **57**, 353-361.
 - [40] Ferrari, S.L.P., Cribari-Neto, F. (2002). Corrected modified profile likeli-

- hood heteroskedasticity tests. *Statistics and Probability Letters*, **57**, 353-361.
- [41] 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.
 - [42] Fisher, R.A. (1925). Theory of statistical estimation. *Proceeding of Cambridge Philosophical Society*, **22**, 700-725.
 - [43] Fisk, P.R. (1961). The graduation of income distributions. *Econometrica*, **29**, 171-185.
 - [44] Godfrey, L.G., Orme, C.D. (1999). The robustness, reliability and power of heteroskedasticity tests. *Econometric Reviews*, **18**, 169-194.
 - [45] Gupta, A.K., Varga, T. (1993). *Elliptically Contoured Models in Statistics*. Kluwer Academic Publishers.
 - [46] Hayakawa, T. (1977). The likelihood ratio criterion and the asymptotic expansion of its distribution. *Annals of the Institute of Statistical Mathematics A*, **29**, 359-378.
 - [47] Hayakawa, T. (1987). Correction to ‘The likelihood ratio criterion and the asymptotic expansion of the its distribution.’ *Annals of the Institute of Statistical Mathematics A*, **39**, 681.
 - [48] Harris, P. (1985). An asymptotic expansion for the null distribution of the efficient score statistic. *Biometrika*, **72**, 653-659.
 - [49] Harris, P. (1986). A note on Bartlett adjustments to likelihood ratio tests. *Biometrika*, **73**, 735-737.
 - [50] Hastings, N.A.J., Peacock, J.B. (1975). *Statistical Distributions*. New York: Wiley.
 - [51] Ihaka, R., Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of Computational Graphics and Statistics*, **5**, 299-314.
 - [52] Kelker, D., (1970). Distribution theory of spherical distributions and a location-scale parameter generalization. *Sankhya A*, **32**, 419-430.
 - [53] Knuth, D. (1986). *The T_EXbook*. New York: Addison-Wesley.

- [54] 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.
- [55] Lawley, D.N (1956). A general method for approximating to the distribution of the likelihood ratio criteria. *Biometrika*, **71**, 233-244.
- [56] Lehmann, E.L.,(1999). *Elements of Large-Sample Theory*. New York: Springer-Verlag.
- [57] Lehmann, E.L., Casella, E. (1998). *Theory of Point Estimation*, 2nd ed. New York: Springer-Verlag.
- [58] Lemonte, A.J. (2006). *Inferência Sobre os Parâmetros da Distribuição Birnbaum-Saunders Bi-Paramétrica*. Dissertação de Mestrado. Programa de Pós-graduação em Estatística, Universidade Federal de Pernambuco.
- [59] Little, R.J.(1988). Robust estimation of the mean e covariance matrix from data with missing values. *Applied Statistics*, **37**, 23-39.
- [60] Montenegro, L.C.C., Cordeiro, G.M. (2002). Bartlett corrected likelihood ratio tests in location-scale nonlinear models. *Communnications in Statistics, Theory and Methods*, **30**, 1353-1372.
- [61] Muirhead, M. (1980). The effects of elliptical distributions on some standart procedures involving correlation coefficients. *Multivariate Statistical Analysis*, North-Holland, 143-159.
- [62] Muirhead, M. (1982). *Aspects of Multivariate Statistical Theory*. NewYork: John Wiley.
- [63] Nocedal, J.,Wright, S.J. (1999). *Numerical Optimization*. New York: Springer-Verlag.
- [64] Pearl, R., Reed, L.J. (1920). On the rate of growth of the population of the United States since 1970 and its nathemactical representation. *Proceedings of the National Academy of Sciences*, **6**, 275-288
- [65] Pearl, R., Reed, L.J. (1924). *Studies in Human Biology*. Baltimore: Williams and Wilkins.
- [66] Pearl, R., Reed, L.J., kish, J.F. (1940). The logistic curve and the census count of 1940. *Science*, **92**, 486-488.

- [67] Plackett, R.L. (1959). The analysis of life-test data. *Technometrics*, **1**, 9-19.
- [68] Rao, B.L.S.P. (1990). Remarks on univariate elliptical distributions. *Statistics and Probability Letters*, **10**, 307-315.
- [69] Rao, C.R. (1947). Large sample test of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society*, **44**, 50-57.
- [70] Rao, C.R. (1973). *Linear Statistical Inference and its Applications*, 2nded. New York: John Wiley.
- [71] Schultz, H. (1930). The standart error of a forecast from a curve. *Journal of the American Statistical Association*, **25**, 139-185.
- [72] Taylor, J. (1992). Properties of modeling the erros distribution with an extra shape parameter. *Computational Statistical & Data Analysis*, **13**, 33-46.
- [73] Uribe, M.A.O (1997). *Aperfeiçoamento de Testes Estatísticos em Várias Famílias de distribuição* . Tese de Doutorado, Programa de pós-graduação em Estatística, Universidade de São Paulo..
- [74] Verhulst, P.J. (1838). Notice sur la loi que la population suit dans sons accroissement. *Corr. Math. et Physique*, **10**, 113-121.
- [75] Verhulst, P.J. (1845). Recherches mathématique sur la loi d' accroissement de la population. *Académie de Bruxelles*, **18**, 1-38.
- [76] 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.