KATIA PIRES DO NASCIMENTO

CORREÇÃO TIPO–BARTLETT EM MODELOS NÃO LINEARES SIMÉTRICOS HETEROSCEDÁSTICOS

RECIFE – FEVEREIRO/2010

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO

PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO

EM BIOMETRIA E ESTATÍSTICA APLICADA

CORREÇÃO TIPO–BARTLETT EM MODELOS NÃO LINEARES SIMÉTRICOS HETEROSCEDÁSTICOS

Dissertação apresentada ao Programa de Pós-Graduação em Biometria e Estatística Aplicada como exigência parcial à obtenção do título de Mestre.

Área de Concentração:

Modelagem Estatística e Computacional

Orientadora: Profa. Dra. Laélia P. B. Campos dos Santos.

Co-orientadora: Profa. Dra. Audrey Helen Mariz de Aquino Cysneiros

RECIFE – FEVEREIRO/2010.

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO

PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

CORREÇÃO TIPO–BARTLETT EM MODELOS NÃO LINEARES SIMÉTRICOS HETEROSCEDÁSTICOS

KATIA PIRES DO NASCIMENTO

Dissertação julgada adequada para obtenção do título de mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 25/02/2010 pela Comissão Examinadora.

Orientador:

Profa. Dra. Laélia P. B. Campos dos Santos
Universidade Federal Rural de Pernambuco

Banca Examinadora:

Profa. Dra. Audrey Helen Mariz de Aquino Cysneiros
Universidade Federal de Pernambuco

Prof. Dr. Francisco José de Azevêdo Cysneiros
Universidade Federal de Pernambuco

Prof. Dr. Mário de Castro Andrade Filho
Universidade de São Paulo

*Aos meus pais **José e Lúcia** e ao meu companheiro **Vinicius**, com muito amor.*

Agradecimentos

*Agradeço primeiramente a **DEUS**, por me conceder a vida.*

*Aos meus pais **José Pires** e **Lúcia Soares** pelo amor, carinho e dedicação na formação de meu caráter, e por tudo que sou.*

*Ao meu companheiro **Vinicius**, pelo amor, ajuda, companherismo, apoio incondicional.*

*Aos meus irmãos **Amanda Pires** e **Luciano Pires**, pelo carinho e apoio.*

*À minha co-orientadora, **Audrey Helen Mariz de Aquino Cysneiros**, por ter me concedido um trabalho maravilhoso, pela sua excelente orientação, paciência, disposição, assistência, apoio e amizade, sempre.*

*Ao Prof. **Gauss Moutinho Cordeiro** por ensinar de forma a estimular nos alunos a necessidade de fazer pesquisa.*

*À Profª. **Laélia P. B. Campos dos Santos** pela assistência.*

*Ao Prof. **Borko Stošić** pelo ensino e assistência.*

*Ao Prof. **Tatijana Stošić** pelo ensino e assistência.*

*Ao Secretário **Marco Antônio dos Santos** pela sua assistência.*

*À **Zuleide** pela amizade e carinho durante esses dois anos.*

*Aos colegas, **Alessandra Esteban**, **Amanda Lira**, **Edleide**, **Francisco Guedes** e **Leila**.*

*Ao **REUNE** pelo apoio financeiro e ao **Programa de Pós-Graduação em Biometria e Estatística Aplicada** pelo suporte logístico e intelectual.*

Resumo

Essa dissertação tem dois objetivos. O primeiro é a obtenção de expressões matriciais para o fator de correção tipo–Bartlett para a estatística escore nos modelos não–lineares simétricos heteroscedásticos, com funções de ligação quaisquer para a média e para o parâmetro de dispersão. O segundo é apresentar resultados de simulação de forma a verificar a influência da correção nos modelos em estudo.

Palavras-chave: Correção tipo–Bartlett, Distribuições simétricas, Modelos heteroscedásticos, Modelos não–lineares, Teste escore.

Abstract

This manuscript has two aims. First, we derive general matrix formulae to Bartlett–type correction to the score statistic in a class of heteroscedastic symmetric nonlinear regression models, with link functions any for both mean and dispersion parameter. In the second part Monte Carlo simulations are also performed to assess the influence of the correction in the models studied.

Keywords: Bartlett-type correction, Heteroscedastic model, Nonlinear model, Score test, Symmetric distribution.

Lista de Tabelas

Sumário

1 Introdução

A suposição de normalidade sempre foi muito atrativa para os erros de modelos de regressão com resposta contínua e, mesmo quando não era alcançada, procurava-se alguma transformação na resposta no sentido de obter-se pelo menos a simetria. Contudo, com o passar do tempo, verificou-se que as estimativas obtidas para os coeficientes dos modelos normais mostraram-se sensíveis a observações extremas, comumente chamadas de observações aberrantes, incentivando o desenvolvimento de metodologias robustas contra 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 alternativas à modelagem com erros normais como, por exemplo, o uso de distribuições simétricas (ou elípticas). Grande parte desses resultados podem ser encontrados em Fang et al. (1990) e Fang e Anderson (1990). Esta classe de distribuições contempla distribuições de caudas leves e pesadas, tais como, t de Student, Logística tipo I e II, Normal, Normal Contaminada, dentre outras. Sob a suposição de heteroscedastidade, Cysneiros et al. (2005) propuseram a classe de modelos de regressão lineares simétricos heteroscedásticos. Nesta direção, como estamos interessados em fazer inferências que envolvam alguns, mas não todos os parâmetros do modelo não–linear simétrico heteroscedástico, refinamentos de testes são necessários. Neste caso dizemos que os parâmetros envolvidos são parâmetros de interesse, enquanto que os demais são chamados de parâmetros de perturbação.

Em problemas regulares, a estatística escore (S_R) tem, sob a hipótese nula, uma distribuição χ^2_q aproximadamente, em grandes amostras, ou seja, χ^2 com q graus de liberdade, em que q é a diferença entre as dimensões dos espaços paramétricos sob as hipóteses alternativa e nula. Em geral, há uma grande dificuldade em se determinar a distribuição exata da estatística S_R, razão pela qual os testes têm sidos construídos com base em resultados assintóticos. Os testes são comumente baseados na comparação das estatísticas com valores críticos obtidos na distribuição χ^2 de referência para níveis de significância nominais fixados. Em pequenas amostras ou mesmo em amostras de tamanho moderado, a aproximação pode não ser satisfatória, podendo conduzir a taxa de rejeição, sob a hipótese nula, bastante distorcida. Para melhorar a qualidade da aproximação da distribuição da estatística S_R pela distribuição qui-quadrado é utilizada a chamada correção tipo–Bartlett.

Uribe–Opazo (1997) obteve fatores de correção de Bartlett e tipo–Bartlett para as estatísticas da razão de verossimilhanças e escore nos modelos lineares simétricos homoscedásticos, respectivamente. Fioresi (2000) obteve fatores de correção de Bartlett e tipo–Bartlett para vários testes de hipóteses em modelos normais lineares heteroscedásticos, considerando funções de ligação quaisquer para a média e para a variância. Cordeiro (2004) obteve um fator de correção de Bartlett para a estatística da razão de verossimilhanças em modelos não–lineares simétricos homoscedásticos. Cysneiros et al. (2008) apresentaram, em notação matricial, um fator de correção tipo–Bartlett para a estatística escore nos modelos não–lineares simétricos homoscedásticos. Brito (2007) obteve um fator de correção de Bartlett para a estatística da razão de verossimilhanças em modelos não–lineares simétricos heteroscedásticos. Dando continuidade a estes trabalhos, temos como objetivo principal desta dissertação a obtenção de um refinamento para testes de hipóteses em modelos não–lineares simétricos heteroscedásticos. Mais especificamente, obtemos um fator de correção tipo–Bartlett para a estatística escore original nesta classe de modelos. Estudos de simulação de Monte Carlo serão desenvolvidos para avaliar e comparar numericamente os desempenhos dos testes em amostras finitas.

Esta dissertação de mestrado está organizada da seguinte forma:

No Capítulo 2, revisamos os principais resultados teóricos relacionados com os modelos de regressão não–lineares simétricos heteroscedásticos. Em particular, discutimos a aplicação do teste escore a esta classe de modelos. No Capítulo 3, desenvolvemos um fator de correção tipo–Bartlett para a estatística escore, recorrendo a proposta de Cordeiro e Ferrari (1991), nos modelos não lineares simétricos heteroscedásticos. A fórmula da correção é dada em notação matricial e pode ser implementada em um sistema de computação algébrica para se obter expressão em forma fechada quando aplicada a modelos especiais. Este fator de correção obtido generaliza o resultado em Cysneiros et al. (2008), já que estes autores desenvolveram um fator de correção tipo–Bartlett para a estatística escore nos modelos não–lineares simétricos homoscedásticos. Vale salientar que este capítulo é a principal contribuição teórica desta dissertação.

No Capítulo 4, apresentamos resultados de simulação para avaliar o desempenho dos testes escore usual, escore corrigido via correção tipo–Bartlett e suas versões alternativas. O desempenho dos testes foi avaliado segundo a probabilidade do erro tipo I, em um estudo de simulação. Avaliamos também o poder dos testes em estudo sob algumas situações. As simulações foram realizadas usando a linguagem de programação matricial `Ox` (Doornik, 2001), versão 4.10 para o sistema operacional Windows.

Finalmente, no Capítulo 5, apresentamos algumas conclusões com os principais resultados e contribuições desta dissertação.

Os desenvolvimentos algébricos dos capítulos citados acima se encontram nos apêndices.

2 Modelos Não–lineares Simétricos Heteroscedásticos

2.1 Introdução

A modelagem de dados simétricos é, frequentemente, baseada na suposição de variância constante para os erros. Contudo, em muitas situações práticas essa suposição é dificilmente verificada. A procura de uma transformação na variável resposta para estabilizar a variância nem sempre tem seu sucesso alcançado ou mesmo é recomendável. Quando a suposição de homoscedasticidade do modelo não é verificada, modelos heteroscedásticos são propostos em que a variância do modelo está relacionada, através de uma função de ligação, com um conjunto de variáveis explicativas. Esta é uma das formas de lidar com o problema. A modelagem da variância tem sido largamente discutida principalmente na área de Econometria. Sob erros normais, por exemplo, Cook e Weisberg (1982) e Atkinson (1985) apresentam alguns métodos gráficos para detectar heteroscedasticidade. Importante passo foi dado por Aitkin (1987) que desenvolveu rotinas computacionais no GLIM para a estimação de máxima verossimilhança para modelagem da variância sob erros normais. Verbyla (1993) compara as estimativas de máxima verossimilhança completa e residual, em que o primeiro método estima todos os parâmetros envolvidos no modelo, e o segundo estima os parâmetros de variância. baseando-se na deleção de casos e no afastamento da verossimilhança. Taylor e Verbyla (2004) propõem a modelagem conjunta dos parâmetros de locação e escala no modelo de regressão linear com erros t de Student.

Na classe dos modelos não–lineares simétricos, abordamos a situação em que os parâmetros de dispersão não são constantes para todas as observações, havendo assim uma estrutura heteroscedástica. Analogamente à estrutura estabelecida para a variância no modelo linear, admitimos a determinação de uma forma funcional que relaciona os parâmetros de dispersão com alguns parâmetros desconhecidos, que não dependem do vetor de parâmetros de regressão, e algumas variáveis auxiliares. Sendo assim, temos como um dos objetivos dessa dissertação obter refinamento de testes de hipóteses nesta classe de modelos.

2.2 Definição

No modelo de distribuição, tratado neste texto, as variáveis aleatórias $Y_1, \cdots, Y_n$ são assumidas como sendo independentes e cada observação Y_l tem uma distribuição simétrica, com parâmetros de locação $\mu_l \in \mathbb{R}$ e de escala $\phi_l > 0$, dado por

$$\pi(y_l; \mu_l, \phi_l) = \frac{1}{\sqrt{\phi_l}} g\left(u_l\right), y_l \in \mathbb{R}, \tag{2.1}$$

sendo $g : \mathbb{R} \to [0, \infty)$ tal que $\int_0^\infty g(u)du < \infty$ e $u_l = \phi_l^{-1}(y - \mu_l)^2$. A função $g(\cdot)$ é tipicamente conhecida como função geradora de densidades, com $g(u) \geq 0$. Será denotado que $Y_l \sim S(\mu_l, \phi_l, g)$.

A função característica $\psi_y(t) = E\left(e^{itY_l}\right)$ é dada por $\psi_y(t) = e^{it\mu_l}\varphi(t^2\phi_l)$, $t \in \mathbb{R}$, para alguma função φ, com $\varphi(x) \in \mathbb{R}$ e $x > 0$. Desde que os dois primeiros momentos existam, $E(Y_l) = \mu_l$ e $Var(Y_l) = k\phi_l$, em que $k = -2\varphi^{'}(0)$ é uma constante positiva que não depende de μ_l e ϕ_l e $\varphi^{'}(0) = d\varphi(u)/du|_{u=0}$. Vale ressaltar que, para encontrar o primeiro momento da variável aleatória Y, deve existir o primeiro momento de u; e para encontrar o segundo momento da variável aleatória Y, deve existir o segundo momento de u. Mais detalhes podem ser escontrados em Fang et al. (1990). Por exemplo, tem-se a distribuição t de Student com $\upsilon > 2$ graus de liberdade para $k = \upsilon/(\upsilon - 2)$. Além disso, se a distribuição $S(\mu_l, \phi_l, g)$ tiver r momentos, então $x^{-(r+1)/2}g(x)$ é integrável. A função densidade de probabilidade de $Z_l = (Y_l - \mu_l)/\sqrt{\phi_l}$ é dada por $\pi(u; 0, 1) = g(u^2)$, $u \in \mathbb{R}$, isto é $Z_l \sim S(0, 1, g)$. É introduzida a notação $t(Z_l) = log\left\{g(Z_l^2)\right\}$ e $t^{(k)}(Z_l) = \frac{\partial^k t(Z_l)}{\partial Z_l^k}$, $k = 1, 2, 3, \cdots$ e $\alpha_{r,s} = E\left\{t^r(Z_l)Z_l^s\right\}$ e $\beta_{r,s} = E\left\{t^r(z_l)t^s(Z_l)\right\}$ para $r, s = 0, 1, 2, 3$, α e β dependem de l, $l = 1 \cdots, n$.

A família simétrica (2.1) mantém a estrutura da distribuição normal, mas elimina a forma específica da densidade normal. Isto inclui todas as distribuições simétricas contínuas com caudas mais pesadas do que a normal e tem uma vasta gama de aplicações em vários campos tais como engenharia, biologia, medicina e economia. A classe de distribuições simétricas definidas em (2.1) tem sido estudada por diversos autores (Kelker, 1970; Cambanis et al., 1981). As propriedades destas distribuições, foram exploradas por Muirhead (1980), Berkane e Bentler (1986) e Fang et al. (1990), Johson, et al. (1995). Uma revisão de diferentes áreas em que as distribuições simétricas são aplicados é dada por Chmielewski (1981).

As distribuições simétricas mais conhecidas são a normal ou gaussiana e a t de Student, mas também existem outras distribuições simétricas, como por exemplo, exponencial potencia, Kotz, Kotz generalizada e a t de Student generalizada, sendo importante ressaltar que algumas das propriedades clássicas da distribuição normal são válidas para todas as distribuições simétricas.

Na classe dos modelos simétricos (2.1), serão introduzidas duas estruturas de regressão. Primeiramente, é assumido que a resposta média seja $\mu = (\mu_1, \cdots, \mu_n)^\top$ com

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

em que β é um vetor coluna, com parâmetros de regressão desconhecidos a serem estimados, dado por

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}^\top,$$

sendo β_1 vetor de parâmetros de interesse de dimensão $p_1 \times 1$, com $p_1 \leq p$ e β_2 vetor de parâmetros de perturbação de dimensão $(p - p_1) \times 1$, $f(.;.)$ é uma função, possivelmente não–linear no segundo argumento, contínua e duplamente diferenciável em β e $x_l = (x_{l1}, \cdots, x_{lm})^\top$ é um vetor de dimensão $m \times 1$ de variáveis explicativas, associadas com a i-ésima resposta. Além disso, a matriz de derivadas, de dimensão $n \times p$, de μ com respeito a β, denotada por $\widetilde{X} = \partial \mu / \partial \beta$ é suposta tendo posto completo, isto é, posto$(\tilde{X}) = p$, para todo β. A matriz $\tilde{X}$ tem elementos que são, em geral, funções do vetor de parâmetros β desconhecidos.

A segunda estrutura de regressão a ser introduzida, é uma componente sistemática para o vetor de parâmetro de dispersão $\phi = (\phi_1, \cdots, \phi_n)^\top$ dado por

$$\phi_l = h(\tau_l),$$

sendo que $h(\cdot)$ é uma função monótona conhecida, contínua e diferenciável, do preditor linear de dispersão definida por $\tau_l = z_l^\top \gamma$, sendo $z_l = (z_{l1}, \cdots, z_{lq})^\top$ um vetor de dimensão $q \times 1$ de variáveis explicativas que podem ter componentes em comum com x_l e γ, um vetor de parâmetros desconhecidos a serem estimados, dado por

$$\gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}^\top,$$

com γ_1 sendo o vetor de parâmetros de interesse de dimensão $q_1 \times 1$, com $q_1 \leq q$ e γ_2 sendo o vetor de parâmetros de perturbação de dimensão $(q - q_1) \times 1$. A função $h(\cdot)$ é habitualmente chamada de função de ligação de dispersão ϕ_l, com $h(\cdot) > 0$. Uma possível escolha da função $h(\cdot)$ seria $h(\tau) = exp(\tau)$. As variáveis de dispersão $z_l's$ não são necessariamente as mesmas variáveis de locação $x_l's$. Denotamos por $Y_l \sim S(\mu_l, \phi_l, g)$ e denominamos de variável aleatória simétrica, para $l = 1, \cdots, n$, com ambos os parâmetros de locação μ_l e parâmetros de dispersão $\phi_l > 0$ variando com as observações. Consideremos um modelo não–linear simétrico heteroscedástico definido por (2.1) e com as duas estruturas não–lineares dadas acima, em relação a μ_l e ϕ_l.

2.3 Função de Verossimilhança e Função Escore

Sejam $Y_1, \cdots, Y_n$ n variáveis aleatórias independentes tais que $y_l \sim S(\mu_l, \phi_l, g)$ com a função de densidade de probabilidade dada por (2.1). Dado o vetor de observações $y_1, \cdots, y_n$ do modelo (2.1), sendo $\theta = (\beta^\top, \phi^\top)^T$ a função de verossimilhança é dada por

$$L(\theta) = \prod_{l=1}^{n} \pi(y_l; \mu_l, \phi_l).$$

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

$$\ell(\theta) = \sum_{l=1}^{n} log\left(\frac{g(u_l)}{\sqrt{\phi_l}}\right),$$

com $u_l = \frac{(y_l - \mu_l)^2}{\phi_l}$, $\mu_l = f(x_l; \beta)$ e $\phi_l = h(\tau_l)$ sendo $\tau_l = z_l^\top \gamma$. Temos que $\ell(\theta)$ pode ser escrita da forma

$$\ell(\theta) = -\frac{1}{2} \sum_{l=1}^{n} log\phi_l + \sum_{l=1}^{n} t(z_l) \tag{2.2}$$

em que $t(z_l) = \log g\left(z_l^2\right)$, com $z_l = \sqrt{u_l} = \frac{y_l - \mu_l}{\sqrt{\phi_l}}$.

A função $\ell(\theta)$ é assumida regular (vide Cox e Hinkley, 1974, cap. 9) com respeito às derivadas dos componentes de β e ϕ até a quarta ordem.

Para a obtenção dos estimadores de máxima verossimilhança, da estatística escore e do fator de correção tipo–Bartlett é necessário calcular as derivadas do logaritmo da função de verossimilhança com relação aos parâmetros desconhecidos e, também, alguns momentos destas derivadas. Então, assume–se que tais derivadas e momentos existem. Todavia, não serão consideradas, nesta dissertação, distribuições simétricas que não satisfazem as condições de regularidade, como por exemplo a exponencial dupla. Foram reservados os índices minúsculos $r, s, \upsilon, \cdots$, como notação padrão adotado nas derivadas da função suporte, com relação aos componentes de β; e os índices maiúsculos R, S, T, etc, para denotar componentes do vetor γ, então

$$U_r = \partial\ell(\theta)/\partial\beta_r, U_{rS} = \partial^2\ell(\theta)/\partial\beta_r\partial\gamma_S, U_{rST} = \partial^3\ell(\theta)/\partial\beta_r\partial\gamma_S\partial\gamma_T, etc,$$

e, de maneira análoga, é possível calcular as demais derivadas, com relação aos parâmetros β e γ. Os cumulantes conjunto das derivadas do logaritmo da função de verossimilhança são dados por: $\kappa_{rs} = E(U_{rs})$, $\kappa_{r,s} = E(U_r U_s)$, $\kappa_{rsT} = E(U_{rsT})$, etc, em que todos os $\kappa' s$ referem-se a um total em cima da amostra e são, em geral, de ordem n. Além disso, são definidas as derivadas dos cumulantes por $\kappa_{rs}^{(\upsilon)} = \partial\kappa_{rs}/\partial\beta_\upsilon$, $\kappa_{rs}^{(T)} = \partial\kappa_{rs}/\partial\gamma_T$, etc. As notações utilizadas para indicar as derivadas de μ com relação aos parâmetros de β são

$$(r)_l = \partial\mu_l/\partial\beta_r, (rs)_l = \partial^2\mu_l/\partial\beta_r\partial\beta_s, (r,s\upsilon)_l = \partial\mu_l/\partial\beta_r\partial^2\mu_l/\partial\beta_s\partial\beta_\upsilon, etc,$$

e com relação aos parâmetros de γ,

$$(R)_l = \partial\tau_l/\partial\gamma_R, (RS)_l = \partial^2\tau_l/\partial\gamma_R\partial\gamma_S, (RS,T)_l = \partial^2\tau_l/\partial\gamma_R\partial\gamma_S\partial\tau_l/\partial\gamma_T.$$

As primeiras derivadas do logaritmo da função de de verossimilhança em (2.2) com relação à β e a γ são obtidas, respectivamente, por

$$U_r = \frac{\partial\ell(\theta)}{\partial\beta_r} = -\sum_{l=1}^{n} t_{z_{(l)}}^{(1)} \frac{1}{\sqrt{\phi_l}}(r)_l,\ r = 1,\cdots,p,$$

e

$$U_R = \frac{\partial\ell(\theta)}{\partial\gamma_R} = -\frac{1}{2}\sum_{l=1}^{n}\frac{h_l'}{\phi_l}(R_l) - \frac{1}{2}\sum_{l=1}^{n} t^{(1)}(z_l)\frac{h_l'}{\phi_l}(R_l),\ R = 1,\cdots,q,$$

sendo $t_{(z_l)}^{(1)} = \partial t\,(z_l)/\partial z_l$, com $l = 1,\cdots,n$

As funções escore para β e γ, em notação matricial, tomam, respectivamente, as formas

$$U_r = \tilde{X}^\top S\Lambda\,(y-\mu)\,, \tag{2.3}$$

e

$$U_R \;=\; -\frac{1}{2}\tilde{P}^\top\Lambda F_1 1 + \frac{1}{2}\tilde{P}^\top\Lambda S F_1 u = -\frac{1}{2}\tilde{P}^\top\Lambda\,(SF_1 u - F_1\iota)\,, \tag{2.4}$$

em que P é uma matriz $n \times q$ com linhas $z_l^\top$, $\tilde{P} = \frac{\partial\phi}{\partial\gamma}$, $\tilde{X} = \frac{\partial\mu}{\partial\beta}$ $\Lambda = \mathrm{diag}\left\{\frac{1}{\phi_1},\cdots,\frac{1}{\phi_n}\right\}$, $S = \mathrm{diag}\left\{\mathrm{s}_1,\cdots,\mathrm{s}_n\right\}$, $s_l = \frac{-2g'(u_l)}{g(u_l)}$ com $u_l = \frac{(y_l-\mu_l)^2}{\phi_l}$, $F_1 = \mathrm{diag}\left\{\mathrm{h}_1',\cdots,\mathrm{h}_n'\right\}$, $u = (u_1,\cdots,u_n)^\top$, e ι é um vetor $n \times 1$ de uns, e de agora em diante o "$'$" em cima de h_l denota a derivada com respeito a τ_l para $l = 1,\cdots,n$.

Características de algumas distribuições simétricas são apresentadas na Tabela 2.1.

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

	Normal	Cauchy	t de Student	t de Student Generalizada	Logística TipoI	Logística TipoII
$g(z)$	$\frac{\exp^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}$	$\frac{1}{\pi}(1+z^{2-1})$	$\frac{\nu^{\frac{\nu}{2}}[\nu+z^2]^{-\frac{\nu+1}{2}}}{B(1/2,\nu/2)}$	$\frac{w^{\frac{r}{2}}[w+z^2]^{-\frac{r+1}{2}}}{B(1/2,r/2)}$	$c\frac{\exp^{-z^2}}{(1+\exp^{-z^2})^2}$	$\frac{\exp^{-z}}{(1+\exp^{-z})^2}$
$\frac{g'(z)}{g(z)}$	$-\frac{1}{2}$	$-\frac{1}{1+z^2}$	$-\frac{\nu+1}{2(\nu+z^2)}$	$-\frac{r+1}{2(w+z^2)}$	$-\frac{1-\exp-z^2}{1+\exp-z^2}$	$-\frac{1}{2}\frac{(\exp\lvert z\rvert-1)}{(\lvert z\rvert\exp\lvert z\rvert+1)}$
s	1	$\frac{2}{1+z^2}$	$\frac{\nu+1}{(\nu+z^2)}$	$\frac{r+1}{(w+z^2)}$	$\frac{2(1-\exp-z^2)}{1+\exp-z^2}$	$\frac{(\exp\lvert z\rvert-1)}{(\lvert z\rvert\exp\lvert z\rvert+1)}$

2.4 Informação de Fisher

Nesta seção, será calculada a matriz de informação de Fisher, que será utilizada no processo iterativo de estimação através do método escore de Fisher, mais a diante. As segundas derivadas do logaritmo da função da verossimilhança têm as formas

$$\frac{\partial^2 \ell(\theta)}{\partial \beta_r \partial \beta_s} = \sum_{l=1}^{n} t_{(z_l)}^{(2)} \frac{1}{\phi_l} (r,s)_l - \sum_{l=1}^{n} t_{z_l}^{(1)} \frac{1}{\sqrt{\phi_l}} (rs)_l \,,$$

$$\frac{\partial^2 \ell(\theta)}{\partial \beta_r \partial \phi_S} = \frac{1}{2} \sum_{l=1}^{n} t_{(z_l)}^{(2)} z_l \frac{h_l'}{\phi_l^{3/2}} (S)_l (r)_l + \frac{1}{2} \sum_{l=1}^{n} t_{(z_l)}^{(1)} \frac{h_l'}{\phi_l^{3/2}} (S)_l (r)_l \,,$$

e

$$\frac{\partial^2 \ell(\theta)}{\partial \phi_R \partial \phi_S} = -\frac{1}{2} \sum_{l=1}^{n} \frac{h_l'' \phi_l - h_l'^2}{\phi_l^2} + \frac{1}{4} \sum_{l=1}^{n} t_{(z_l)}^{(2)} z_l^2 \frac{h_l'^2}{\phi_l^2} + \frac{1}{4} \sum_{l=1}^{n} t_{(z_l)}^{(1)} z_l \frac{2h_l'' \phi_l - 3h_l'^2}{\phi_l^2}.$$

A matriz de informação de Fisher de dimensão quadrada $(q+p)\times(q+p)$, contida nos dados y, é definida por

$$K(\theta) = E\left\{U(\theta)U(\theta)^\top\right\}, \tag{2.5}$$

sendo $U(\theta)$ o vetor escore da superfície suporte de θ.

A matriz de primeiras derivadas da função escore com sinal negativo $J = -\frac{\partial U(\theta)^\top}{\partial \theta} = -\frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^\top}$ é denominada matriz de informação observada. Obtém-se a matriz de informação para θ calculando $E(J) = K(\theta)$. Deste modo, a matriz de informação total de Fisher para θ na classe dos modelos simétricos é dada por

$$K(\theta) = -E\begin{bmatrix} \frac{\partial^2 \ell(\theta)}{\partial \beta_r \beta_s} & \frac{\partial^2 \ell(\theta)}{\partial \beta_r \partial \phi_S} \\ \frac{\partial^2 \ell(\theta)}{\partial \phi_R \partial \beta_s} & \frac{\partial^2 \ell(\theta)}{\partial \phi_R \phi_S} \end{bmatrix}.$$

Deste modo, a matriz de informação total de Fisher $K = K(\beta,\gamma)$ para $(\beta^\top,\gamma^\top)^\top$ é diagonal em blocos com submatrizes

$$K_{\beta,\beta} = -K_{\beta\beta} = E\left[-\frac{\partial^2 \ell(\theta)}{\partial \beta_r \partial \beta_s}\right] = \delta_{(2,0,0,0,0)} \widetilde{X}^\top \Lambda \widetilde{X}, \tag{2.6}$$

$$K_{\gamma,\gamma} = -K_{\gamma\gamma} = E\left[\frac{\partial^2 \ell(\theta)}{\partial \phi_R \partial \phi_S}\right] = \frac{\left(\delta_{(2,0,0,0,2)} - 1\right)}{4} P^\top \Lambda_5 P = P^\top V P, \tag{2.7}$$

e

$$E\left[\frac{\partial^2 \ell(\theta)}{\partial \phi_R \partial \beta_s}\right] = 0, \tag{2.8}$$

em que $\widetilde{X}$ foi definida na seção 2.2 , P foi definida na seção 2.3, $V = \mathrm{diag}\{\mathrm{v}_1, \cdots, \mathrm{v}_\mathrm{n}\}$ e $v_l = \frac{(\delta_{(2,0,0,0,2)}-1)}{4}\Lambda_5$, $\Lambda_5 = \left\{\frac{h_l'^2}{\phi_l^2}\right\}$, para $l = 1, \cdots, n$.

Os parâmetros β e γ são globalmente ortogonais e os estimadores de máxima verossimilhança $\widehat{\beta}$ e $\widehat{\gamma}$ são assintoticamente independentes.

Substituindo os resultados obtidos acima na matriz de informação tem-se

$$K(\theta) = \begin{bmatrix} \delta_{(2,0,0,0,0)}\widetilde{X}^\top \Lambda \widetilde{X} & 0 \\ 0 & P^\top V P \end{bmatrix},$$

uma vez que está sendo utilizada a seguinte notação

$$\delta\left(a,b,c,d,e\right) = E\left\{t^{(1)a}t^{(2)b}t^{(3)c}t^{(4)d}z_l^e\right\}, \tag{2.9}$$

com $a,b,c,d,e = 1,2,3,4$, $t^{(k)} = \partial^k t/\partial z^k$, $k = 1,\cdots,4$. Pelas condições de regularidade (vide Cox e Hinkley, 1974, Cap. 9), tem-se que $\delta_{(1,0,0,0,0)} = 0$ e $\delta_{(0,1,0,0,0)} = -\delta_{(2,0,0,0,0)}$. A função $\delta_{(a,b,c,d,e)}$ definida em (2.9) é de grande utilidade para os cálculos das correções tipo–Bartlett para a estatística escore na família de distribuições simétricas. No Apêndice B são dados os $\delta's$ necessários à obtenção de algorítmos para o cálculo em algumas distribuições.

A seguir apresentaremos alguns algoritmos iterativos para a obtenção dos estimadores de máxima verossimilhança.

2.5 Métodos Iterativos

O processo iterativo de Newton-Raphson para obter a estimativa de máxima verossimilhança de θ é definido expandindo-se a função escore, em função de θ, em série de Taylor em torno de um valor inicial $\theta^{(0)}$ de modo que

$$U\left(\theta\right) \cong U\left(\theta^{(m)}\right) + U'\left(\theta^{(m)}\right)\left(\theta - \theta^{(m)}\right),$$

$m = 0,1,\cdots$, sendo $U'\left(\theta\right)$ correspondente à primeira derivada de $U(\theta)$ com respeito a θ. Assim, como $U(\hat{\theta}) = 0$, chega-se ao seguinte processo iterativo

$$\theta^{(m+1)} = \theta^{(m)} + \left\{-U'\left(\theta^{(m)}\right)\right\}^{-1} U\left(\theta^{(m)}\right),$$

com $m = 0,1,\cdots$. Como a matriz $-U'\left(\theta^{(m)}\right)$ pode não ser positiva definida, a aplicação do método escore de Fisher substituindo a matriz $-U'\left(\theta^{(m)}\right)$ pelo correspondente valor esperado, pode ser mais adequado. Resultando no seguinte processo iterativo

$$\theta^{(m+1)} = \theta^{(m)} + K\left(\theta^{(m)}\right)^{-1} U\left(\theta^{(m)}\right),$$

$m = 0, 1, \cdots$, ou equivalentemente

$$\begin{cases} \beta^{(m+1)} = \beta^{(m)} + K\left(\beta^{(m)}\right)^{-1} U(\beta^{(m)}) \\ \gamma^{(m+1)} = \gamma^{(m)} + K\left(\gamma^{(m)}\right)^{-1} U(\gamma^{(m)}), \end{cases}$$

que em forma matricial é dada por

$$\beta^{(m+1)} = \beta^{(m)} + \frac{1}{\delta(0,1,0,0,0)} \left\{ \widetilde{X}^{(m)\top} \Lambda^{(m)} \widetilde{X}^{(m)} \right\}^{-1} \widetilde{X}^{(m)\top} S^{(m)} \Lambda^{(m)} \left(y - \mu^{(m)}\right) \tag{2.10}$$

e

$$\gamma^{(m+1)} = \left(P^\top V^m P\right)^{-1} P^\top V^{(m)} \left(\eta^{(m)} + \delta^{(m)}\right), \tag{2.11}$$

$m = 0, 1, \cdots$, sendo $\eta = \widetilde{Z}\gamma$, ι é um vetor de uns, as matrizes S, F_1, Λ e os vetores s_l, u_l foram definidos na seção 2.3, a matriz V foi definido na Seção 2.4.

Além disso, se o logaritmo da função de verossimilhança, definida na seção 2.4, satisfaz as condições de regularidade sob as quais o EMV de $\theta = (\beta^\top, \phi^\top)^\top$ é assintoticamente normal, então

$$\left(\widehat{\beta}^\top, \widehat{\phi}^\top\right)^\top \sim AN_{p+1}\left(\left(\beta^\top, \phi^\top\right)^\top, K^{-1}\right),$$

em que $K^{-1} = \mathrm{Diag}\left\{\mathrm{K}_{\beta,\beta}^{-1}, \mathrm{K}_{\phi,\phi}^{-1}\right\}$, com $K_{\beta,\beta}$ e $K_{\phi,\phi}$ dadas em (2.6) e (2.7), respectivamente, na matriz de informação.

As equações (2.7) e (2.8) mostram que qualquer linguagem que tenha uma rotina de regressão linear ponderada pode ser usada para calcular os EMVs $\hat{\beta}$ e $\hat{\gamma}$ iterativamente. Aproximações iniciais $\beta^{(1)}$ e $\gamma^{(1)}$ para o algoritmo iterativo são usados para avaliar $\widetilde{X}^{(1)}$, $\mu^{(1)}$, $\Lambda^{(1)}$, $V^{(1)}$, $\eta^{(1)}$, $F^{(1)}$, $u^{(1)}$ e $\gamma^{(1)}$ das quais essas equações são utilizadas para obter as próximas estimativas de $\beta^{(2)}$ e $\gamma^{(2)}$. Esses novos valores atualizam $\widetilde{X}$, μ, Λ, V, η, F, u e γ e, então, as iterações continuam até a convergência ser atingida. As matrizes de covariâncias assintótica de $\widehat{\beta}$ e $\widehat{\gamma}$ são $\widehat{K_\beta}^{-1} = -\delta_{2,0}^{-1}\left(\widetilde{X}^\top \widehat{\Lambda}^{-1} \widetilde{X}\right)^{-1}$ e $\widehat{K_\gamma}^{-1} = Z^\top \widehat{V} Z^{-1}$, respectivamente.

Vale, ainda, acrescentar que quando existirem nos modelos outros parâmetros tais como graus de liberdade, é necessário obter a matriz de informação para todos os parâmetros e estimá-los. Outra alternativa, talvez, seria repetir o processo iterativo para uma gama de valores para os parâmetros extras e escolher aquele valor que produz o maior valor para a função de verossimilhança. Assumimos então que nosso modelo satisfaz as suposições habituais da teoria de verossimilhança em amostras de tamanho grande.

2.6 Teste Escore em Modelos Não–lineares Simétricos Heteroscedásticos

Nesta seção, obteremos a estatística escore para testes simultâneos sobre a média e o parâmetro de precisão, ou seja, sobre $\left(\beta_1^\top,\gamma_1^\top\right)^\top$, dos modelos não–lineares simétricos heteroscedásticos, definidos em (2.1). Considere o vetor $y=(y_1,\cdots,y_n)^\top$ representando n observações independentes em que cada y_l tem uma função densidade na família simétrica definida em (2.1). Assumimos que a função $\ell(\beta)$ em (2.2) é regular com respeito às derivadas em relação aos componentes de β até a quarta ordem, consideremos as seguintes partições $\beta=\left(\beta_1^\top,\beta_2^\top\right)^T$, em que $\beta_1=(\beta_1,\cdots,\beta_{p_1})^\top$, $\beta_2=\left(\beta_{p_1+1},\cdots,\beta_p\right)^\top$, $\gamma=\left(\gamma_1^\top,\gamma_2^\top\right)^T$, sendo $\gamma_1=\left(\gamma_1,\cdots,\gamma_{q_1}\right)^\top$ e $\gamma_{q_1+1}=\left(\gamma_1,\cdots,\gamma_q\right)^\top$, com $p_1\leq p$ e $q_1\leq q$.

Neste trabalho não consideraremos modificações no valor crítico da estatística escore, restringiremos o escopo do estudos de correções na própria estatística e teremos resultados com relação à monotonicidade destas estatística apenas via simulação.

2.6.1 Testes simultâneos sobre a média e o parâmetro de precisão

Para o modelo (2.1) estamos interessados em testar a hipótese $\mathscr{H}_0^1:\beta_1=\beta_1^{(0)}$, $\gamma_1=\gamma_1^{(0)}$, contra $\mathscr{H}_1^1$: pelo menos uma das igualdades é violada, em que $\beta_1^{(0)}$ e $\gamma_1^{(0)}$ são vetores fixos de dimensões $p_1\times 1$ e $q_1\times 1$, respectivamente. Considerando $p_1=p-1$ e $q_1=q-1$, queremos testar, em particular, a hipótese $\mathscr{H}_0^1:\beta_1=0$, $\gamma_1=0$, com $x_{lp}=1$ e $z_{lp}=1$, com $l=1,\cdots,n$. O teste de $\mathscr{H}_0^1$ equivale a testar se as variáveis $y_1,\cdots,y_n$ são equivalentes, ou seja i.i.d.

Seguindo as partições induzidas por $\mathscr{H}_0^1$, considere $X=(X_1,X_2)$ e $P=(P_1,P_2)$ sendo as matrizes particionadas correspondentes ao modelo, em que X_1, X_2, P_1 e P_2 são respectivamente, $n\times p_1$, $n\times(p-p_1)$, $n\times q_1$ e $n\times(q-q_1)$ matrizes de posto completo. A função escore, correspondente é dada por

$$U=\Big(U_{\beta_1}^\top(\beta_1,\beta_2,\gamma_1,\gamma_2),\ U_{\beta_2}^\top(\beta_1,\beta_2,\gamma_1,\gamma_2),\ U_{\gamma_1}^\top(\beta_1,\beta_2,\gamma_1,\gamma_2),\ U_{\gamma_2}^\top(\beta_1,\beta_2,\gamma_1,\gamma_2)\Big)^\top.$$

A matriz de informação total de Fisher $K=K(\beta,\gamma)$, para $(\beta^\top,\gamma^\top)^\top$ é diagonal em blocos com submatrizes

$$K_{\beta,\beta}=\begin{pmatrix} K_{\beta_{11}} & K_{\beta_{12}} \\ K_{\beta_{21}} & K_{\beta_{22}} \end{pmatrix}\ e\ K_{\gamma,\gamma}=\begin{pmatrix} K_{\gamma_{11}} & K_{\gamma_{12}} \\ K_{\gamma_{21}} & K_{\gamma_{22}} \end{pmatrix},$$

sendo $K_{\beta_{11}}=\delta_{(2,0,0,0,0)}X_1^\top\Lambda X_1$, $K_{\beta_{12}}=K_{\beta_{21}}^\top=\delta_{(2,0,0,0,0)}X_1^\top\Lambda X_2$, $K_{\beta_{22}}=\delta_{(2,0,0,0,0)}X_2^\top\Lambda X_2$,

$K_{\gamma_{11}} = (\delta_{(2,0,0,0,2)} - 1)/4 P_1^\top \Lambda_5 P_1$ e $K_{\gamma_{12}} = K_{\gamma_{21}}^\top = (\delta_{(2,0,0,0,2)} - 1)/4 P_1^\top \Lambda_5 P_2$,
$K_{\gamma_{22}} = (\delta_{(2,0,0,0,2)} - 1)/4 P_2^\top \Lambda_5 P_2$.

A estatística escore para testar a hipótese $\mathscr{H}_0^1 : \beta_1 = \beta_1^{(0)}$, $\gamma_1 = \gamma_1^{(0)}$ é

$$S_R = U(\widetilde{\theta})^\top \widetilde{K}^{-1} U(\widetilde{\theta}),$$

que pode ser escrita como a soma de duas formas quadráticas, a saber

$$S_R = \widetilde{r}^\top \widetilde{X}_1 (\widetilde{X}_1^\top \Lambda \widetilde{X}_1)^{-1} \widetilde{X}_1^\top \widetilde{r} + \widetilde{\zeta}^\top \widetilde{P}_1 \left(\widetilde{P}_1^\top \Lambda_5 \widetilde{P}_1\right)^{-1} \widetilde{P}_1^\top \widetilde{\zeta}, \tag{2.12}$$

em que $r = (r_1, \cdots, r_n)^\top$, com $r_l = \Lambda^{1/2} s_l z_l \delta_{(2,0,0,0,0)}^{-1/2}$, $\zeta = (\zeta_1, \cdots, \zeta_n)^\top$, sendo $\zeta_l = (u^\top F_1 S^\top - F_1^\top 1)(\delta_{(2,0,0,0,2)} - 1)^{-1/2} \Lambda$. Os vetores s_l, z_l e Λ foram definidos na Seção 2.3, Λ_5 na Seção (2.4), bem como as matrizes F_1 e S foram definidas na Seção 2.6.

A estatística escore para testar a hipótese $\mathscr{H}_0^1 : \beta_1 = \beta_1^{(0)}$, $\gamma_1 = \gamma_1^{(0)}$, originalmente sugerida por Rao (1947), é dada por

$$S_R = U_1^\top (\widetilde{\theta}) K^{11} (\widetilde{\theta}) U_1 (\widetilde{\theta}), \tag{2.13}$$

em que $U_1^\top(\widetilde{\theta})$ e $K^{11}(\widetilde{\theta})$ são, respectivamente, a função escore e a inversa da matriz de informação total de Fisher para β e γ avaliadas sob $\mathscr{H}_0^1$, com $\widetilde{\theta} = \left(\widetilde{\beta}^\top, \widetilde{\gamma}^\top\right) = (\beta_1^{(0)^\top}, \widetilde{\beta}_2^\top, \gamma_1^{(0)^\top}, \widetilde{\gamma}_2^\top)^\top$, sendo $\widetilde{\beta}_2$ e $\widetilde{\gamma}_2$ os estimador de máxima verossimilhança de β_2 e γ_2, respectivamente, sob a hipótese nula. Assintoticamente e sob a hipótese nula, temos que

$$S_R \xrightarrow{D} \chi^2_{p_1+q_1}, \quad \text{quando } n \to \infty,$$

ou equivalentemente, que a estatística escore converge em distribuição para a distribuição qui-quadrado com $p_1 + q_1$ graus de liberdade.

2.6.2 Testes de hipóteses sobre o parâmetro de precisão

Para o modelo (2.1) estamos interessados em testar a hipótese $\mathscr{H}_0^2 : \gamma_1 = \gamma_1^{(0)}$ contra a hipótese alternativa $\mathscr{H}_1^2 : \gamma_1 \neq \gamma_1^{(0)}$, sendo $\gamma_1^{(0)}$ vetor fixado de dimensão $q_1 \times 1$. Supondo $p_1 = 0$ e $q_1 = q - 1$, um caso especial, diz respeito a testar a hipótese $\mathscr{H}_0'^2 : \gamma_1 = 0$, com $z_{lq} = 1$, para $l = 1, \cdots, n$. A hipótese $\mathscr{H}_0'^2$ equivale à homoscedasticidade, ou seja, $Y_1, \cdots, Y_n$ têm a mesma variância.

A estatística escore S_R para testar a hipótese nula $\mathscr{H}_0^2 : \gamma_1 = \gamma_1^{(0)}$ contra a hipótese alternativa $\mathscr{H}_1^2 : \gamma_1 \neq \gamma_1^{(0)}$, com β representando o parâmetro de perturbação e $\gamma_1^{(0)}$ é um valor fixo positivo, tem a forma

$$S_R = U_1^\top(\widetilde{\gamma})\widetilde{K}^{11}(\widetilde{\gamma})U_1(\widetilde{\gamma}), \tag{2.14}$$

em que a função escore $U_1(\tilde{\gamma})$ e a inversa da matriz de informação total de Fisher para γ $K_{11}(\tilde{\gamma})$ estão avaliadas em $\left(\widetilde{\beta}^\top, \gamma_1^{(0)T}, \widetilde{\gamma}_2^\top\right)^\top$, sendo $\widetilde{\beta}$ e $\widetilde{\gamma}_2$ os estimadores de máxima verossimilhança de β e γ_2 sob $\mathscr{H}_0^{(2)}$. Assintoticamente e sob a hipótese nula, temos que

$$S_R \xrightarrow{D} \chi^2_{q_1}, \quad \text{quando } n \to \infty,$$

sendo q_1 o número de graus de liberdade. Assim, substituindo os valores de $U_1(\gamma)$ e $K_{11}^{-1}(\gamma)$ na equação 2.14, tem-se

$$S_R = \widetilde{\zeta}^\top \widetilde{P}_1 \left(\widetilde{P}_1^\top \Lambda_5 \widetilde{P}_1\right)^{-1} \widetilde{P}_1^\top \widetilde{\zeta}, \tag{2.15}$$

sendo o vetor ζ definido na Subseção 2.6.1, as matrizes P e Λ_5 definidas na Seção 2.3 e 2.4, respectivamente.

2.6.3 Testes de hipóteses sobre a média

O principal objetivo é testar a hipótese composta $\mathscr{H}_0^3 : \beta_1 = \beta_1^{(0)}$ contra $\mathscr{H}_1^3 : \beta_1 \neq \beta_1^{(0)}$ em que o vetor de parâmetros fixos β com p componentes é particionado como $\beta = (\beta_1^\top, \beta_2^\top)^\top$, sendo $\beta_1 = (\beta_1, \beta_2, \cdots, \beta_{p_1})^\top$ o vetor de parâmetros de interesse e $\beta_2 = (\beta_{p_1+1}, \beta_{p_1+2}, \cdots, \beta_p)^\top$ e $\gamma = (\gamma_1, \cdots, \gamma_q)$ são vetores de parâmetros de perturbação e $\beta_1^{(0)}$ é um vetor especificado de dimensão $p_1 \times 1$, com $(p_1 \leq p)$. A matriz de planejamento é particionada de acordo com a partição de β, isto é, $\widetilde{X} = \left(\widetilde{X}_1, \widetilde{X}_2\right)$. Denota-se que $\widehat{\theta} = \left(\widehat{\beta}_1^\top, \widehat{\beta}_2^\top, \widehat{\gamma}^\top\right)^\top$ é o estimador de máxima verossimilhança irrestrito de θ e por $\widetilde{\theta}$ o estimador de máxima verossimilhança de θ restrito à hipótese nula.

A função escore para β pode ser particionada como $U(\beta) = \left(U_1^\top(\beta), U_2^\top(\beta)\right)^\top$, sendo

$$U_1(\beta) = \tilde{X}_1^\top S \Lambda (y - \mu) \quad \text{e} \quad U_2(\beta) = \tilde{X}_2^\top \Lambda S (y - \mu);$$

a matriz Λ foi definida na Seção 2.4. A matriz de informação de Fisher correspondente ao parâmetro de β supondo ϕ conhecido é dada por

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

com a matriz $K(\beta)$ positiva e definida da seguinte forma

$$K_{11}(\beta) = -\delta_{(0,1,0,0,0)}\widetilde{X}_1^\top \Lambda \widetilde{X}_1,$$
$$K_{22}(\beta) = -\delta_{(0,1,0,0,0)}\widetilde{X}_2^\top \Lambda \widetilde{X}_2,$$

e

$$K_{21}(\beta) = K_{12}(\beta) = -\delta_{(0,1,0,0,0)}\widetilde{X}_1^\top \Lambda \widetilde{X}_2.$$

em que

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

sendo

$$K^{11}(\beta) = -\delta^{-1}_{(0,1,0,0,0)}\left(\widetilde{X}_1^\top \Lambda \widetilde{X}_1\right)^{-1},$$
$$K^{12}(\beta) = K^{21}(\beta) = -\delta^{-1}_{(0,1,0,0,0)}\left(\widetilde{X}_1^\top \Lambda \widetilde{X}_2\right)^{-1},$$

e

$$K_{22}(\beta) = -\delta^{-1}_{(0,1,0,0,0)}\left(\widetilde{X}_2^\top \Lambda \widetilde{X}_2\right)^{-1}.$$

Para testar a hipótese nula $\mathscr{H}_0^3$ versus a alternativa $\mathscr{H}_1^3$ será utilizada a estatística escore S_R, que em problemas regulares, tem, segundo a hipótese nula $\mathscr{H}_0^3$, distribuição assintótica qui-quadrado com q graus de liberdade, sendo que q é a diferença entre as dimensões dos espaços paramétricos sob a hipótese alternativa e nula.

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

$$S_R = U_1^\top(\widetilde{\theta})K^{11}(\widetilde{\theta})U_1(\widetilde{\theta}), \tag{2.16}$$

em que $\widetilde{\theta} = \left(\beta_1^{(0)\top}, \tilde{\beta}_2^\top, \tilde{\gamma}^\top\right)^\top$ e $\tilde{\beta}_2^\top$ e $\tilde{\gamma}^\top$ são os estimadores de máxima verossimilhança de θ sob $\mathscr{H}_{(0)}^3$. Assintoticamente e sob a hipótese nula, temos que

$$S_R \xrightarrow{D} \chi^2_{p_1}, \quad \text{quando}\, n \to \infty,$$

sendo p o número de graus de liberdade.

Então, substituindo os valores de $U_1(\tilde{\beta})$ e $K^{11}(\tilde{\beta})$ na estatística escore (2.16) para testar $\mathscr{H}_0^3$ versus $\mathscr{H}_1^3$, tem-se

$$S_R = \widetilde{r}^\top \widetilde{X}_1(\widetilde{X}_1^\top \Lambda \widetilde{X}_1)^{-1}\widetilde{X}_1^\top \widetilde{r}, \tag{2.17}$$

em que a função escore $U_1(\tilde{\beta})$ e o inversa da matriz de informação total de Fisher para β $K^{11}(\tilde{\beta})$ estão avaliadas em $\left(\beta_1^{(0)T}, \tilde{\beta}_2^\top, \tilde{\gamma}^\top\right)^\top$, sendo $\widetilde{\beta}_2$ e $\tilde{\gamma}$ é o estimador de máxima verossimilhança de θ sob $\mathscr{H}_{(0)}^3$, o vetor r foi dado na Subseção $2.6.1$.

No próximo capítulo desenvolvemos um fator de correção tipo–Bartlett, em notação matricial, para a estatística escore, via Cordeiro e Ferrari (1991).

3 Correção Tipo–Bartlett Para a Estatística Escore

3.1 Introdução

Este capítulo visa a obter ajustes para estatísticas de teste. Mais especificamente, enfocaremos a estatística escore, S_R, (Rao, 1947). Nos casos em que a estimação sob a hipótes alternativa é complicada, o teste baseado na estatística escore apresenta vantagem computacional em relação a outros testes pois requer apenas a estimação dos parâmetros sob a hipótese nula. Sabe-se que os testes baseados nas estatísticas da razão de verossimilhanças (LR), S_R e Wald (W) são equivalentes em grandes amostras e, em problemas regulares, convergem segundo a hipótese nula $\mathscr{H}_0$ para a distribuição χ^2_q, em que q é o número de restrições impostas por $\mathscr{H}_0$. Entretanto, em pequenas amostras, a aproximação da distribuição da estatística de teste pela distribuição χ^2 pode não ser satisfatória. A primeira idéia para melhorar as propriedades de estatísticas de testes foi proposta por Bartlett (1953) considerando apenas a estatística da razão de verossimilhanças, computando o seu valor esperado segundo a hipótese nula até a ordem n^{-1}, em que n é o tamanho da amostra. Harris (1985) obteve uma expansão assintótica para a distribuição da estatística escore S_R até ordem n^{-1}. Honda (1988) derivou a correção do valor crítico da estatística S_R para o teste de homoscedasticidade no modelo normal heteroscedástico.

Através do trabalho de Harris (1985), Cordeiro e Ferrari (1991) demonstraram que, sob condições gerais de regularidade, qualquer estatística cuja distribuição assintótica é qui-quadrado pode ser aperfeiçoada por um fator de correção multiplicativo expresso como um polinômio de segundo grau na própria estatística. Baseado no trabalho de Cordeiro e Ferrari (1991), muitos resultados têm sido publicados envolvendo estatísticas escore aperfeiçoadas por fatores de correções tipo–Bartlett em várias classes de modelos de regressão. Ferrari e Cordeiro (1994) desenvolveram expressões matriciais para o fator de correção tipo–Bartlett da estatística escore em problemas com parâmetros globalmente ortogonais. Cribari-Neto e Ferrari (1995a) aperfeiçoaram o teste escore nos modelos lineares heteroscedásticos. Similarmente, correções tipo–Bartlett para a estatística escore em alguns modelos de regressão multivariados foram obtidas por Cribari-Neto e Zarkos (1999).

Cribari-Neto e Ferrari (1995b) obtiveram um fator de correção tipo–Bartlett para a estatística escore em modelos normais lineares heteroscedásticos. Uribe-Opazo (1997) obteve um fator de correção de Bartlett para a estatística da razão de verossimilhanças, generalizando o trabalho de Ferrari e Arellano-Valle (1996); e, ainda, a correção tipo–Bartlett nos modelos lineares simétricos homoscedásticos, sob o enfoque de Cordeiro e Ferrari (1991). Cordeiro et al. (2000) obtiveram a correção de viés dos estimadores de máxima verossimilhança na classe dos modelos de regressão não–lineares simétricos homoscedásticos. Fioresi (2000) obteve fatores de correção de Bartlett e tipo–Bartlett para as estatísticas da razão de verossimilhanças e escore, respectivamente, na classe dos modelos normais heterocedásticos com funcões de ligação quaisquer para a média e para o parâmetro de precisão. Um fator de correção tipo–Bartlett para a estatística escore em modelos de regressão não–lineares simétricos homoscedásticos foram desenvolvidas por Cysneiros et al. (2008).

Neste capítulo, obteremos um fator de correção tipo–Bartlett para a estatística escore para vários testes de hipóteses: sobre a média e/ou o parâmetro de precisão, em modelos não–lineares simétricos heteroscedásticos. Generalizamos, portanto, os resultados de Uribe-Opazo et al. (2008), Fioresi (2000) e Cysneiros et al. (2008).

3.2 Correção Tipo–Bartlett

Cordeiro e Ferrari (1991) mostraram que, em problemas regulares, a estatística escore, S_R, pode ser melhorada por uma correção tipo–Bartlett que não é exatamente a correção de Bartlett porque envolve um polinômio de segundo grau na estatística original, produzindo uma estatística escore modificada ajustada com distribuição χ^2 até ordem n^{-1}, segundo a hipótese nula. Cordeiro e Ferrari (1991) propuseram a estatística escore modificada, dada por

$$S_R^* = S_R\left[1 - \frac{A_3}{12u(u+2)(u+4)}S_R^2 - \frac{A_2 - 2A_3}{12(u+2)}S_R - \frac{A_1 - A_2 + A_3}{12u}\right], \tag{3.1}$$

sendo A_1, A_2, e A_3 funções de cumulantes conjuntos de derivadas do logaritmo da função de verossimilhança. Fórmulas matriciais para estas funções relativas aos testes escores são dadas, em generalidade, por Ferrari e Cordeiro (1994). Os coeficientes A_1, A_2 e A_3, na forma matricial, podem ser escritos como

$$
\begin{aligned}
A_1 &= 3\sum_{l}\beta,\gamma\left(\kappa_{l_1l_2l_3}+2\kappa_{l_1,l_2l_3}\right)\left(\kappa_{l_4l_5l_6}+2\kappa_{l_4,l_5l_6}\right)a_{l_1l_2}\,a_{l_5l_6}\,m_{l_3l_4}\\
&\quad -6\sum_{l}\beta,\gamma\left(\kappa_{l_1l_2l_3}+2\kappa_{l_1,l_2l_3}\right)\kappa_{l_4,l_5,l_6}a_{l_1l_2}\,a_{l_3l_4}\,m_{l_5l_6}\\
&\quad +6\sum_{l}\beta,\gamma\left(\kappa_{l_1,l_2l_3}-2\kappa_{l_1,l_2,l_3}\right)\left(\kappa_{l_4l_5l_6}+2\kappa_{l_4l_5,l_6}\right)a_{l_2l_5}\,a_{l_3l_6}\,m_{l_1l_4}\\
&\quad +6\sum_{l}\beta,\gamma\left(\kappa_{l_1l_2l_3l_4}+\kappa_{l_1,l_2,l_3l_4}\right)a_{l_3l_4}m_{l_1l_2},\\
&= A_{11}+A_{12}+A_{13}+A_{14},
\end{aligned} \tag{3.2}
$$

$$
\begin{aligned}
A_2 &= -3\sum_{l}\beta,\gamma\kappa_{l_1,l_2,l_3}\kappa_{l_4,l_5,l_6}a_{l_3l_4}m_{l_1l_2}m_{l_5l_6}\\
&\quad +6\sum_{l}\beta,\gamma\left(\kappa_{l_1l_2l_3}+2\kappa_{l_1,l_2l_3}\right)\kappa_{l_4,l_5,l_6}a_{l_1l_2}m_{l_3l_4}m_{l_5l_6}\\
&\quad -6\sum_{l}\beta,\gamma\kappa_{l_1,l_2,l_3}\kappa_{l_4,l_5,l_6}a_{l_3l_6}m_{l_1l_4}m_{l_2l_5}\\
&\quad +3\sum_{l}\beta,\gamma\kappa_{l_1,l_2,l_3l_4}m_{l_1l_2}m_{l_3l_4},\\
&= A_{21}+A_{22}+A_{23}+A_{24},
\end{aligned} \tag{3.3}
$$

$$
\begin{aligned}
A_3 &= 3\sum_{l}\beta,\gamma\kappa_{l_1,l_2,l_3}\kappa_{l_4,l_5,l_6}m_{l_1l_2}m_{l_3l_4}m_{l_5l_6}\\
&\quad +2\sum_{l}\beta,\gamma\kappa_{l_1,l_2,l_3}\kappa_{l_4,l_5,l_6}m_{l_1l_4}m_{l_2l_5}m_{l_3l_6},\\
&= A_{31}+A_{32},
\end{aligned} \tag{3.4}
$$

sendo que os índices $l_1,\cdots,l_6$ variam sobre todos os componentes dos vetores β e γ, e $\sum_{\beta,\gamma}$ denota todas as possíveis combinações de $p+q$ parâmetros de $\beta_1,\cdots,\beta_p$ e $\gamma_1,\cdots,\gamma_q$.

Para o modelo (2.1), as parcelas de A_1, A_2 e A_3 foram desenvolvidos substituindo cumulantes de até quarta ordem. Portanto, um teste escore aperfeiçoado pode ser encontrado utilizando-se a estatística S_R^* e a distribuição χ_q^2 de referência ou usando a estatística S_R juntamente com os valores críticos corrigidos. Entretanto, a estatística escore aperfeiçoada (S_R^*) nem sempre é uma transformação monótona, assim, para solucionar esse problema Kakizawa (1996) sugeriu uma transformação monótona dada por $K=S_R^*+P(S_R)$, envolvendo a própria estatística escore e os coeficientes a, b e c, sendo $P(S_{R1}^*)$ dada por

$$
S_{R1}^*=\frac{1}{4}\left\{c^2S_R+2bcS_R^2+\left(2ac+\frac{4}{3}b^2\right)S_R^3+3abS_R^4+\frac{9}{5}S_R^5\right\}.
$$

Posteriormente, Cordeiro et al. (1998) também apresentaram uma fórmula para a estatística escore aperfeiçoada, de modo que também fosse uma transformação monótona de S_R. A estatística alternativa S_{R2}^* é expressa em termos da função da distribuição normal padrão Φ na forma

$$S_{R2}^* = \begin{cases} \sqrt{\frac{\pi}{3a}} \exp\left(\frac{b^2}{3a} - c\right) \times \\ \left\{\Phi\left(\sqrt{6a}S_R + \sqrt{\frac{2}{3a}}b\right) - \Phi\left(\sqrt{\frac{2}{3a}}b\right)\right\}, se\ a > 0, \\ \frac{1}{2b}\exp(-c)\left\{1 - \exp(-2bS_R)\right\},\ se\, a = 0\ e\ b \neq 0. \end{cases}$$

Observe que se $a = 0$ e $b = 0$, temos $S_{R2}^* = S_R(1-c)$, e não é necessário definir uma estatística escore alternativa.

3.3 Correção tipo–Bartlett à estatística escore envolvendo a média e o parâmetro de precisão

Consideraremos o modelo não–linear simétrico definido em (2.1) e as seguintes partições $\beta = \left(\beta_1^\top, \beta_2^\top\right)^T$, em que $\beta_1 = (\beta_1, \cdots, \beta_{p_1})^\top$, $\beta_2 = \left(\beta_{p_1+1}, \cdots, \beta_p\right)^\top$, $\gamma = \left(\gamma_1^\top, \gamma_2^\top\right)^T$, sendo $\gamma_1 = \left(\gamma_1, \cdots, \gamma_{q_1}\right)^\top$ e $\gamma_2 = \left(\gamma_{q_1+1}, \cdots, \gamma_q\right)^\top$, com $p_1 \leq p$ e $q_1 \leq q$. Tais decomposições induzem as seguintes partições $X = (X_1, X_2)$, $P = (P_1, P_2)$, $U = \left(U_1^\top, U_2^\top\right)^\top$, ou seja

$$U = \left(U_{\beta_1}^\top(\beta_1, \beta_2, \gamma_1, \gamma_2),\ U_{\beta_2}^\top(\beta_1, \beta_2, \gamma_1, \gamma_2),\ U_{\gamma_1}^\top(\beta_1, \beta_2, \gamma_1, \gamma_2),\ U_{\gamma_2}^\top(\beta_1, \beta_2, \gamma_1, \gamma_2)\right)^\top,$$

de modo que X_1, X_2, P_1 e P_2 são matrizes conhecidas de posto completo e dimensões $n \times p_1$, $n \times (p - p_1)$, $n \times q_1$ e $n \times (q - q_1)$, respectivamente.

O objetivo dessa seção é encontrar o fator de correção tipo–Bartlett para a estatística escore dada por (2.10). Estamos interessados em testar a hipótese $\mathscr{H}_0^1 : \beta_1 = \beta_1^{(0)}$, $\gamma_1 = \gamma_1^{(0)}$ contra $\mathscr{H}_1^1$: pelo menos uma das igualdades é violada, em que $\beta_1^{(0)}$ e $\gamma_1^{(0)}$ são vetores fixados de dimensões p_1 e q_1, respectivamente.

A matriz de informação total de Fisher correspondente é dada por

$$K = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} = \begin{pmatrix} K_{\beta_{11}} & 0 & K_{\beta_{12}} & 0 \\ 0 & K_{\gamma_{11}} & 0 & K_{\gamma_{12}} \\ K_{\beta_{21}} & 0 & K_{\beta_{22}} & 0 \\ 0 & K_{\gamma_{21}} & 0 & K_{\gamma_{22}} \end{pmatrix},$$

sendo que as matrizes $K_{\beta_{11}}$, $K_{\beta_{12}} = K_{\beta_{21}}^\top$, $K_{\beta_{22}}$, $K_{\gamma_{11}}$, $K_{\gamma_{12}} = K_{\gamma_{21}}^\top$, $K_{\gamma_{22}}$ estão definidas na Seção (2.8). Definimos, também, as matrizes

$$B = \begin{pmatrix} 0 & 0 \\ 0 & K_{22}^{-1} \end{pmatrix}, \; K_{22}^{-1} = \begin{pmatrix} K^{\beta_{22}} & 0 \\ 0 & K^{\gamma_{22}} \end{pmatrix},$$

e $M = K^{-1} - B$, para

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

Por simplicidade, apresentamos somente as expressões do fator tipo–Bartlett para os modelos não–lineares simétricos heteroscedásticos. Detalhes sobre o desenvolvimento destas deduções encontram-se no Apêndice A. Para os modelos não–lineares simétricos heteroscedásticos, definido em (2.1), os elementos $A_{11}, \cdots, A_{14}$, $A_{21}, \cdots, A_{24}$, A_{31} e A_{32} foram desenvolvidos substituindo os cumulantes apresentados na seção 3.2 nas expressões (3.2) até (3.4).

Após extensas manipulações algébricas, obtemos as seguintes expressões para os $A's$

$$\begin{aligned} A_{11} \;=\; & 3(\delta_{(2,0,0,0,0)})^{-1}(\iota^\top Q_2\Lambda_6(Z_\beta - Z_{2\beta})\Lambda_6 Q_2\iota) \\ & +b_1\Big\{-\left(\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}\right)\iota^\top\Lambda_4 Z_{2\beta d}(Z_\gamma - Z_{2\gamma})Z_{2\beta d}\Lambda_4\iota \\ & +(7 - 9\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)})\iota^\top\Lambda_4 Z_{2\beta d}(Z_\gamma - Z_{2\gamma})Z_{2\gamma d}\Lambda_1\iota \\ & +6(-1 + \delta_{(0,1,0,0,2)})\iota^\top\Lambda_1 Z_{2\beta d}(Z_\gamma - Z_{2\gamma})Z_{2\gamma d}\Lambda_2 1\Big\} \\ & +b_2\Big\{\left(7 - 9\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\right)\iota^\top\Lambda_1 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma})Z_{2\beta d}\Lambda_4\iota \\ & -6(-1 + \delta_{(0,1,0,0,2)})\iota^\top\Lambda_2 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma})Z_{2\beta d}\Lambda_4\iota\Big\} \\ & +b_{16}\iota^\top\Lambda_1 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma})Z_{2\gamma d}\Lambda_1\iota \;+b_{17}\iota^\top\Lambda_1 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma})Z_{2\gamma d}\Lambda_2\iota \\ & +b_{18}\iota^\top\Lambda_2 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma})Z_{2\gamma d}\Lambda_1\iota \;+b_{19}\Lambda_2 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma})Z_{2\gamma d}\Lambda_2\iota, \end{aligned} \tag{3.5}$$

$$\begin{aligned} A_{12} \;=\; & b_{10}\left\{\left((\delta_{(2,0,0,0,2)} - 1)/\delta_{(2,0,0,0,0)}\right)(\delta_{(3,0,0,0,1)} - \delta_{(0,1,0,0,0)})\iota^\top\Lambda_4 Z_{2\beta d}\right. \\ & *Z_{2\gamma}(Z_\beta - Z_{2\beta d})\Lambda_4\iota \\ & \left.+(1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)})\right\}\iota^\top\Lambda_4 Z_{2\beta d}Z_{2\gamma}(Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota \\ & -b_{11}\frac{\delta_{(2,0,0,0,2)}}{(\delta_{(2,0,0,0,2)} - 1)}\left\{(7 - 9\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\right)\iota^\top\Lambda_1 Z_{2\gamma d}Z_{2\gamma} \\ & *(Z_{\beta d} - Z_{2\beta d})\Lambda_4\iota \\ & \left.-6(\delta_{(0,1,0,0,2)} - 1)\iota^\top\Lambda_2 Z_{2\gamma d}Z_{2\gamma}(Z_{\beta d} - Z_{2\beta d})\Lambda_4\iota\right\} \\ & +b_9\left\{(7 - 9\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\right)\iota^\top\Lambda_1 Z_{2\gamma d}Z_{2\gamma}(Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota \\ & \left.+6(-1 + \delta_{(0,1,0,0,2)})\iota^\top\Lambda_2 Z_{2\gamma d}Z_{2\gamma}(Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota\right\}, \end{aligned} \tag{3.6}$$

$$\begin{aligned}
A_{13} &= -6\iota^\top\Lambda_6(Z_\beta - Z_{2\beta})(Z_\beta - Z_{2\beta})\Lambda_6 \odot J\iota \\
&+2b_1\left\{(\delta_{((1,1,0,0,1)} - \delta_{(3,0,0,0,1)})\,\iota^\top\Lambda_4 Z_{2\beta}\odot(Z_\beta - Z_{2\beta})\odot Z_{2\gamma}\Lambda_4\iota\right. \\
&+3(\delta_{(3,0,0,0,1)} - \delta_{(1,0,0,0,1)})\,\iota^\top\Lambda_4 Z_{2\beta}\odot(Z_\gamma - Z_{2\gamma})\odot Z_{2\beta}\Lambda_4\iota \\
&+b_5\ \mathbf{1}^\top\Lambda_1 Z_{2\gamma}\odot(Z_\gamma - Z_{2\gamma})\odot Z_{2\gamma}\Lambda_1\iota + b_6\iota^\top\Lambda_2 Z_{2\gamma}(Z_\gamma - Z_{2\gamma})Z_{2\gamma}\Lambda_2\iota \\
&+b_7\ \mathbf{1}^\top\Lambda_1 Z_{2\gamma}\odot(Z_\gamma - Z_{2\gamma})\odot Z_{2\gamma}\Lambda_2\iota \\
&+b_8\ \mathbf{1}^\top\Lambda_2 Z_{2\gamma}(Z_\gamma - Z_{2\gamma})Z_{2\gamma}\Lambda_1\iota,
\end{aligned} \tag{3.7}$$

$$\begin{aligned}
A_{14} &= b_{15}tr\left\{\Lambda_6 Z_{2\beta d}(Z_{\beta d} - Z_{2\beta d})\right\} \\
&+b_{12}\,tr\left\{\Lambda_7 Z_{2\gamma d}(Z_{\beta d} - Z_{2\beta d})\right\} \\
&-2b_{11}\,\delta_{(2,0,0,0,0)}tr\left\{\Lambda_8 Z_{2\gamma d}(Z_{\beta d} - Z_{2\beta d})\right\} \\
&+b_{13}\,tr\{\Lambda_9 Z_{2\gamma d}(Z_{\gamma d} - Z_{2\gamma d})\} \\
&-\frac{12(2-3\delta_{(2,0,0,0,2)} + \delta_{(3,0,0,0,3)})}{(\delta_{(2,0,0,0,2)} - 1)^2}tr\{\Lambda_3 Z_{2\gamma d}(Z_{\gamma d} - Z_{2\gamma d})\} \\
&+b_{14}\,tr\{\Lambda_7 Z_{2\beta d}(Z_{\gamma d} - Z_{2\gamma d})\},
\end{aligned} \tag{3.8}$$

$$\begin{aligned}
A_{21} &= 2b_{11}\,\iota^\top\Lambda_4(Z_{\beta d} - Z_{2\beta d})Z_{2\gamma}(Z_{\beta d} - Z_{2\beta d})\Lambda_4\iota \\
&+b_9\,(1-3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)})1^\top\Lambda_1(Z_{\gamma d} - Z_{2\gamma d})Z_\gamma \\
&(Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota \\
&-b_{10}\,(\delta_{(2,0,0,0,2)} - 1)(1-3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}) \\
&*\iota^\top\Lambda_1(Z_{\gamma d} - Z_{2\gamma d})Z_\gamma(Z_{\beta d} - Z_{2\beta d})\Lambda_4\iota,
\end{aligned} \tag{3.9}$$

$$\begin{aligned}
A_{22} &= 2b_1\,(\delta_{(3,0,0,0,1)} - \delta_{(0,1,0,0,0)})\iota^\top\Lambda_4 Z_{2\beta d}(Z_\gamma - Z_{2\gamma})(Z_{\beta d} - Z_{2\beta d})\Lambda_4\iota \\
&-\frac{4\,b_1\,\delta_{(2,0,0,0,0)}}{(\delta_{(2,0,0,0,2)} - 1)}(1-3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)})\iota^\top\Lambda_4 Z_{2\beta d}(Z_\gamma - Z_{2\gamma}) \\
&*(Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota \\
&-\frac{b_{11}\,\delta_{(2,0,0,0,0)}}{(\delta_{(2,0,0,0,2)} - 1)}\Big\{(1-9\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)})\iota^\top\Lambda_1 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma}) \\
&*(Z_{\beta d} - Z_{2\beta d})\Lambda_4\iota \\
&+2(-1+8\delta_{(0,1,0,0,2)})\iota^\top\Lambda_4 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma})(Z_{\beta d} - Z_{2\beta d})\Lambda_2\iota\Big\} \\
&-b_9\,\{(1-9\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)})\,\iota^\top\Lambda_1 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma})(Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota \\
&+2(-1+8\delta_{(0,1,0,0,2)})\iota^\top\Lambda_2 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma})(Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota\ \},
\end{aligned} \tag{3.10}$$

$$
\begin{aligned}
A_{23} &= b_{11}\Big\{(\delta_{(3,0,0,0,1)} - \delta_{(0,1,0,0,0)})\iota^\top\Lambda_4(Z_\beta - Z_{2\beta})\odot Z_{2\gamma}(Z_\beta - Z_{2\beta})\Lambda_4\iota \\
&\quad + 2\iota^\top\Lambda_4(Z_\beta - Z_{2\beta})\odot Z_{2\beta}(Z_\gamma - Z_{2\gamma})\Lambda_4\iota\Big\} \\
&\quad + 2b_9\iota^\top\Lambda_1(Z_\gamma - Z_\gamma)\odot Z_{2\gamma}(Z_\gamma - Z_\gamma)\Lambda_1\iota, \qquad (3.11)
\end{aligned}
$$

$$
\begin{aligned}
A_{24} &= \frac{3(\delta_{(4,0,0,0,0)} - 3\delta^2_{(0,1,0,0,0)})}{\delta^2_{(2,0,0,0,0)}} tr\left\{\Lambda_6(Z_{\beta d} - Z_{2\beta d})^2\right\} \\
&\quad + \frac{3\left(\delta_{(0,1,0,0,2)} + 2\delta_{(3,0,0,0,1)} + \delta_{(4,0,0,0,2)}\right)}{(\delta_{(2,0,0,0,2)} - 1)\delta_{(2,0,0,0,0)}} tr\left\{\Lambda_7(Z_{\beta d} - Z_{2\beta d})(Z_{\gamma d} - Z_{2\gamma d})\right\} \\
&\quad + \frac{3}{(\delta_{(2,0,0,0,2)} - 1)^2}\{-6 + \delta_{(2,0,0,0,2)}(12 - 3\delta_{(2,0,0,0,2)}) \\
&\quad + 4\delta_{(3,0,0,0,3)} + \delta_{(4,0,0,0,4)}\} tr\{\Lambda_9(Z_{\gamma d} - Z_{2\gamma d})^2\}, \qquad (3.12)
\end{aligned}
$$

$$
\begin{aligned}
A_{31} &= -\frac{b_{11}}{2\delta_{(2,0,0,0,2)}}(\delta_{(3,0,0,0,1)} - \delta_{(0,1,0,0,0)})\iota^\top\Lambda_5(Z_{\beta d} - Z_{2\beta d})(Z_\gamma - Z_{2\gamma}) \\
&\quad (Z_{\beta d} - Z_{2\beta d})\Lambda_5\iota \\
&\quad - \frac{b_{11}\delta_{(2,0,0,0,2)}}{(\delta_{(2,0,0,0,2)} - 1)}(-1 + 3\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)})\iota^\top\Lambda_1(Z_{\gamma d} - Z_{2\gamma d}) \\
&\quad (Z_\gamma - Z_{2\gamma})(Z_{\beta d} - Z_{2\beta d})\Lambda_4\iota \\
&\quad - \frac{b_{10}\,\delta_{(2,0,0,0,2)}}{18\,(\delta_{(2,0,0,0,2)} - 1)}(1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)})^2\iota^\top\Lambda_1(Z_{\beta d} - Z_{2\beta d}) \\
&\quad (Z_\gamma - Z_{2\gamma})(Z_{\gamma d} - Z_{2\gamma d})\Lambda_4\iota \\
&\quad - b_9\left(1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\right)\iota^\top\Lambda_1(Z_{\gamma d} - Z_{2\gamma d})(Z_\gamma - Z_{2\gamma}) \\
&\quad (Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota, \qquad (3.13)
\end{aligned}
$$

e

$$
\begin{aligned}
A_{32} &= -b_{11}(\delta_{(3,0,0,0,1)} - \delta_{(0,1,0,0,0)})\iota^\top\Lambda_4(Z_\beta - Z_{2\beta})\odot(Z_\gamma - Z_{2\gamma}) \\
&\quad \odot(Z_\beta - Z_{2\beta})\iota \\
&\quad - \frac{2b_9}{3}(1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)})\iota^\top\Lambda_1(Z_\gamma - Z_{2\gamma})\odot \\
&\quad (Z_\gamma - Z_{2\gamma})\odot(Z_\gamma - Z_{2\gamma})\iota. \qquad (3.14)
\end{aligned}
$$

Todas as matrizes envolvidas, são avaliadas sob o modelo restrito à hipótese nula $\mathscr{H}_0^1$, e são definidas do seguinte modo

$$
Z_\beta = X\left(K^{-1}_{\beta\beta}\right)X^\top = \delta^{-1}_{(2,0,0,0,0)}X\left(X^\top\Lambda X\right)^{-1}X^\top,
$$

se $p_1 \leq p$

$$Z_{2\beta} = X_2 \left(K_{\beta_{22}}^{-1}\right) X_2^{\top} = \delta_{(2,0,0,0,0)}^{-1} X_2 \left(X_2^{\top} \Lambda X_2\right)^{-1} X_2^{\top},$$

e

$$Z_{\gamma} = P \left(K_{\gamma\gamma}^{-1}\right) P^{\top} = 4\left(\delta_{(2,0,0,0,2)} - 1\right)^{-1} P \left(P^{\top} \Lambda_5 P\right)^{-1} P^{\top},$$

se $q_1 \leq q$

$$Z_{2\gamma} = P_2 \left(K_{\gamma_{22}}^{-1}\right) P_2^{\top} = 4\left(\delta_{(2,0,0,0,2)} - 1\right)^{-1} P_2 \left(P_2^{\top} \Lambda_5 P_2\right)^{-1} P_2^{\top},$$

que são as matrizes de covariância assintóticas de $X\widehat{\beta}$ correspondentes a Z_β e $Z_{2\beta}$ e $P\widehat{\gamma}$ retativo a Z_γ e $Z_{2\gamma}$. A notação $\odot$ indica a operação de produto direto entre matrizes e ι é um vetor $n \times 1$ de valores iguais a 1.

As matrizes diagonal $\Lambda' s$ presentes nas expressões acima são dadas por

$$\begin{aligned}
\Lambda_1 &= \operatorname{diag}\left\{\frac{\mathrm{h}_1'^3}{\phi_1{}^3}\right\}, \ \Lambda_2 = \operatorname{diag}\left\{\frac{\mathrm{h}_1'\mathrm{h}_1''}{\phi_1{}^2}\right\}, \\
\Lambda_3 &= \operatorname{diag}\left\{\frac{\mathrm{h}_1'\mathrm{h}_1''^2}{\phi_1{}^3}\right\}, \ \Lambda_4 = \operatorname{diag}\left\{\frac{\mathrm{h}_1'}{\phi_1{}^2}\right\}. \\
\Lambda_5 &= \left\{\frac{h_l'^2}{\phi_l{}^2}\right\}, \ \Lambda_6 = \operatorname{diag}\left\{\frac{1}{\phi_1{}^2}\right\}, \\
\Lambda_7 &= \operatorname{diag}\left\{\frac{\mathrm{h}_1'^2}{\phi_1{}^3}\right\}, \ \Lambda_8 = \operatorname{diag}\left\{\frac{\mathrm{h}_1''}{\phi_1{}^2}\right\}, \\
\Lambda_9 &= \operatorname{diag}\left\{\frac{\mathrm{h}_1'^4}{\phi_1{}^4}\right\}.
\end{aligned} \tag{3.15}$$

Com a finalidade de tornar mais atraente as expressões matricias de A_{11} até A_{32}, foram utilizadas substituições para as quantidades b_1 até b_{19} com os valores dados a seguir

$$b_1 = \frac{3(\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)})}{(\delta_{(2,0,0,0,2)}-1)\delta^2_{(2,0,0,0,0)}},\ b_2 = \frac{3(\delta_{(0,0,1,0,1)} - 2\delta_{(0,1,0,0,0)})}{(\delta_{(2,0,0,0,2)}-1)\delta^2_{(2,0,0,0,0)}},$$

$$b_3 = \frac{\left(7 - 9\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\right)}{(\delta_{(2,0,0,0,2)}-1)^3},\ b_4 = \frac{72(-1+\delta_{(0,1,0,0,2)})^2}{(\delta_{(2,0,0,0,2)}-1)},$$

$$b_5 = \frac{6}{(\delta_{(2,0,0,0,2)}-1)^3}\left(-5 + \delta_{(0,1,0,0,2)}(4 + 9\delta_{(0,0,1,0,3)} + 9\delta_{(0,1,0,0,2)} - 5\delta_{(0,0,1,0,3)})\right),$$

$$b_6 = \frac{-12}{(\delta_{(2,0,0,0,2)}-1)^3}(3\delta_{(0,1,0,0,2)} + 2 - 5\delta^2_{(0,1,0,0,2)}),$$

$$b_7 = \frac{12}{(\delta_{(2,0,0,0,2)}-1)^3}(-11\delta_{(0,1,0,0,2)} + 6 + 9\delta^2_{(0,1,0,0,2)}),$$

$$b_8 = \frac{6}{(\delta_{(2,0,0,0,2)}-1)^3}(7\delta_{(0,1,0,0,2)} + 5\delta^2_{(0,1,0,0,2)} + 5\delta_{(0,1,0,0,2)}\delta_{(0,0,1,0,3)}),$$

$$b_9 = \frac{-12\left(1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\right)}{(\delta_{(2,0,0,0,2)}-1)^3},$$

$$b_{10} = \frac{6(\delta_{(1,1,0,0,1)} - \delta_{(0,1,0,0,0)})}{(\delta_{(2,0,0,0,2)}-1)^2\delta_{(2,0,0,0,0)}},\ b_{11} = \frac{-6\left(\delta_{(3,0,0,0,1)} - \delta_{(0,1,0,0,0)}\right)}{\left(\delta_{(2,0,0,0,2)}-1\right)\delta^2_{(2,0,0,0,0)}},$$

$$b_{12} = \frac{-6(3\delta_{(2,0,0,0,0)} + 7\delta_{(3,0,0,0,1)} + \delta_{(4,0,0,0,2)} + \delta_{(2,1,0,0,2)})}{(\delta_{(2,0,0,0,2)}-1)\delta_{(2,0,0,0,0)}},$$

$$b_{13} = \frac{-3}{(\delta_{(2,0,0,0,2)}-1)^2}(40 - 23\delta_{(2,0,0,0,2)} - 2\delta^2_{(2,0,0,0,2)} + 7\delta_{(3,0,0,0,3)} + \delta_{(4,0,0,0,4)} + 2\delta_{(0,0,1,0,3)}$$
$$+2\delta_{(2,1,0,0,4)}),$$

$$b_{14} = \frac{-6(\delta_{(2,0,0,0,2)}\delta_{(0,1,0,0,2)} + 2\delta_{(3,0,0,0,1)} + \delta_{(4,0,0,0,2)})}{(\delta_{(2,0,0,0,2)}-1)\delta_{(2,0,0,0,2)}},\ b_{15} = 12\left(\frac{\delta_{(2,1,0,0,0)}}{\delta^2_{(2,0,0,0,0)}} + 1\right),$$

$$b_{16} = \frac{3}{(\delta_{(2,0,0,0,2)}-1)^3}\left\{1 + 2\delta_{(0,1,0,0,2)}(1 + \delta_{(0,0,1,0,3)}) + \delta_{(0,0,1,0,3)}(2 + \delta_{(0,0,1,0,3)}) + \delta^2_{(0,1,0,0,2)}\right\},$$

$$b_{17} = \frac{6}{(\delta_{(2,0,0,0,2)}-1)^3}\left\{\delta^2_{(0,1,0,0,2)} - 1 + \delta_{(0,0,1,0,3)}(1 - \delta_{(0,1,0,0,2)})\right\},$$

$$b_{18} = \frac{6}{(\delta_{(2,0,0,0,2)}-1)^3}\left\{\delta_{(0,1,0,0,2)}(\delta_{(0,1,0,0,2)} + \delta_{(0,0,1,0,3)}) - \delta_{(0,0,1,0,3)} - 1\right\},$$

e

$$b_{19} = \frac{12}{(\delta_{(2,0,0,0,2)}-1)^3}\left\{\delta_{(0,1,0,0,2)}(\delta_{(0,1,0,0,2)} - 2) + 1\right\}.$$

As expressões dos dependem da distribuição em estudo conforme apresentamos no Apêndice B.

Substituindo os $A's$ nas expressões de (3.2) a (3.4) e consequentemente na expressão (3.1), obteremos a estatística escore aperfeiçoada para testar $\mathscr{H}_0^1 : \beta_1 = \beta_1^{(0)}$, $\gamma_1 = \gamma_1^{(0)}$ contra $\mathscr{H}_1^1$: pelo menos uma das igualdades é violada, em que $\beta_1^{(0)}$ e $\gamma_1^{(0)}$ são vetores fixos de dimensões p_1 e q_1, respectivamente.

Para os modelos normais lineares heteroscedásticos temos que as matrizes $J_2 = Q_2 = 0$ e as quantidades A_1, A_2 e A_3 coincidem com os resultados de Fioresi (2000).

3.4 Correção tipo–Bartlett à estatística escore envolvendo o parâmetro de precisão

Nesta subseção temos como objetivo encontrar o fator de correção tipo Bartlett da estatística escore, em (2.12) para testar a hipótese nula $\mathscr{H}_0^2 : \gamma_1 = \gamma_1^{(0)}$ contra a alternativa $\mathscr{H}_1^2 : \gamma_1 \neq \gamma_1^{(0)}$, em que β e γ_2 são os vetores de parâmetros de perturbação e $\gamma_1^{(0)}$ é um valor especificado para γ_1. Para obter os coeficientes A_1, A_2 e A_3, através das quantidades em (3.2) a (3.4), são utilizadas as expresões (3.5) a (3.14) considerando que $Z_\beta = Z_{2\beta}$. Assim,

$$
\begin{aligned}
A_{11} \;=\;& b_1 \Big\{ -\left(\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)}\right) 1^\top \Lambda_4 Z_{2\beta d}(Z_\gamma - Z_{2\gamma}) Z_{2\beta d} \Lambda_4 \iota \\
&+ (7 - 9\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)})^\top \Lambda_4 Z_{2\beta d}(Z_\gamma - Z_{2\gamma}) Z_{2\gamma d} \Lambda_1 \iota \\
&+ 6(-1 + \delta_{(0,1,0,0,2)}) 1^\top \Lambda_1 Z_{2\beta d}(Z_\gamma - Z_{2\gamma}) Z_{2\gamma d} \Lambda_2 1 \Big\} \\
&+ b_2 \left\{ \left(7 - 9\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\right) \iota^\top \Lambda_1 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma}) Z_{2\beta d} \Lambda_4 \iota \right. \\
&\left. - 6(-1 + \delta_{(0,1,0,0,2)}) \iota^\top \Lambda_2 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma}) Z_{2\beta d} \Lambda_4 \iota \right\} \\
&+ b_{16} \iota^\top \Lambda_1 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma}) Z_{2\gamma d} \Lambda_1 \iota \\
&+ b_{17} \iota^\top \Lambda_1 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma}) Z_{2\gamma d} \Lambda_2 \iota \\
&+ b_{18} \iota^\top \Lambda_2 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma}) Z_{2\gamma d} \Lambda_1 \iota \\
&+ b_{19} \iota^\top \Lambda_2 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma}) Z_{2\gamma d} \Lambda_2 \iota,
\end{aligned}
$$

$$
\begin{aligned}
A_{12} \;=\;& b_{10}(1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)} \iota^\top \Lambda_4 Z_{2\beta d} Z_{2\gamma}(Z_{\gamma d} - Z_{2\gamma d}) \Lambda_1 \iota \\
&+ b_9 \left\{ \left(7 - 9\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\right) \iota^\top \Lambda_1 Z_{2\gamma d} Z_{2\gamma}(Z_{\gamma d} - Z_{2\gamma d}) \Lambda_1 \iota \right. \\
&\left. + 6(-1 + \delta_{(0,1,0,0,2)}) \iota^\top \Lambda_2 Z_{2\gamma d} Z_{2\gamma}(Z_{\gamma d} - Z_{2\gamma d}) \Lambda_1 \iota \right\},
\end{aligned}
$$

$$
\begin{aligned}
A_{13} \;=\;& b_{11}(\delta_{(0,0,1,0,1)} + 2\delta_{(0,1,0,0,0)})\iota^\top\Lambda_4 Z_{2\beta} \odot (Z_\gamma - Z_{2\gamma}) \odot Z_{2\beta}\Lambda_4\iota \\
& + b_5\,\iota^\top\Lambda_1 Z_{2\gamma} \odot (Z_\gamma - Z_{2\gamma}) \odot Z_{2\gamma}\Lambda_1 1 + a_6\iota^\top\Lambda_2 Z_{2\gamma}(Z_\gamma - Z_{2\gamma})Z_{2\gamma}\Lambda_2\iota \\
& + b_7\iota^\top\Lambda_1 Z_{2\gamma} \odot (Z_\gamma - Z_{2\gamma}) \odot Z_{2\gamma}\Lambda_2 1 \\
& + b_8\,\iota^\top\Lambda_2 Z_{2\gamma}(Z_\gamma - Z_{2\gamma})Z_{2\gamma}\Lambda_1\iota,
\end{aligned}
$$

$$
\begin{aligned}
A_{14} \;=\;& b_{13}\,tr\{\Lambda_9 Z_{2\gamma d}(Z_{\gamma d} - Z_{2\gamma d}) + a_{14}\,tr\{\Lambda_7 Z_{2\beta d}(Z_{\gamma d} - Z_{2\gamma d})\} \\
& - 12\frac{(2 - 3\delta_{(2,0,0,0,2)} + \delta_{(3,0,0,0,3)})}{(\delta_{(2,0,0,0,2)} - 1)^2}\,tr\{\Lambda_3 Z_{2\gamma d}(Z_{\gamma d} - Z_{2\gamma d})\},
\end{aligned}
$$

$$
\begin{aligned}
A_{21} \;=\;& b_9\,(1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)})\iota^\top\Lambda_1(Z_{\gamma d} - Z_{2\gamma d})Z_\gamma(Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota \\
& - b_{10}\,(\delta_{(2,0,0,0,2)} - 1)(1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}) \\
& * \iota^\top\Lambda_1(Z_{\gamma d} - Z_{2\gamma d})Z_\gamma(Z_{\beta d} - Z_{2\beta d})\Lambda_4\iota,
\end{aligned}
$$

$$
\begin{aligned}
A_{22} \;=\;& -\frac{4\,b_1\,\delta_{(2,0,0,0,0)}}{(\delta_{(2,0,0,0,2)} - 1)}(1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)})\iota^\top\Lambda_4 Z_{2\beta d}(Z_\gamma - Z_{2\gamma}) \\
& * (Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota \\
& - b_9\,\big\{(1 - 9\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)})\,\iota^\top\Lambda_1 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma})(Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota \\
& + 2(-1 + 8\delta_{(0,1,0,0,2)})\,\iota^\top\Lambda_2 Z_{2\gamma d}(Z_\gamma - Z_{2\gamma})(Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota\,\big\},
\end{aligned}
$$

$$
A_{23} \;=\; 2\,b_9\,\iota^\top\Lambda_1(Z_\gamma - Z_\gamma) \odot Z_{2\gamma}(Z_\gamma - Z_\gamma)\Lambda_1\iota,
$$

$$
\begin{aligned}
A_{24} \;=\;& \frac{3}{(\delta_{(2,0,0,0,2)} - 1)^2}\{-6 + \delta_{(2,0,0,0,2)}(12 - 3\delta_{(2,0,0,0,2)}) + 4\delta_{(3,0,0,0,3)} \\
& + \delta_{(4,0,0,0,4)}\,\}\,tr\{\Lambda_9(Z_{\gamma d} - Z_{2\gamma d})^2\},
\end{aligned}
$$

$$
\begin{aligned}
A_{31} \;=\;& -b_9\,\big(1 - 3\delta_{(0,1,0,0,2)} - \delta_{(0,0,1,0,3)}\,\big)\,\iota^\top\Lambda_1(Z_{\gamma d} - Z_{2\gamma d})(Z_\gamma - Z_{2\gamma}) \\
& * (Z_{\gamma d} - Z_{2\gamma d})\Lambda_1\iota,
\end{aligned}
$$

e

$$
\begin{aligned}
A_{32} &= \frac{-2b_9}{3}(1-3\delta_{(0,1,0,0,2)}-\delta_{(0,0,1,0,3)})\iota^\top\Lambda_1(Z_\gamma - Z_{2\gamma}) \\
&\odot(Z_\gamma - Z_{2\gamma})\odot(Z_\gamma - Z_{2\gamma})\iota.
\end{aligned}
$$

Todas as matrizes envolvidas estão definidas na seção (3.3), sendo avaliadas sob a hipótese $\mathscr{H}_0^2$. Substituindo os valores dos $A's$ nas expressões (3.2) e consequentemente na expressão (3.1), obteremos a estatística escore aperfeiçoada para testar $\mathscr{H}_0^2 : \gamma_1 = \gamma_1^{(0)}$ contra a alternativa $\mathscr{H}_1^2 : \gamma_1 \neq \gamma_1^{(0)}$, onde β é um vetor de parâmetros de perturbação e $\gamma_1^{(0)}$ é um valor especificado para γ_1.

As expressões dos $A\prime s$ de γ ,para os modelos lineares simétricos, coincidem com os resultados obtidos por Uribe-Opazo (1997); e para os modelos normais heteroscedásticos as quantidades $A_{1,\gamma}$, $A_{2,\gamma}$ e $A_{3,\gamma}$ coincidem com os resultados de Fioresi (2000), pág. 42, equações (3.23) a (3.32).

3.5 Correção tipo–Bartlett à estatística escore envolvendo a média

Nesta seção estamos interessados em testar a hipótese $\mathscr{H}_0^3 : \beta_1 = \beta_1^{(0)}$ contra $\mathscr{H}_1^3 : \beta_1 \neq \beta_1^{(0)}$ utilizando a estatística escore em (2.13). As quantidades em (3.3) a (3.5) para a obtenção dos coeficientes A_1, A_2 e A_3 são obtidas a partir das expressões (3.6) a (3.15) considerando que $Z_\gamma = Z_{2\gamma}$. Então os $A\prime s$ são dados a seguir

$$
A_{11} = 3(\delta_{(2,0,0,0,0)})^{-1}(\iota^\top Q_2\Lambda_6(Z_\beta - Z_\beta)\Lambda_6\,\iota),
$$

$$
\begin{aligned}
A_{12} &= b_{10}\frac{(\delta_{(2,0,0,0,2)}-1)}{\delta_{(2,0,0,0,0)}}(\delta_{(3,0,0,0,1)}-\delta_{(0,1,0,0,0)})\iota^\top\Lambda_4 Z_{2\beta d}Z_{2\gamma}(Z_\beta - Z_{2\beta d})\Lambda_4\iota \\
&-b_{11}\frac{\delta_{(2,0,0,0,2)}}{(\delta_{(2,0,0,0,2)}-1)}\Big\{(7-9\delta_{(0,1,0,0,2)}-\delta_{(0,0,1,0,3)})\iota^\top\Lambda_1 Z_{2\gamma d}Z_{2\gamma} \\
&*(Z_{\beta d} - Z_{2\beta d})\Lambda_4\iota \\
&-6(\delta_{(0,1,0,0,2)}-1)\iota^\top\Lambda_2 Z_{2\gamma d}Z_{2\gamma}(Z_{\beta d} - Z_{2\beta d})\Lambda_4\iota\Big\},
\end{aligned}
$$

$$
\begin{aligned}
A_{13} &= 2b_1\Big\{(\delta_{((1,1,0,0,1)}-\delta_{(3,0,0,0,1)})\iota^\top\Lambda_4 Z_{2\beta}\odot(Z_\beta - Z_{2\beta})\odot Z_{2\gamma}\Lambda_4\iota\Big\} \\
&-6\iota^\top\Lambda_6(Z_\beta - Z_{2\beta})\Lambda_6\otimes J\iota,
\end{aligned}
$$

$$
\begin{aligned}
A_{14} &= b_{15} tr\left\{\Lambda_6 Z_{2\beta d}(Z_{\beta d}-Z_{2\beta d})\right\}+b_{12}\, tr\left\{\Lambda_7 Z_{2\gamma d}(Z_{\beta d}-Z_{2\beta d})\right\}\\
&\quad -2b_{11}\,\delta_{(2,0,0,0,0)} tr\left\{\Lambda_8 Z_{2\gamma d}(Z_{\beta d}-Z_{2\beta d})\right\},
\end{aligned}
$$

$$
A_{21} = 2b_{11}\,\iota^\top \Lambda_4 (Z_{\beta d}-Z_{2\beta d}) Z_{2\gamma}(Z_{\beta d}-Z_{2\beta d})\Lambda_4 \iota,\ A_{22}=0,
$$

$$
\begin{aligned}
A_{23} &= b_{11}\{(\delta_{(3,0,0,0,1)}-\delta_{(0,1,0,0,0)})\iota^\top \Lambda_4 (Z_\beta - Z_{2\beta})\odot\\
&\quad * Z_{2\gamma}(Z_\beta - Z_{2\beta})\Lambda_4 \iota,
\end{aligned}
$$

$$
A_{24} = \frac{3\left(\delta_{(4,0,0,0,0)}-3\delta^2_{(0,1,0,0,0)}\right)}{\delta^2_{(2,0,0,0,0)}} tr\left\{\Lambda_6 (Z_{\beta d}-Z_{2\beta d})^2\right\},\ e\ A_{31}=A_{32}=0.
$$

Para os modelos não–lineares simétricos homoscedásticos as quantidades dos $A\prime s$ coincidem com os resultados obtidos por Cysneiros et al. (2008) e com a tese de doutorado de Uribe-Opazo (1997) para o caso de modelos lineares simétricos homoscedásticos.

4 Estudo de Simulação

Apresentaremos alguns resultados de simulações para avaliar a eficácia da correção tipo–Bartlett para os testes escore, em modelos não lineares simétricos heteroscedásticos. Comparamos os desempenhos de quatro estatísticas de testes, isto é, o teste escore usual S_R, escore corrigido via correção tipo–Bartlett $S_R{}^*$ e suas versões alternativas S_{R2}^* e S_{R1}^*. Os desempenhos são avaliados em função da proximidade das probabilidades de rejeição da hipótese nula, sendo verdadeira (probalidade de ocorrer o erro tipo I), aos respectivos níveis nominais dos testes.

A componente sistemática utilizada em todas as simulações é dada por

$$\mu_l = \beta_0 + \beta_1 x_1 + \exp(\beta_2 x_2),$$

com $h(z_l{}^\top\gamma) = \exp(z_l{}^\top\gamma)$, sendo $z_l{}^\top\gamma = \tau$ e $\tau = \gamma_1 + \gamma_2 z_2 + \gamma_3 z_3$. A hipótese nula considerada e $\mathscr{H}_0^2 : \gamma_1 = \gamma_1^{(0)}$ e a hipótese alternativa é $\mathscr{H}_1^2 : \gamma_1 \neq \gamma_1^{(0)}$, para $\gamma_1^{(0)}$ o vetor de parâmetros de interesse de dimensão $q_1 \times 1$. Sem perda de generalidade, tomamos os seguintes valores para os parâmetros da regressão: $\beta_1 = 5$, $\beta_2 = 2$, $\beta_3 = 1$, $\gamma_1 = 0$, $\gamma_2 = 0,3$ e $\gamma_3 = 0,5$. Todas as variáveis em X e Z são independentes e geradas de uma distribuição $U(0,1)$.

O número de réplicas foi fixado em 10.000 para tamanhos de amostras n = 20, 30, 40, 50, e foram considerados os níveis nominais $\alpha = 10\%, 5\%, 1\%, 0,5\%$. As simulações foram realizadas utilizando a linguagem de programação matricial `Ox` (Doornik, 2001)

Para cada tamanho de amostra e cada nível considerado, foram estimadas as taxas de rejeição de cada teste, ou seja, $P(S_R \geq x_\alpha)$, $P(S_R{}^* \geq x_\alpha)$,$P(S_{R1}{}^* \geq x_\alpha)$ $P(S_{R2}{}^* \geq x_\alpha)$ sendo x o quantil $(1-\alpha)$ apropriado da distribuição χ_q^2. Todas as entradas das tabelas apresentadas correspondem a porcentagens, exceto a Tabela 4.7 que faz comparações da média e variância, para diferentes modelos. O estudo de simulação visa analisar o melhoramento da estatística escore S_R, nos modelos não–lineares simétricos heteroscedásticos, quando aplicado a ela um fator de correção tipo–Bartlett. Foram realizadas simulações para os seguintes modelos de distribuições: normal, logístico tipo I e exponencial potência para k=0,1.

A Tabela 4.1 apresenta as taxas de rejeição para diferentes tamanhos de amostras $n = 20, 30, 40, 50$, fixando-se os valores dos parâmetros de interesse $p = 1$ e perturbação $q = 5$. Nota-se que as taxas de rejeição dos testes baseados nas estatísticas S_R, S_R^*, S_{R1}^* e S_{R2}^* se aproximam dos níveis nominais, ao passo que aumenta o tamanho da amostra, conforme o previsto; e ainda, as taxas de rejeição da estatística alternativa S_{R2}^* estão mais próximas do nível nominal do que das outras estatísticas envolvidas no teste. Deste modo, o teste baseado na estatística S_{R2}^* para o modelo normal não–linear heteroscedástico apresentou o melhor resultado em relação aos outros testes estatísticos.

Na Tabela 4.2 são apresentadas as taxas de rejeição para diferentes tamanhos de amostras $n = 25, 30, 35, 40, 45, 100$, relativo à distribuições logística tipo I, fixando-se os valores dos parâmetros de interesse $p = 1$ e dos parâmetros de perturbação $q = 5$ e diferentes níveis nominais α. Para todos os valores de $\alpha = (10\%, 5\%, 1\%, 0,5\%)$, observamos que as taxas de rejeição do teste baseados em S_R^* são bem mais próximas dos correspondentes níveis nominais do que as dos testes baseados em S_R. A tabela mostra que para tamanho de amostra pequeno ou mesmo moderado todos os valores dos testes apresentaram-se conservativos. Os teste baseado na estatística corrigida obtiveram seus valores mais próximos do nível nominal fixado do que os testes baseados na estatística usual.

A Tabela 4.3 mostra resultados de simulações, para os modelos não–lineares logístico tipo I e exponencial potência, considerando a hipótese alternativa para os valores de n, citados anteriormente, e diferentes valores para γ foram considerados com níveis nominais de 5% e 10%. A análise desta tabela mostra que comparando os testes baseado nas estatísticas S_{R1}^* e S_{R2}^* com os demais testes,para 5% e 10%, eles obtiveram seus valores mais próximos do nível nominal.

Nas Tabelas 4.4–4.6 mostramos os resultados de simulações, para os modelos não–lineares normal, logístico tipo I, ambos para tamanho de amostra $n = 40$ e exponencial potência para $n = 30$, fixando-se os valores dos parâmetros de interesse $p = 1$ e perturbação $q = 5$ e diferentes valores para γ foram considerados com níveis nominais de 5% e 10%. A análise destas tabelas mostraram que comparando o poder dos testes baseados nas estatísticas S_R, S_R^*, S_{R1}^* e S_{R2}^* os resultados mostram que não há nenhuma perda de poder decorrente de se utilizar o fator de correção tipo–Bartlett. Todavia, o poder dos testes para as diferentes estatísticas, em análise, parecem semelhantes, com pequenas distorções de tamanho.

Na Tabela 4.7 apresentamos comparações da média e variância, para os modelos:

normal não–linear com $n = 40$, $p = 1$ e $q = 5$, logístico tipo I não–linear com $n = 30$, $p = 1$ e $q = 5$, e exponencial potência não–linear com $k = 0.1$, $n = 40$, $p = 1$ e $q = 5$, das estatísticas S_R, ${S_R}^*$, S^*_{R2}, S^*_{R1} e da distribuição χ^2_1. Os resultados dessas tabelas mostram que todas as estatísticas obtiveram médias e variâncias próximas à da distribuição χ^2_1. Entretanto, as estatísticas referentes a distribuição normal obtiveram todos os valores das médias iguais e variâncias próximas à da distribuição χ^2_1.

Tabela 4.1: Tamanho dos testes – modelo normal não–linear com $p = 1$ e diversos valores para (n, α).

	Nível nominal	Estatística de teste			
n	$\alpha(\%)$	S_R	${S_R}^*$	S^*_{R1}	S^*_{R2}
	10,0	7,3	8,4	8,7	8,9
	5,0	3,1	4,1	4,1	4,3
20	1,0	0,5	0,7	0,7	0,8
	0,5	0,2	0,4	0,4	0,4
	10,0	8,2	9,0	9,1	9,1
	5,0	3,9	4,3	4,3	4,4
30	1,0	0,7	0,8	0,8	0,8
	0,5	0,4	0,4	0,4	0,4
	10,0	8,8	9,7	9,7	9,8
	5,0	4,5	4,8	4,8	4,8
40	1,0	0,8	0,8	0,8	0,9
	0,5	0,5	0,4	0,4	0,4
	10,0	8,6	9,8	9,8	9,9
	5,0	3,8	4,6	4,6	4,6
50	1,0	0,6	0,8	0,8	0,8
	0,5	0,3	0,4	0,4	0,4

Tabela 4.2: Tamanho dos testes – modelo logístico tipo I não–linear com $p = 1$ e diversos valores para (n, α).

	Nível nominal	Estatística de teste			
n	$\alpha(\%)$	S_R	$S_R{}^*$	S^*_{R1}	S^*_{R2}
	10,0	5,2	6,2	6,4	6,3
30	5,0	2,3	2,7	2,8	2,8
	10,0	5,2	7,2	7,4	7,3
35	5,0	2,3	3,2	3,3	3,3
	10,0	6,4	7,7	7,9	7,8
40	5,0	2,7	3,5	3,5	3,5
	10,0	6,5	8,0	8,1	8,1
45	5,0	2,8	3,6	3,6	3,6
	10,0	8,9	9,9	10,0	9,9
100	5,0	4,1	4,5	4,6	4,6

Tabela 4.3: Tamanho dos testes – modelos não–lineares logístico tipo I e exponencial potência, respectivamente, com $k = 0,1$ $p = 1$ $n = 40$ e α.

	10%		5%		1%		0.5%	
S_R	6,4	8,6	2,7	4,2	0,3	0,7	0,2	0,5
$S_R{}^*$	7,7	9,6	3,5	5,0	0,5	1,3	0,2	0,7
S^*_{R2}	7,9	10,0	3,5	5,0	0,5	1,4	0,2	0,7
S^*_{R1}	7,8	10,0	3,5	5,0	0,5	1,3	0,2	0,7

Tabela 4.4: Poder dos testes – modelo normal não–linear com $n = 40$, $p = 1$, $\alpha = 5\%$ e $\alpha = 10\%$, respectivamente.

γ	S_R	$S_R{}^*$	S^*_{R1}	S^*_{R2}	S_R	$S_R{}^*$	S^*_{R1}	S^*_{R2}
0,0	4,5	4,8	4,8	4,8	8,8	9,7	9,7	9,8
1,0	17,3	18,3	18,1	18,1	26,4	27,0	27,1	27,1
2,0	60,8	62,8	61,8	61,8	73,3	73,7	73,7	73,8
3,0	88,7	88,3	88,2	88,3	88,7	88,2	88,3	88,3
4,0	99,0	99,9	99,9	98,9	99,6	99,5	99,5	99,6

Tabela 4.5: Poder dos testes – modelo logístico tipo I não–linear com $n = 30$, $p = 1$, $\alpha = 5\%$ e $\alpha = 10\%$, respectivamente.

γ	S_R	$S_R{}^*$	S^*_{R1}	S^*_{R2}	S_R	$S_R{}^*$	S^*_{R1}	S^*_{R2}
$0,0$	$2,5$	$4,9$	$4,0$	$4,3$	$5,5$	$8,9$	$9,3$	$9,9$
$1,0$	$6,5$	$9,1$	$8,8$	$9,0$	$13,1$	$16,2$	$16,4$	$16,6$
$2,0$	$17,8$	$16,4$	$16,1$	$16,3$	$29,6$	$28,8$	$29,0$	$29,3$
$3,0$	$44,5$	$58,3$	$57,5$	$58,3$	$60,2$	$69,8$	$70,5$	$70,5$
$4,0$	$82,0$	$86,7$	$86,3$	$86,5$	$89,3$	$91,7$	$91,8$	$91,9$
$5,0$	$99,8$	$99,8$	$99,8$	$99,8$	$99,3$	$99,5$	$99,5$	$99,6$

Tabela 4.6: Poder dos testes – modelo exponencial potência não–linear com $k = 0,1$ $p = 1$ $n = 40$, $\alpha = 5\%, 10\%$, respectivamente.

γ	S_R	$S_R{}^*$	S^*_{R1}	S^*_{R2}	S_R	$S_R{}^*$	S^*_{R1}	S^*_{R2}
$0,0$	$4,2$	$5,0$	$5,0$	$5,0$	$8,6$	$9,9$	$10,0$	$10,0$
$1,0$	$15,1$	$16,9$	$16,8$	$26,9$	$25,3$	$26,7$	$26,9$	$26,9$
$2,0$	$49,4$	$54,5$	$54,1$	$54,3$	$63,1$	$66,5$	$66,7$	$66,9$
$3,0$	$84,6$	$84,6$	$85,5$	$84,5$	$90,9$	$90,8$	$90,8$	$90,8$
$4,0$	$97,6$	$97,8$	$97,8$	$97,8$	$99,9$	$99,0$	$99,0$	$99,0$

Tabela 4.7: Médias e variâncias – modelos não–lineares normal, logístico tipo I e exponencial potência, respectivamente.

Momentos	χ^2_1	S_R			$S_R{}^*$			S^*_{R1}			S^*_{R2}		
Média	1,0	1,0	0,9	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0	1,0
Variância	2,0	1,9	1,8	1,9	1,9	2,2	2,1	1,9	1,9	2,0	1,9	1,9	2,0

5 Conclusões

A principal contribuição teórica desta dissertação foi a obtenção, em notação matricial, de um fator de correção tipo–Bartlett para a estatística escore na classe dos modelos não–lineares simétricos heteroscedásticos generalizando assim os resultados de Fioresi (2000), Uribe-Opazo et al. (1999) e Cysneiros et al. (2008). Várias estatísticas de testes de hipóteses foram propostas considerando funções de ligação quaisquer para a média e variância.

Além dessa contribuição teórica, analisamos estudos de simulação para verificar o desempenho das estatísticas de teste S_R, S_R^*, S_{R1}^* e S_{R2}^*, relembrando que S_{R1}^* e S_{R2}^* são as estatísticas escore propostas por Kakizawa (1996) e a estatística escore proposta por Cordeiro et al. (1998), respectivamente. Os resultados de simulação mostraram que os testes baseados nas estatísticas escore corrigidas via correção tipo–Bartlett (S_R^*, S_{R1}^* e S_{R2}^*) apresentaram melhor performance do que a estatística escore não corrigida, uma vez que suas taxas de rejeição foram mais próximas dos níveis nominais.

Dentre os quatro testes apresentados, o teste baseado na estatística S_{R2}^* obteve, sempre, o melhor desempenho, apresentando taxas de rejeição mais próximas dos níveis nominais, concluindo ser o melhor teste para o modelo em estudo.

APÊNDICE A

Neste apêndice apresentamos a obtenção de alguns cumulantes conjuntos de derivadas do logaritmo da função de verossimilhança nos modelos não–lineares simétricos heteroscedásticos definidos em (2.1), necessários aos cálculos das correções tipo–Bartlett para a estatística escore S_R.

Seja o $\ell(\beta,\gamma)$ o logaritmo da função de verossimilhança do vetor de parâmetros $\theta = \left(\beta^\top,\gamma^\top\right)^\top$, dado o vetor de observações $(y_1,\cdots,y_n)$ nos modelos não–lineares simétricos heteroscedásticos em (2.1), que tem a forma

$$\ell(\theta) = -\frac{1}{2}\sum_{l=1}^{n} log\phi_l + \sum_{l=1}^{n} t\left(z_l\right)$$

em que $t\left(z_l\right) = \log g\left(z_l^2\right)$, com $z_l = \sqrt{u_l}$ e $u \geq 0, = \frac{y_l-\mu_l}{\sqrt{\phi_l}}$, $\mu_l = f(x_l,\beta)$, $\phi_l = h(\tau_l)$, com $\tau_l = z_l^\top\gamma$ e para $l = 1,\cdots,n$. Diferenciando $\ell(\theta)$ temos os seguintes resultados

$$U_r = \frac{\partial\ell(\theta)}{\partial\beta_r} = -\sum_{l=1}^{n}\frac{1}{\sqrt{\phi_l}}\left(r\right)_l,$$

$$U_{rs} = \frac{\partial^2\ell(\theta)}{\partial\beta_r\partial\beta_s} = \sum_{l=1}^{n} t_{(z_l)}^{(2)}\cdot\frac{1}{\phi_l}\left(r,s\right)_l - \sum_{l=1}^{n} t_{(z_l)}^{(1)}\frac{1}{\sqrt{\phi_l}}\left(rs\right)_l,$$

$$\begin{aligned} U_{rst} = \frac{\partial^3\ell(\theta)}{\partial\beta_r\partial\beta_s\partial\beta_\top} &= -\sum_{l=1}^{n} t_{(z_l)}^{(3)}\frac{1}{\phi_l^{\frac{3}{2}}}\left(r,s,t\right)_l \\ &+ \sum_{l=1}^{n} t_{(z_l)}^{(2)}\frac{1}{\phi_l}\left[(rs,t)+(r,st)+(rt,s)\right]_l \\ &- \sum_{l=1}^{n} t_{(z_l)}^{(1)}\frac{1}{\sqrt{\phi_l}}\left(rst\right)_l, \end{aligned}$$

$$U_R = \frac{\partial\ell(\theta)}{\partial\gamma_R} = -\frac{1}{2}\sum_{l=1}^{n}\frac{h_l^{'}}{\phi_l}\left(R\right)_l - \frac{1}{2}\sum_{l=1}^{n} t_{(z_l)}^{(1)} z_l\frac{h_i^{'}}{\phi_l}\left(R\right)_l,$$

$$
\begin{aligned}
U_{RST} &= \frac{\partial^3\ell(\theta)}{\partial\gamma_R\partial\gamma_S\partial\gamma_T} \\
&= \{-\sum_{l=1}^{n}\frac{h_l^{'3}}{\phi_l^3}+\frac{3}{2}\sum_{l-1}^{n}\frac{h_l^{'}h_l^{''}}{\phi_l^2}-\frac{1}{2}\sum_{l-1}^{n}\frac{h_l^{(3)}}{\phi_l}-\frac{15}{8}\sum_{l=1}^{n}\frac{h_l^{'3}}{\phi_l^3}t^{(1)}(z_l)z_l \\
&+ \frac{9}{4}\sum_{l-1}^{n}\frac{h_l^{'}h_l^{''}}{\phi_l^2}t^{(1)}(z_l)z_l-\frac{1}{2}\sum_{l-1}^{n}\frac{h_l^{(3)}}{\phi_l}t^{(1)}(z_l)z_l-\frac{9}{8}\sum_{l=1}^{n}\frac{h_l^{'3}}{\phi_l^3}t^{(2)}(z_l)z_l^2 \\
&+ \frac{3}{4}\sum_{l-1}^{n}\frac{h_l^{'}h_l^{''}}{\phi_l}t^{(2)}(z_l)z_l^2-\frac{1}{8}\sum_{l=1}^{n}\frac{h_l^{'3}}{\phi_l^3}t^{(3)}(z_l)z_l^3\}(R,S,T)_l,
\end{aligned}
$$

$$
U_{Rs} = \frac{\partial^2\ell(\theta)}{\partial\gamma_R\partial\beta_s}=\frac{1}{2}\sum_{l=1}^{n}t_{(z_l)}^{(2)}z_l\frac{h_l^{'}}{\phi_l^{\frac{3}{2}}}(R,s)_l+\frac{1}{2}\sum_{l=1}^{n}t_{(z_l)}^{(1)}\frac{h_l^{'}}{\phi_l^{\frac{3}{2}}}(R,s)_l.
$$

Os cumulantes dessas derivadas são dados do seguinte modo: $\kappa_r = E(U_r)$, $\kappa_{rs} = E(U_{rs})$, $\kappa_{rS} = E(U_{rs})$, $\kappa_{rs,TU} = E(U_{rs}U_{TU}) - \kappa_{rs}\kappa_{TU}$, etc. Deste modo, apresentaremos alguns dos resultados dos cumulantes, referentes aos modelos de distribuições não–lineares simétricos heteroscedásticos, utilizados na correção tipo–Bartlett, temos

$$
\begin{aligned}
\kappa_{rs} &= \delta_{(0,1,0,0,0)}\sum_l\frac{(r,s)_l}{\phi_l}, \ \ \kappa_{rS}=0 \\
\kappa_{rst} &= \delta_{(0,1,0,0,0)}\sum_l\frac{1}{\phi_l}[(rs,t),(rt,s),(st,r)]_l, \\
\kappa_{RS} &= \frac{\left(\delta_{(0,1,0,0,0)}-1\right)}{4}\sum_l\frac{h_l'^2}{\phi_l{}^2}z_{iR}z_{iS}, \\
\kappa_{Rst} &= -\frac{\left(-2\delta_{(0,1,0,0,0)}-\delta_{(0,0,1,0,1)}\right)}{2}\sum_l\frac{h_l'}{\phi_l{}^2}(r,s)_l z_{iR}, \\
\kappa_{RS}^{(T)} &= \frac{\partial\kappa_{RS}}{\partial\gamma_T}=\frac{\left(\delta_{(0,1,0,0,2)}-1\right)}{2}\sum_l\frac{\left(\phi_l h_l'h_l''-h_l'^3\right)}{\phi_l{}^3}z_{lR}z_{lS}z_{lT},
\end{aligned}
$$

$$
\begin{aligned}
\kappa_{rs}^{(t)} &= \frac{\partial \kappa_{rs}}{\partial \beta_{\top}} = \delta_{(0,1,0,0,0)} \sum_{l} \frac{1}{\phi_l}[(rt,s)+(r,st)]_l,\\
\kappa_{rsT} &= \frac{\left(\delta_{(1,1,0,0,1)}-\delta_{(0,1,0,0,0)}\right)}{2} \sum_{l} \frac{h_l'}{\phi_l{}^2}(r,s)_l z_{lT},\\
\kappa_{(r,sT)} &= -\frac{\left(\delta_{(1,1,0,0,1)}-\delta_{(0,1,0,0,0)}\right)}{2} \sum_{l} \frac{h_l'}{\phi_l{}^2}(r,s)_l z_{lT},\\
\kappa_{(R,s,t)} &= -\frac{\left(\delta_{(3,0,0,0,1)}-\delta_{(0,1,0,0,0)}\right)}{2} \sum_{l} \frac{h_l'}{\phi_l{}^2}(s,t)_l z_{lR},\\
\kappa_{(R,ST)} &= \frac{\left(-3+\delta_{(0,0,1,0,3)}\right)}{8} \sum_{l} \frac{h_l'^3}{\phi_l{}^3} + \frac{\left(2+5\delta_{(0,1,0,0,2)}\right)}{8} \sum_{l} \frac{h_l' h_l''}{\phi_l{}^2} z_{lR} z_{lS} z_{lT},\\
\kappa_{(r,Ts)} &= -\frac{\left(\delta_{(1,1,0,0,1)}+\delta_{(2,0,0,0,0)}\right)}{2} \sum_{l} \frac{h_l'}{\phi_l{}^2}(r,s)_l z_{lT},\\
\kappa_{(rSt)} &= \left\{-\frac{\delta_{(0,0,1,0,1)}}{2} \sum_{l} \frac{h_l'}{\phi_l{}^2} + \delta_{(2,0,0,0,0)} \sum_{l} \frac{h_l'}{\phi_l{}^2}\right\}(r,t)_l z_{lS},\\
\kappa_{(rs,T)} &= -\frac{\left(\delta_{(0,1,0,0,0)}+\delta_{(1,1,0,0,0)}\right)}{2} \sum_{l} \frac{h_l'}{\phi_l{}^2}(r,s)_l z_{lT},\\
\kappa_{r,s,T,U} &= \frac{1}{4}\left\{\delta_{(2,0,0,0,0)}\delta_{(0,1,0,0,2)}+2\delta_{(3,0,0,0,1)}+\delta_{(4,0,0,0,2)}\right\} \sum_{l} \frac{h_l'^2}{\phi_l{}^3}(r,s)_l z_{lT} z_{lU},\\
\kappa_{r,s,TU} &= \{ \frac{\delta_{(2,0,0,0,0)}}{2} \sum_{l} \left(\frac{-h_l''}{\phi_l{}^2} + \frac{h_l'^2}{2\phi_l{}^3}\right) - \frac{\delta_{(3,0,0,0,1)}}{2} \sum_{l} \frac{h_l''}{\phi_l{}^2}\\
&\quad + \frac{\delta_{(2,1,0,0,2)}+3\delta_{(3,0,0,0,1)}}{4} \sum_{l} \frac{h_l'^2}{\phi_l{}^3} + \frac{\delta_{(2,0,0,0,0)}\delta_{(2,0,0,0,2)}}{4} \sum_{l} \frac{h_l'^2}{\phi_l{}^3} \}(r,s)_l z_{lT} z_{lU},\\
\kappa_{rSt} &= -\frac{\left(\delta_{(0,1,0,0,0)}-\delta_{(1,1,0,0,1)}\right)}{2} \sum_{l=1}^{n} \frac{h_l'}{\phi_l^2}(r,S,t)_l,\\
e\ \kappa_{R,ST} &= \{ \frac{\left(-3+\delta_{(0,0,1,0,3)}\right)}{8} \sum_{l=1}^{n} \frac{h_l'^3}{\phi_l^3} + \frac{\left(2+5\delta_{(0,1,0,0,2)}\right)}{8} \sum_{l=1}^{n} \frac{h_l' h_l''}{\phi_l^2} \}(R,S,T)_l.
\end{aligned}
$$

APÊNDICE B

Neste apêndice serão apresentados os valores dos $\delta' s$, associados à correção tipo–Bartlett para cada distribuição pertencente à classe dos modelos simétricos. As notações $t^k(z_l)$ e $\delta_{(a,b,c,d,e)}$ estão definidas, respectivamente, nas Seções 2.2 e 2.5. Em nosso texto, consideramos $z_l = \frac{y_l - \mu_l}{\sqrt{\phi_l}}$, com $\mu_l = f(x_l;\beta)$.

B.1 Distribuição Normal

Seja $y_l \sim N(\mu_l, \phi_l)$ com função de densidade da forma

$$\pi(y_l, \beta_l, \phi_l) = \frac{1}{\sqrt{\phi_l 2\pi}} \exp\left\{-\frac{1}{2}\frac{(y_l - \mu_l)^2}{\phi_l}\right\}, y \in \mathbb{R},$$

para $\beta \in \mathbb{R}^p$ e $\phi_l = h(z_l^\top \gamma)$, com $\phi_l > 0$. O logaritmo da função de verossimilhança é dado por (2.2). Assim, $t(z_l) = -z_l^2/2$, sendo $g(z_l)$ dada na Tabela 2.1. Deste modo as quatro primeiras derivadas de t com relação a z_l são

$$t^{(1)}_{(z_l)} = -z, \; t^{(2)}_{(z_l)} = -1, \; t^{(3)}_{(z_l)} = t^{(4)}_{(z_l)} = 0.$$

Usando a notação $\delta_{(a,b,c,d,e)}$, dada na equação (2.6), segue que

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

B.2 Distribuição de Cauchy

Seja $y_l \sim C(\mu_l, \phi_l)$ com função densidade dada da seguinte forma

$$\pi(y_l, \phi_l) = \frac{1}{\pi\sqrt{\phi_l}}\left[1 + \frac{(y_l - \mu_l)^2}{\phi_l}\right]^{-1}, y_l \in \mathbb{R},$$

para $\beta \in \mathbb{R}^p$ e $\phi_l = h(z_l^\top \gamma)$, com $\phi_l > 0$. O logaritmo da função de verossimilhança é dado por (2.2), assim $t(z_l) = \log\left(\frac{1}{\pi}\left[1+z^2\right]^{-1}\right)$, sendo $g(z_l)$ e s dados na Tabela 2.1. Deste modo as quatro primeiras derivadas de t com relação à z_l são

$$\begin{aligned}
&t^{(1)}_{(Z_l)} = 2zs,\ t^{(2)}_{(Z_l)} = 2s + 4z^2s^2,\\
&t^{(3)}_{(Z_l)} = 12zs^2 + 16z^3s^3,\ e\ t^{(4)}_{(Z_l)} = 12s^2 + 96z^2s^2 + 96z^4s^4.
\end{aligned}$$

Usando a notação $\delta_{(a,b,c,d,e)}$, dada na equação (2.6), segue que

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

B.3 Distribuição t de Student

Seja $y_l \sim t(\mu_l, \phi_l, \nu)$ com a função densidade

$$\pi(y_l, \mu_l, \phi_l, \nu) = \frac{\nu^{\nu/2}}{\sqrt{\phi}B(1/2, \nu/2)}\left[\nu + \frac{(y_l - \mu_l)^2}{\phi_l}\right]^{-\frac{\nu+1}{2}}, y_l \in \mathbb{R},$$

para ν conhecido, $\beta \in \mathbb{R}^p$ e $\phi_l = h(z_l^\top \gamma)$, com $\phi_l > 0$. O logaritmo da função de verossimilhança é dado por (2.2) em que $t(z_l) = \log\left(\frac{\nu^{\frac{\nu}{2}}}{B(1/2,\nu/2)}\left[\nu + z_l^2\right]^{\frac{-(\nu+1)}{2}}\right)$, sendo $g(z_l)$, s dadas na Tabela 2.1 e $B(1/2, \nu/2)$ é a função Beta, deste modo as quatro primeiras derivadas de

t com relação à z_l são

$$t^{(1)}_{(Z_l)} = 2zs,\ t^{(2)}_{(Z_l)} = 2s + 4\frac{8z^2s^2}{\nu+1},$$
$$t^{(3)}_{(Z_l)} = \frac{24}{\nu+1}zs^2 + \frac{64}{\nu+1}z^3s^3,\ e\ t^{(4)}_{(Z_l)} = \frac{24}{\nu+1}s^2 + \frac{384}{(\nu+1)^2}z^2s^3 + \frac{768}{(\nu+1)^3}z^4s^4.$$

Usando a notação $\delta_{(a,b,c,d,e)}$ dada na equação (2.6), temos que

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

sendo que obtivemos o resultado de $\delta_{(2,1,0,0,4)} = 125/77$, para $\nu = 4$.

B.4 Distribuição t de Student generalizada

Seja $y_l \sim t\left(y_l; \mu_l, \sqrt{\phi_l}, r, w\right)$ com função densidade descrita como

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

em que r, w são conhecidos e positivos, $\beta \in \mathbb{R}^p$ e $\phi_l = h(z_l^\top \gamma)$, com $\phi_l > 0$. O logaritmo da função de verossimilhança é dado por (2.2)

com $t(Z_l) = \log\left(\frac{w^{\frac{r}{2}}}{B(1/2,r/2)}\left[w + z_l^2\right]^{\frac{-(r+1)}{2}}\right)$, sendo que $g(Z_l)$ e s dadas na Tabela 2.1.

Deste modo as quatro primeiras derivadas de t com relação à z_l são

$$
\begin{aligned}
&t^{(1)}_{(Z_l)} = 2zs, \quad t^{(2)}_{(Z_l)} = 2s + \frac{8z^2s^2}{r+1},\\
&t^{(3)}_{(Z_l)} = \frac{24}{r+1}zs^2 + \frac{64}{r+1}z^3s^3, \; e \;\; t^{(4)}_{(Z_l)} = \frac{24}{r+1}s^2 + \frac{384}{(r+1)^2}z^2s^3 + \frac{768}{(r+1)^3}z^4s^4.
\end{aligned}
$$

Usando a notação $\delta's$, dada na equação (2.6), temos que

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

B.5 Distribuição Logística Tipo I

Seja $y_l \sim LI\,(\mu_l, \phi_l)$ com funçao de densidade escrita como

$$
\pi\,(y_l, \mu_l, \phi_l) = \frac{c}{\sqrt{\phi_l}} \frac{\exp\left\{-\frac{(y_l-\mu_l)^2}{\phi_l}\right\}}{\left(1+\exp\left\{-\frac{(y_l-\mu_l)^2}{\phi_l}\right\}\right)^2}
$$

em que $c \approx 1,484300029$, $\beta \in \mathbb{R}^p$ e $\phi_l = h(z_l^\top \gamma)$, com $\phi_l > 0$. O logaritmo da função de verossimilhança é dado por (2.2) com $t(z_l) = \log\left(\frac{e^{-z_l^2}}{1+e^{-z_l^2}}\right)$, sendo que $g(z_l)$ e s são dadas na Tabela 2.1. Deste modo as quatro primeiras derivadas de t com relação a z_l são

$$t_{(Z_l)}^{(1)} = 2zs,\ t_{(Z_l)}^{(2)} = 2s + 2z^2(s^2-1),$$
$$t_{(Z_l)}^{(3)} = 6z(s^2-1) + 4z^3s(s^2-1),\ e\ t_{(Z_l)}^{(4)} = 6(s^2-1) + 24sz^2(s^2-1) + 4z^4(s^2-1)(3s^2-1).$$

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

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

com $w = 1 - \frac{e^{z^2}}{(1+e^{-z^2})^2}$. Assim, os $\delta's$ para a distribuição logística tipo I, são dados por

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

B.6 Distribuição Logística Tipo II

Seja $y_l \sim LII(\mu_l, \phi_l)$ com função de densidade

$$\pi(y_l,\mu_l,\phi_l) = \frac{1}{\sqrt{\phi_l}} \frac{\exp\left\{-\frac{(y_l-\mu_l)^2}{\phi_l}\right\}}{\left(1+\exp\left\{-\frac{(y_l-\mu_l)^2}{\phi_l}\right\}\right)^2},$$

em que $\beta \in \mathbb{R}^p$ e $\phi_l = h(z_l^\top \gamma)$, com $\phi_l > 0$. O logaritmo da função de verossimilhança é dado por (2.2), com $t(z_l) = \log\left(\frac{e^{z_l}}{\left(1+e^{-z_l^2}\right)^2}\right)$, sendo que $g(z_l)$ e s dados na Tabela 2.1. Deste modo as quatro primeiras derivadas de t com relação a z_l são

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

Assim, os $\delta' s$ são dados por

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

B.7 Exponencial Potência

Seja $y_l \sim EP(\mu_l, \phi_l, k)$ com função de densidade

$$\pi(y_l, \mu_l, \phi_l) = \frac{C(k)}{\sqrt{\phi_l}} \exp\left\{-\frac{1}{2}\left[\left|\frac{(y_l-\mu_l)^2}{\phi_l}\right|\right]^{\frac{1}{1+k}}\right\}, y_l \in \mathbb{R},$$

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

$$\ell(\beta, \gamma) = n \log C(k) - \frac{1}{2}\sum_{l=1}^{n} \log \phi_l + \sum_{l=1}^{n} t(z_l),$$

com $t(z_l) = -\frac{1}{2}|z_l|^{2/(1+k)}$, sendo que $g(z_l)$ e s são dados na Tabela 2.1. Deste modo as quatro primeiras derivadas de t com relação a z_l são

$$
\begin{aligned}
t^{(1)}_{(Z_l)} &= -\frac{1}{(1+k)}\,|z_l|^{\frac{1-k}{1+K}}\, sinal(z_l), \;\; se \;\; -1 < k \leq 1,\\
t^{(2)}_{(Z_l)} &= -\frac{1-k}{(1+k)^2}\,|z_l|^{-\frac{2k}{1+K}}, \;\; se \;\; k > 0,\\
t^{(3)}_{(Z_l)} &= \frac{2(1-k)k}{(1+k)^3}\,|z_l|^{-\frac{1+3k}{1+K}}\, sinal(z_l), \;\; se \;\; k < -1/3,\\
e \;\; t^{(4)}_{(Z_l)} &= -\frac{2(1-k)(1+3k)k}{(1+k)^4}\,|z_l|^{-\frac{2(1+2k)}{1+K}}, \;\; se \;\; k < -1/2.
\end{aligned}
$$

Assim, os $\delta' s$ são dados por

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

Referências Bibliográficas

Atkinson, A.C. (1985). Plots, Transformation and Regression. Oxford: Clarendom Press.

Berkane e Bentler (1986). *Moments of elliptical distributed random variates*. Statistics and Probability Letters, 4, 333–335.

Brito, C.C.R. (2007). *Correção de Bartlett em modelos não–lineares simétricos heteroscedásticos*. 184p. Dissertação de Mestrado em Estatística. Universidade Federal de Pernambuco.

Cambanis, S.,Huang, S. e Simons, G. (1981). On the theory of elliptically countoured distributions. *International Statistical Review*. **49**. 67–74.

Chmielewski, M.A. (1981). Elliptically symmetric distributions: a review and bibliography. *International Statistical Review*, **49**, 67–74.

Cook, R. D. and Weisberg, S. (1982). *Residuals e Influence in Regression*. New York: Chapman & Hall.

Cordeiro, G.M. (2004). Corrected LR tests in symmetric nonlinear regression models. *Journal of Statistical Computation and Simulation*. **74**, 8, 609–620.

Cordeiro, G.M.; Cysneiros, A.H.M.A. and Cysneiros, F.J.A. (2008). Bias-Corrected Maximum Likelihood Estimators in Nonlinear Heteroscedastic Models. Communications in Statistics. Theory and Methods, **87**, p. 1-26, 2009.

Cordeiro, G.M. and Ferrari, S.L.P. and Cysneiros, A.H.M.A. (1998). A formula to improve score test statistics. Journal of Statistical Computation and Simulation, Estados Unidos, **61**, p. 123-136.

Cordeiro, G.M. Ferrari, S.L.P. and Paula, G.M. (1993). Improved score tests for generalized linear models. Journal of the Royal Statistical Society, Series B, Inglaterra, **55**, p. 661-674.

Cordeiro, G.M., Ferrari, S.L.P., Uribe–Opazo, M.A. and Vasconcellos, K.L.P. (2000). Corrected maximum likelihood estimation in a class of symmetric nonlinear regression models. *Statistics and Probability Letters*, **46**, 317-328.

Cox, D.R. and Hinkley, D.V. (1974). *Theoretical Statistics*. London: Chapman and Hall.

Cribari–Neto, F. e Cordeiro, G.M. (1996). On Bartlett and Bartlett–type corrections. *Econometric Reviews*, **15**, 339–367.

Cribari–Neto, F. e Ferrari, S.L.P. (1995a). An improved Lagrange multiplier test for heteroskedasticity. *Comunications in Ststistics-Computation and Simulation*, **24**, 31–44.

Cribari–Neto, F. e Ferrari, S.L.P. (1995b). Bartlett–corrected tests for heteroskedastic linear models. *Economics Letters*, **48**, 113–118.

Cribari–Neto, F. e Zarkos, S.G. (1999). Bootstrap methods for heteroskedastics regression models: evidence on estimation and testing. *Economics Reviews*, **18**, 211–228.

Cysneiros, A.H.M.A. (2004). *Refinamentos para testes de hipóteses em modelos lineares e não–lineares heteroscedásticos*. 149p. Tese (Doutorado). Instituto de Matemática e Estatística - Universidade de São Paulo.

Cysneiros, A.H.M.A. e Cordeiro, G.M. (2002). On improving the χ^2 approximation of score tests in location-scale nonlinear models. *Communications in Statistics-Theory and Methods*, 31, **10**, 1709-1732.

Cysneiros, F.J.A., Paula, G.A. and Galea, M. (2005). *Modelos Simétricos Aplicados*. ABE: São Paulo - IX Escola de Modelos de Regressão, 100p.

Cysneiros, A. H. M. A., Rodrigues, K.S.P., Cordeiro, G.M e Ferrari, S.L.P. (2008). Three-Bartlett-Type Correction for Score Statistics in Symmetric Nonlinear Regression Models. *Statistical Papers*, **00**, p. 1-13.

Cysneiros, F.J.A., Cordeiro, G.M. e Cysneiros, A.H.M.A. (2009). Bias–corrected maximum likelihood estimators in nonlinear heteroscedastic symmetric regression models. *Communications in Statistics - Theory and Methods*.

Doornik, J.A. (2001). *Ox: An Object-oriented Matrix Programming Language* (Timberlake Consultants, London, 4th ed., and Oxcountford: `http://www.doornik.com`).

Fang, K.T., Kotz, S. and Ng, K.W. (1990). *Symmetric Multivariate and Related Distributions*. London: Chapman & Hall.

Fang, K.T. and Anderson, T.W. (1990). *Statistical Inference in Elliptical Contoured and Related Distributions*. New York: Allerton Press.

Ferrari, S.L.P., Cysneiros, A.H.M.A., Cribari-Neto, F. (2004). An improved test for heteroskedasticity using adjusted modified profile likelihood inference. *Journal of Statistical Planning and Inferences*, **57**, 353–361.

Ferrari, S. L. P., CORDEIRO, G. M., (1994). Matrix formulae for computing improved score tests. *Journal of Statistical Computation and Simulation*, Estados Unidos, **49**, 195–206.

Ferrari, S. L. P. e Arellano–Valle, R. B. . *Modified likelihood ratio and score tests In regression models using the t distribution*. Brazilian Journal of Probability and Statistics, v. 10, p. 15-33, 1996.

Fioresi, D. B., (2000). *Correção de Bartlett e tipo*$_{Bartlettemmodelosnormaishete}-$ $roscedsticos.184\ DissertaodeMestradoemEstatstica.UniversidadedeSoPaulo.$

Harris, P. (1985). An asymptotic expansion for the null distribution of the efficient score statistic. *Biometrika*, **72**, 653–659.

Johnson, N.L., Kotz, S. e Balakrishna, N. (1995). *Continuous Univariate Distributions*. Vol. 2, 2nd Edition, Wiley, New York.

Kelker, D. (1970). Distribution theory of spherical distributions and a location-escala parameter generalization. *Sankhya A*, **32**, 419-430. 923–927.

Muirhead, R. (1980). The effects of symmetric distributions on some standard procedures involving correlation coefficients. *Multivariate Statistical Analysis* (Ed. R.P Gupta) North–Holland, pp.143–159. New York: Springer-Verlag.

Rao, C.R. (1947). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society*, **44**, 50–57.

Serfling, R.J. (1980). Approximation Theorems of Mathematical Statistics. New York: John Wiley.

Verbyla, A.P. (1993). Modelling variance heterogeneity: residual maximum likelihood and diagnostics. *Journal of the Royal Statistical Society*, B **55**, 493-508.

Uribe–Opazo, M.A. (1997). Aperfeiçoamento de Testes Estatísticos em Várias Famílias de Distribuições. Tese de Doutorado. Departamento de Estatística. Universidade de São Paulo, Brasil.

Uribe–Opazo, M.A., Ferrrari, S.L.P. and Cordeiro, G.M. (2008). Corrected score tests in symmetric linear regression models. *Communications in Statistics – Theory and Methods*, **37**, 1–16.